FACTORS THAT AFFECT GENERALIZATION OF ADAPTATION

by

DAE-YONG LEE

A DISSERTATION

Presented to the Department of Linguistics
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

December 2022

DISSERTATION APPROVAL PAGE

Student: Dae-yong Lee

Title: Factors that Affect Generalization of Adaptation

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Linguistics by:

Melissa M. Baese-Berk	Chairperson
Melissa Redford	Core Member
Vsevolod Kapatsinski	Core Member
Kaori Idemaru	Institutional Representative

and

Krista Chronister	Vice Provost for Graduate Studies

Original approval signatures are on file with the University of Oregon Division of Graduate Studies.

Degree awarded December 2022

DISSERTATION ABSTRACT

Dae-yong Lee

Doctor of Philosophy

Department of Linguistics

December 2022

Title: Factors that affect generalization of adaptation

As there is a growing population of non-native speakers worldwide, facilitating communication involving native and non-native speakers has become increasingly important. While one way to help communication involving native and non-native speakers is to help non-native speakers improve proficiency in their target language, another way is to help native listeners better understand non-native speech. Specifically, while it may be initially difficult for native listeners to understand non-native speech, the listeners may become better at this skill after short training sessions (i.e., adaptation) and they may better understand novel non-native speakers (i.e., generalization). However, it is not well-understood how native listeners adapt and generalize to a novel speaker. This dissertation investigates how speaker and listener characteristics affect generalization to a novel speaker. Specifically, we examine how acoustic characteristics and talker information interact in generalization of adaptation, how accentedness of non-native speech affects generalization to a novel speaker, and how listeners' linguistic experience affects generalization of adaptation.

The results suggest that acoustic similarity between speakers may help generalization and that listeners' reliance on talker information is down-weighted, as long as speakers that listeners are trained with and tested with have similar acoustic characteristics. Furthermore, the results show that exposure to more accented non-native speech disrupts generalization of adaptation

compared to exposure to less accented non-native speech, suggesting that having exposure to non-native speakers does not always help generalization. The results also show that having extended linguistic experience with non-native speakers may disrupt generalization to a novel non-native speaker.

The results of the present study have implications for how speaker- and listener-related factors affect generalization of adaptation. Specifically, we suggest that, at least in the early stages of learning, generalization of adaptation is constrained by acoustic similarity and that generalization to a non-native speaker utilizes mechanisms that are general to speech perception, rather than specific to this type of adaptation. We suggest that exposure to non-native accented speech that is too different from the speech that listeners are familiar with may disrupt generalization. Further, we suggest that the representation of non-native accents becomes less malleable with extended linguistic experience.

CURRICULUM VITAE

NAME OF AUTHOR: Dae-yong Lee

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
Hankuk University of Foreign Studies, Seoul

DEGREES AWARDED:

Doctor of Philosophy, Linguistics, 2022, University of Oregon
Master of Arts, Linguistics, 2016, Hankuk University of Foreign Studies
Bachelor of Arts, Linguistics, 2012, Hankuk University of Foreign Studies

AREAS OF SPECIAL INTEREST:

Speech Perception
Speech Production
Phonetics
Psycholinguistics

PROFESSIONAL EXPERIENCE:

Graduate Employee (teaching assistant), Department of Linguistics, University of Oregon, 2018-2021
Graduate Employee (teaching assistant), Department of Linguistics, University of Oregon, 2016-2017

GRANTS, AWARDS, AND HONORS:

Fulbright Graduate Study Award, 2016-2018

PUBLICATIONS:

Baese-Berk, M. M., Drake, S., Foster, K., Lee, D., Staggs, C., & Wright, J. M. (2021). Lexical diversity, lexical sophistication, and predictability for speech in multiple listening conditions. *Frontiers in psychology*, *12*, 661415.

Lee, D., & Baese-Berk, M. M. (2020). The maintenance of clear speech in naturalistic conversations. *The Journal of the Acoustical Society of America*, *147*(5), 3702-3711.

Lee, D., & Baese-Berk, M. M. (2021). Non-native English listeners' adaptation to native English speakers. *JASA Express Letters*, *1*(10), 105201.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisor, Melissa Baese-Berk, for the guidance and support that she has provided me throughout the doctoral program. She is a researcher that I admire, and she has pushed me to better myself as well. I really appreciate her guidance, and I hope to become a successful researcher like her.

I am also deeply grateful to my dissertation committee members Lisa Redford, Volya Kapatsinski, and Kaori Idemaru for their invaluable feedback and suggestions on this dissertation. They really helped me refine my arguments.

In addition to my advisor and dissertation committee members, I would like to thank my former professors in the Department of English Linguistics at Hankuk University of Foreign Studies in Seoul, Korea: Professors Sookyung Cho, Hohsung Choe, Sung-Hoon Hong, Jee Eun Kim, Ji Hyon Kim, Kwang-sup Kim, Yookang Kim, Iksoo Kwon, Jeong-Woon Park, Seongha Rhee, and Kyung-Hee Suh for helping me develop as a linguist. I would especially like to thank my MA thesis advisor Tae-Yeoub Jang who inspired me to pursue a career in phonetics.

I would further like to express my appreciation to the members of the Speech Perception and Production Lab at the University of Oregon. Attending lab meetings was always exciting, and the collaboration really helped improve the quality of my work. In addition, I am grateful to Professors Jongbum Ha, Haeyeon Kim, and Kwan Young Oh for providing support and advice when they were visiting scholars in Eugene.

I would also like to thank my fellow graduate students Charlie Farrington, Kurtis Foster, Xuan Guan, Jeff Kallay, Misaki Kato, Marie Pons, Cece Staggs, Matt Stave, Allison Taylor-Adams, Amos Teo, and Jonathan Wright for making my life at the University of Oregon special.

It was enjoyable spending time together, and I also learned a lot. Special thanks to my great friend Kaylynn Gunter. I am grateful for everything. To my non-linguist friends, Yujung Kang and Seyon Jo, thank you for providing different insights and perspectives on life! I also thank Yeonjoo Lee for her companionship and always having faith in me.

I am grateful to the Fulbright Foreign Student Program for providing financial and administrative support during my initial years at the University of Oregon. Receiving the Fulbright scholarship allowed me to focus on my research and helped me complete the doctoral program successfully.

Finally, I would like to thank my parents, Jong Kun Lee and Hyunhee Seo, and my brother, Soo-Yong Lee, for encouraging and emotionally supporting me. I rarely express how thankful I am, but I am fortunate to have such a wonderful family. It would not have been possible to complete this work without their love.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

I. INTRODUCTION

Communication involving native (i.e., speakers who have learned and used a language from birth) and non-native (i.e., speakers who have not acquired a target language as their first language) English speakers can be challenging. Previous studies demonstrate that native English listeners have difficulty understanding non-native English speech (Munro & Derwing, 1995; Bent & Bradlow, 2003) and non-native English listeners have difficulty understanding native English speech (Bent & Bradlow, 2003). These difficulties can be a challenge in the United States as there is a growing population of non-native English speakers. According to a United States Census Bureau report, over 60 million people speak a language other than English at home. Among the 60 million people, around 25 million people report that they are not able to speak English "very well" (U.S. Census Bureau, 2015). Therefore, it is necessary to investigate factors that could facilitate communication involving native and non-native English speakers.

There are multiple ways to facilitate communication between native and non-native speakers. One way is to train non-native English speakers to acquire phonetic categories that do not exist in their native languages and improve their production and perception of these non-native phonetic categories, with the goal of being more comprehensible or intelligible to a listener and understanding the speech they encounter. Previous studies have shown that non-native English speakers can improve perception and production of their non-native language with training (e.g., Logan, Lively, & Pisoni, 1991; Lively, Logan, & Pisoni, 1993; Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999; Wang, Spence, Jongman, & Sereno 1999). For example, non-native English speakers can learn novel phonetic categories that do not exist in their language with training in a laboratory (Logan, Lively, & Pisoni, 1991) and retain these newly learned phonetic categories in the long term (i.e.,

3 and 6 months; Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994). However, it is important to note that training naïve listeners to perceive and produce novel categories is extraordinarily time consuming, and still often results in perceptions of talkers as being heavily accented, and not comprehensible or intelligible to a listener.

Another way to improve communication between native and non-native speakers is to train listeners so that they can better understand non-native speakers. Native English listeners often have difficulty understanding unfamiliar speech, including non-native English speech (Munro & Derwing, 1995; Bent & Bradlow, 2003) but they are able to adapt rapidly to unfamiliar speech with exposure to the unfamiliar speech (Bradlow & Bent, 2008; Clarke & Garrett, 2004; Sidaras, Alexander, & Nygaard, 2009; Xie et al., 2018). That is, native listeners demonstrate higher accuracy transcribing non-native speech (Bradlow & Bent, 2008; Sidaras, Alexander, & Nygaard, 2009) and respond faster in cross-modal matching tasks (Clarke & Garrett, 2004; Xie et al., 2018) after having exposure to non-native speech. Further, after adapting to non-native speakers, listeners are able to generalize this adaptation to novel speakers from the same (Bradlow & Bent, 2008; Sidaras, Alexander, & Nygaard, 2009; Xie et al., 2018) and different language backgrounds (Baese-Berk, Bradlow, & Wright, 2013), depending on their exposure.

Generalization to novel speakers is a crucial piece in facilitating communication involving native and non-native speakers. Specifically, as there is a growing population of non-native English speakers worldwide, it is unlikely that native English speakers communicate with a single non-native speaker. Rather, speakers will likely have conversations with multiple non-native speakers and encounter non-native speakers that they are not familiar with. Therefore, understanding underlying mechanisms of generalization of adaptation to novel speakers would

help communication involving native and non-native English speakers. In order to better understand underlying mechanisms of generalization of adaptation to a novel non-native English speaker, this dissertation explores how characteristics of non-native English speakers and native English listeners affect generalization of adaptation to novel non-native English speakers.

In this introduction, we provide an overview of the relevant literature and of the research questions investigated in the dissertation. Specifically, we discuss previous literature examining how variable speech production is and how listeners successfully understand speech despite its variability (Section 1.1). Next, we review studies on how listeners adapt to types of speech that may be challenging for listeners to understand (e.g., non-native speech; Section 1.2). Finally, we review studies on how listeners generalize their adaptation to novel non-native speakers (Section 1.3).

1.1. Listeners' perception of unfamiliar speech

Speakers produce speech in a variable manner. When speakers produce a single sentence, different speakers have different ways of producing the sentence and even when one speaker produces a sentence twice, the speaker can produce the sentence differently. Speakers may vary in the way they speak depending on a variety of factors, including their geographical origin, age, and gender. For example, in one study, researchers show that speakers from Wisconsin tend to speak faster than speakers from North Carolina, individuals in their 40s tend to speak faster than individuals in younger and older age groups, and men tend to speak faster than women (Jacewicz, Fox, & Wei, 2010). Further, speakers may change the way they speak within a single conversation. For example, native English speakers' speech become less intelligible over the

course of conversations with both native and non-native English interlocutors (e.g., Lee & Baese-Berk, 2020).

Further, speakers may change the way they speak depending on the interlocutor. Native English speakers produce the same sentences more clearly (i.e., easier to understand) when they are asked to speak as if they are speaking to a listener who may have difficulty understanding them, such as a hearing-impaired or non-native English listener than when they are asked to speak as if they are talking to a close friend (e.g., Picheny et al., 1985, 1986; Krause & Braida, 2002, 2004; Biersack et al., 2005; Maniwa et al., 2008, 2009). Further, they do this in naturalistic conversations with listeners who have difficulty understanding them (Lee & Baese-Berk, 2020). These findings suggest that speech production involves both between-speaker and within-speaker variability. This variability that lies within and between speakers may pose a challenge for listeners. That is, since every speaker differs in the way they speak, listeners have to adapt to each novel speaker in order to successfully understand speakers.

However, speech perception involves flexible systems that are capable of processing variable speech that is produced by different speakers. That is, even though speech is variable, listeners rarely experience communicative failure in conversations with familiar or novel native speakers and listeners can easily adapt and understand speech that they have not encountered before. Specifically, listeners become better at understanding speakers as they have more exposure from those speakers. For example, listeners are better at understanding novel words that are produced by familiar speakers than novel words that are produced by unfamiliar speakers (e.g., Nygaard, Sommers, & Pisoni, 1994; Nygaard & Pisoni, 1998). Further, listeners are better at recognizing words when the words are spoken by the same speaker in the training and test

than when the words are spoken by different speakers (Bradlow, Nygaard, & Pisoni, 1999) suggesting that listeners quickly adapt to characteristics of speakers.

Even when listeners do have difficulty understanding speech, they are able to better understand the speech as they have more exposure with speakers. For example, Bradlow & Pisoni (1999) demonstrate that listeners tend to have difficulty understanding "hard" words (i.e., words that are not used frequently and have many phonologically similar words) than "easy" words (i.e., words that are used more frequently than hard words and have few phonologically similar words) and have difficulty understanding words that are produced faster than words that are produced slower. However, listeners become better at understanding hard words and words that are produced with fast speaking rates as listeners get more exposure from specific speakers (Best et al., 2015; Maye, Aslin, & Tanenhaus, 2008; Floccia, Goslin, Girard, & Konopczynski, 2006).

Another type of speech that may be difficult for listeners to understand is unfamiliar regional accents. Regional accents may initially be difficult for naïve listeners to understand because of their unfamiliar acoustic characteristics. That is, regional accents may have different segmental and suprasegmental characteristics than familiar accents for a listener. Clopper & Pisoni (2004), for example, demonstrate that South Midland, Southern, and Western speakers have a more fronted /u/ in 'suit' than New England speakers and South Midland and Western American English speakers tend to be more r-full when pronouncing 'dark' than New England speakers. While these different characteristics may initially disrupt perception of novel regional accents, listeners are able to adapt to these accents. For example, Australian English listeners become better at categorizing unfamiliar British English accents (i.e., London and Yorkshire English) after hearing a short story read by a speaker of the unfamiliar British English accent

(Best et al., 2015). Similar findings are also demonstrated with an artificial accent. Specifically, after listeners hear a short story of which the vowels are acoustically modified to simulate a regional accent, listeners adapt to the artificial accented vowels (Maye, Aslin, & Tanenhaus, 2008).

Further, listeners are able to adapt to dysarthric speech. Dysarthric speech is the result of a speech-motor disorder and often involves unpredictable acoustic variation (Borrie et al., 2012). Because of its irregular acoustic signals, dysarthric speech may initially be difficult to understand for listeners that are not familiar with it. However, previous studies demonstrate that listeners become better at understanding dysarthric speech after having exposure to it in a laboratory (Borrie et al., 2012; Borrie, Lansford, & Barrett, 2017; Borrie, McAuliffe, Liss, O'Beirne, & Anderson, 2012; Borrie & Schäfer, 2015). For example, Borrie et al., (2012) demonstrate that listeners become better at transcribing hypokinetic dysarthric speech (i.e., perceptually rapid speaking rate, mono-pitch, mono-loudness, reduced syllable stress, imprecise consonants, and weak and breathy voice) after listening to it in the lab.

Similarly, previous studies demonstrate that listeners are able to adapt to special types of speech that listeners are not familiar with, including time-compressed speech, noise-vocoded speech, and computer synthesized speech (Davis, Johnsrude, Hervais-Ademan, Taylor, & McGettian, 2005; Dupoux & Green, 1997; Greenspan, Nusbaum, & Pisoni, 1988; Pallier et al., 1998). For example, listeners initially are able to transcribe fewer than 10% of words when they first hear noise-vocoded speech. However, listeners become better at transcribing this speech sentences after hearing a small number of noise-vocoded sentences (Davis, Johnsrude, Hervais-Adelman, Taylor, & McGettigan, 2005).

Taken together, these studies demonstrate that listeners tend to have difficulty understanding types of speech that they are not familiar with. However, they are able to adapt to unfamiliar speech after exposure. These results suggest that listeners' speech perception is flexible enough to adapt both to novel speakers and speech that listeners are unfamiliar with.

1.2. Adaptation to non-native speech

In addition to the types of speech reviewed above, listeners frequently have difficulty understanding non-native speech (Bent & Bradlow, 2003; Ferguson, Jongman, Sereno, & Keum, 2010; Gordon-Salant, Yeni-Komshian, & Fitzgibbons, 2010; Munro & Derwing, 1995, Munro & Derwing, 1999). For example, it takes longer for native listeners to process sentences produced by non-native speakers than sentences produced by native speakers (Munro & Derwing, 1995) and native listeners demonstrate greater number of transcription errors when transcribing sentences read by non-native speakers than sentences read by native speakers (Munro & Derwing, 1999).

While there are various factors that add to the difficulty of processing non-native speech, one of the significant factors is distinct characteristics of non-native speech that may be unfamiliar to native listeners. Non-native speech often has characteristics of speech that are different than native speech that differ as a function of language background or by target language (Flege & Eefting, 1987; Flege, 1991; Guion, Flege, Liu, & Yeni-Komshian, 2000; Kang & Guion, 2006; Laturnus, 2020; Mok & Dellwo, 2008; Oh et al, 2011; Wayland, 1997; Yang, 1996) on top of speech characteristics that vary as a function of individual speaker properties (Bradlow, Blasingame, & Lee, 2018; Bradlow, Kim, & Blasingame, 2017; Jacewicz, Fox, & Wei, 2010). These characteristics of non-native speech include both segmental and

suprasegmental characteristics. For example, non-native speech tends to be slower than native speech (Guion, Flege, Liu, & Yeni-Komshian, 2000) and stops produced by non-native English speakers have different acoustic characteristics (e.g., VOT, H1-H2, and F0) than stops produced by native English speakers (Kang & Guion, 2006). As a result of these characteristics of non-native speech, perception of non-native speech may be challenging for native listeners.

While these characteristics of non-native speech may initially disrupt perception of non-native speech, native listeners are able to adapt to non-native English speech (i.e., improve their understanding of non-native English speech) with training. Previous studies have shown that listeners are able to adapt to non-native English speakers within a very short exposure (Bradlow & Bent, 2008; Clarke & Garrett, 2004; Xie et al., 2018). Specifically, native listeners' processing time of non-native speech is reduced after brief exposure to a non-native speaker. For example, Clarke & Garrett (2004) demonstrate that native listeners' response time in a cross-modal matching task is greatly reduced within a minute of exposure to non-native speech.

Further, listeners demonstrate better performance of transcribing non-native speech after having exposure to non-native speech (Bradlow & Bent, 2008; Gass & Varonis, 1984; Gordon-Salant, Yeni-Komshian, Fitzgibbons, & Schurman, 2010; Mitterer & McQueen, 2009; Pinet, Iverson, & Evans, 2011; Sidaras, Alexander, & Nygaard, 2009). For example, Bradlow & Bent (2008) demonstrate that listeners become better at transcribing sentences read by a non-native speaker within two days of exposure to the non-native speaker. Similarly, Sidaras, Alexander, & Nygaard (2009) show that listeners demonstrate better performance in transcribing novel words produced by a non-native speaker after a short training period. These studies suggest that speech perception systems are flexible enough to rapidly adapt to non-native speech.

1.3. Generalization of adaptation to a novel non-native speaker

One important aspect of adaptation to non-native speech is that after listeners adapt to a non-native accent, listeners are able to successfully understand novel non-native speakers from the same language background in certain conditions (i.e., "generalization"; Bradlow & Bent, 2008; Xie et al., 2018). Generalization of adaptation to novel speakers is important for communication involving native and non-native speakers as native speakers are likely to communicate with more than a single non-native speaker. That is, training native speakers to adapt to one non-native speaker will facilitate native speakers' communication with the non-native speaker but it does not necessarily help communication with novel non-native speakers outside of the lab. However, if native speakers are trained to adapt to non-native speakers and generalize their adaptation to novel non-native speakers, they will likely have more successful communication with non-native speakers than speakers who are trained to adapt to a single speaker. Thus, understanding the underlying mechanisms of generalization to novel speakers would greatly help native speakers understand non-native speech in real-world communications.

Understanding how listeners generalize their adaptation to a novel speaker also helps learning the underlying mechanisms of speech perception. A significant challenge in speech perception is the variation that listeners encounter. As discussed in Section 1.1., different speakers have different speech characteristics caused by multiple factors (e.g., age, region, gender, native language) and even the same speaker often produces the same sentence differently under different circumstances. This variability may initially pose a challenge for speech perception. For example, Mullennix, Pisoni, & Martin (1989) demonstrate that high variability in terms of number of talkers makes word identification more challenging than low variability. That is, it is more difficult to identify words when the words are presented by a bigger number of

speakers than when the words are presented by a smaller number of speakers demonstrating the difficulty that high variability poses to listeners.

However, listeners are still able to have successful communication with other speakers despite this variability. In order to understand how listeners overcome this challenge, one line of research suggests that while speech is variable, it includes invariable cues that help listeners distinguish one sound from another (e.g., Blumstein & Stevens, 1979; Blumstein & Stevens, 1980; Stevens & Blumstein, 1978; Wade, Wayland, & Wong, 2000). Indeed, Wade, Wayland, & Wong (2000) demonstrate that certain acoustic characteristics of fricatives (i.e., spectral and amplitude properties of fricatives) serve as important factors for distinguishing one fricative to another.

Although the studies discussed above demonstrate that invariable cues exist in speech and posit that these invariable cues may help distinguish different phonetic categories, it is unlikely that listeners rely solely on these cues in speech perception. That is, while acoustic characteristics of phonetic categories are indeed important factors for speech perception, a number of studies suggest that speech-external characteristics such as talker information (e.g., speaker identity or speaker background) or non-linguistic information (e.g., having a doll in the lab) also has an effect on speech perception (Hay & Drager, 2010; Hay, Nolan, & Drager, 2006; Niedzielski, 1999). Specifically, even when the acoustic characteristics of a phonetic category remain the same, listeners may perceive the phonetic categories as different sounds depending on talker background. For example, Niedzielski (1999) shows that listeners identify the same synthesized vowel differently depending on whether the listeners think the speaker is from the same region as themselves or from a different region. Therefore, in order to account for how listeners successfully understand speech despite of its variability, it is important to examine how

acoustic characteristics of speech and talker information together affect speech perception. Thus, we examine how acoustic similarity between talker and talker information affect generalization of adaptation to a novel speaker to better understand how acoustic characteristics of speakers and talker information interact in speech perception.

Further, the variable realization of speech is often viewed as an obstacle for speech perception. As discussed above, earlier studies that examine the underlying mechanisms of speech perception often consider variability as a factor that disrupts speech perception (e.g., Blumstein & Stevens, 1979; Blumstein & Stevens, 1980; Stevens & Blumstein, 1978). Specifically, the assumption is that listeners focus on constant acoustic cues and discard variable information for speech perception to be successful. This approach is understandable as numerous studies show that variability disrupts speech perception (e.g., Mullennix, Pisoni, & Martin, 1989; Mullennix & Pisoni, 1990). However, it is not the case that variability is necessarily detrimental for speech perception and exposure to variability may facilitate speech perception (e.g., Baese-Berk, Bradlow, & Wright, 2013; Bradlow & Bent, 2008; Lively, Logan, & Pisoni, 1993; Sidaras, Nygaard, & Alexander, 2009; Tzeng, Alexander, Sidaras, & Nygaard, 2016; Xie et al., 2018). Specifically, exposure to multiple talkers may help listeners better understand novel non-native speakers. For example, Bradlow & Bent (2008) show that listeners who have exposure to multiple Mandarin learners of English are better at transcribing a novel Mandarin learner of English than listeners who have exposure to a single Mandarin learner of English and listeners who do not have exposure to Mandarin learners of English. Further, Baese-Berk, Bradlow, & Wright (2013) demonstrate that listeners who have exposure to non-native speakers from multiple language backgrounds are better at transcribing a non-native English speaker from a novel language background than listeners who do not have exposure to non-native English

speakers from multiple language backgrounds. These studies suggests that exposure to variable types of speech is not always detrimental for speech perception. That is, it is possible that there is learnable structure in variability and exposure to variability helps listeners learn the structure. For example, as Baese-Berk, Bradlow, & Wright (2013) suggest, non-native speakers may share common characteristics that are likely caused by speaking a non-native language such as speaking slower than native speakers or demonstrating less reduction when producing unstressed vowels than native speakers.

Although previous studies suggest that variability is helpful for speech perception, it is much less understood how variability is helpful for speech perception. Specifically, it is unclear what type of variability is helpful for speech perception. For example, as discussed above, Bradlow & Bent (2008) demonstrate that exposure to multiple Mandarin learners of English help better understand a novel Mandarin learner of English. However, exposure to a single Mandarin learner of English does not help better understand a novel Mandarin learner of English, suggesting that exposure to multiple speakers plays an important factor for generalization to a novel non-native speaker. On the other hand, Weil (2001) demonstrates that exposure to a single speaker facilitates generalization to a novel non-native speaker from the same language background. Specifically, the difference in post-test intelligibility scores between listeners who have training with a Marathi learner of English and listeners who do not have training with a Marathi learner of English is the same whether the listeners are trained and tested with the same Marathi learner of English or trained with one Marathi learner of English and tested with a novel Marathi learner of English. As the study does not include a condition in which listeners are trained with multiple Marathi learners of English, it is difficult to directly compare the results of Weil (2001) to that of Bradlow & Bent (2008), but these results suggest that training listeners

with a single non-native speaker facilitates generalization to a novel speaker from the same language background. Bradlow & Bent (2008) suggest that the contrary results may be caused by the type and amount of training given to the listeners. Specifically, while Bradlow & Bent (2008) train listeners with sentences over the course of two training sessions, Weil (2001) trains listeners with words, sentences, and passages over the course of three training sessions. While it is possible that different types of variability have different effects on speech perception, it is unclear when variability becomes beneficial for speech perception. As previous studies demonstrate that variability plays a significant role in adaptation and its generalization, it is important to examine how variability may or may not be helpful for speech perception.

There is some evidence suggesting that listeners may benefit from exposure to variability by highlighting the common characteristics shared by non-native speakers. That is, it is possible that systematic variability exists in non-native speech as suggested in previous studies (Baese-Berk, Bradlow, & Wright, 2013; Laturnus, 2018). For example, previous studies show that non-native speech tends to have distinct characteristics from native speech (Baker, Trofimovich, Flege, Mack, & Halter, 2008; Flege, 1987; Oh et al., 2011) and that non-native speakers share common segmental and suprasegmental characteristics (Guion, Flege, Liu, & Yeni-Komshian, 2000; Laturnus, 2020; Munro & Derwing, 1995; Toivola, Lennes, & Aho, 2009). Specifically, Laturnus (2020) demonstrates that Farsi and Italian learners of English have significantly shorter voiced VOT (i.e., voice onset time) durations than native English speakers and Farsi, Korean, and Thai learners of English tend to produce schwas as unreduced vowels. In terms of suprasegmental features, Guion et al., (2000) demonstrate that non-native English speakers' speech rate is often slower than native speaker' speech. These studies on non-native speakers' production of speech in their non-native language suggest that non-native speakers may have

common characteristics, regardless of the talkers' language background, that stem from using a non-native language.

On top of characteristics that non-native speakers share regardless of their language backgrounds, non-native speakers from the same language background may also share common characteristics with one another. Specifically, previous studies suggest that non-native speakers who share the same native language may have common characteristics that are transferred from or are a result of the speakers' native language and its relationship to the target language (Bent & Bradlow, 2003; Hayes-Harb, Smith, Bent, & Bradlow, 2008; Major, Fitzmaurice, Bunta, & Balasubramanian, 2002). For example, Bent & Bradlow (2003) demonstrate that speakers' intelligibility (i.e., percent of keywords correctly transcribed) is the same when listeners transcribe a proficient non-native speaker from the same language background as when the listeners transcribe a native speaker, suggesting that listeners benefit from the knowledge they have about their native speech when processing non-native speakers who share the same language background. That is, it is possible that non-native speech has common characteristics transferred from speakers' native language. If it is the case that non-native speech has common characteristics that are transferred from non-native speakers' native languages, highlighting the characteristics would help native listeners adapt and generalize to novel non-native speakers. One possible characteristic of non-native speech that may highlight the common features of non-native speech is accentedness of non-native speech. Previous studies demonstrate that more accented non-native speech deviates more from native speech than less accented non-native speech (Munro, 1993; Porretta, Kyröläinen, & Tucker, 2015). Thus, if listeners generalize to a novel speaker by learning common characteristics of non-native speech, it is possible that exposure to more accented non-native speech facilitates generalization of adaptation by

highlighting features of the accent that a learner may adapt to. On the other hand, it is also possible that exposure to more accented non-native speech disturbs generalization. Previous studies suggest that learning in difficult environments make it difficult for the learning to generalize (e.g., Ahissar & Hochstein, 2004). As more accented non-native speech is likely to have more distinct characteristics from the speech that listeners are familiar with than less accented non-native speech (Munro, 1993; Porretta, Kyröläinen, & Tucker, 2015), training with more accented non-native speakers may be more difficult than training with less accented non-native speakers and this may disrupt generalization to a novel speaker. Understanding how accentedness of non-native speech affects generalization of adaptation will provide a better understanding of what aspects of variability facilitates speech perception. Thus, the current work examines how accentedness of non-native speech affects generalization of adaptation to a novel non-native speaker.

While speaker characteristics (e.g., acoustic similarity, talker information, and accentedness) may play a significant role in speech perception, it is also important to investigate how listener characteristics affect generalization to a novel speaker. Specifically, exposure to the same amount of variability may have different effects on speech perception based on listeners' linguistic experience. There are, indeed, previous studies that demonstrate that listener characteristics have an impact on speech perception (Adank & Janse, 2010; Banks, Gowen, Munro, & Adank, 2015; Bent & Bradlow, 2003; Gordon-Salant, Yeni-Komshian, Fitzgibbons, & Schurman, 2010; Laturnus, 2018; Peelle & Wingfield, 2005). For example, listeners' age has an effect on generalization of adaptation to time-compressed speech (Adank & Janse, 2010) and listeners' cognitive abilities affect their adaptation to non-native speech (Banks, Gowen, Munro, & Adank, 2015). These studies demonstrate that speech perception is not solely driven by

speaker characteristics and that listener characteristics play a significant role. Further, Laturnus (2018) demonstrates that listeners who have greater lifetime experience with non-native English speakers are better at transcribing sentences read by non-native English speakers than listeners who have less lifetime experience highlighting the influence of listeners' linguistic experience on perception of non-native speech. Laturnus (2018) suggests that non-native speech may have some systematicity and listeners who have extended experience with non-native speakers may learn the systematicity from repeated exposure to non-native speech. However, it is unknown how different types of linguistic experience affect adaptation and generalization to a novel speaker. That is, it is not clear how exposure to different types of variability affects how listeners adapt and generalize their adaptation to a novel non-native speaker. Examining how listeners' linguistic experience affects adaptation and generalization would help better understand how exposure to different types of variability may or may not be helpful for speech perception. Thus, the current works examines the effect of linguistic experience on generalization of adaptation to a novel non-native speaker.

1.4. Current research

The goal of this dissertation is to better understand how talkers' acoustic characteristics and talker information together affect speech perception, as well as to better understand when variability is beneficial for speech perception. Specifically, in order to examine how acoustic characteristics and talker information interact in speech perception, we examine the effect of acoustic similarity between speakers and talker information on generalization of adaptation to a novel speaker. We further examine how variability may be helpful for speech perception by

examining the effect of accentedness of non-native speech and the effect of listeners' linguistic experience on generalization of adaptation.

1.4.1. Hypotheses explored in the dissertation

In this dissertation, we explore how acoustic characteristics of talkers and talker information together affect generalization of adaptation and what type of variability may be helpful for generalization to a novel speaker. Specifically, one question we ask is how acoustic similarity between speakers and talker information affect generalization to a novel speaker. One possible outcome is that both acoustic similarity between speakers and talker information affect generalization to a novel speaker, as previous studies suggest that listeners' perception of the talker has an effect on speech perception (Hay & Drager, 2010; Hay, Nolan, & Drager, 2006; Niedzielski, 1999). On the other hand, given previous studies suggesting that generalization of phonetic retuning is constrained by acoustic similarity between speakers (e.g., Reinisch & Holt, 2014; Xie & Myers, 2017), it is also possible that generalization of adaptation to a novel non-native speaker is strictly constrained by acoustic similarity between speakers and talker information (i.e., perceived talker change) does not play a significant role as long as speakers have similar acoustic characteristics.

The current work also explores what types of variability may be helpful for generalization of adaptation. Specifically, we examine the effect of accentedness of non-native speech on generalization to a novel speaker. As previous studies suggest that non-native speakers from the same language background share common characteristics of L2 speech that are transferred from their L1 (e.g., Flege, Schirru, & MacKay, 2003; Flege, Takagi, & Mann, 1995), it is possible that highlighting these characteristics facilitates generalization to a novel non-native

speaker. That is, having exposure to more accented non-native speakers may help listeners learn the common characteristics of non-native speakers and generalize to a novel speaker than having exposure to less accented non-native speakers. However, it is also possible that exposure to more accented non-native speakers disrupts generalization of adaptation. Previous studies suggest that exposure to high variability does not guarantee generalization (e.g., Perrachione, Lee, Ha, & Wong, 2011). Specifically, exposure to stimuli that are highly variable may in fact disrupt listeners from learning the characteristics of non-native speech. Similarly, more accented non-native speech may be too different than types of speech that native listeners are familiar with and this gap between more accented non-native speech and speech that listeners are familiar with may disrupt generalization of adaptation.

Further, we explore how extended exposure to variability affects speech perception. Specifically, we examine how different types of linguistic experience affect generalization of adaptation. Previous studies demonstrate that extended linguistic experience with non-native speakers helps listeners better understand a novel non-native speaker (e.g., Laturnus, 2018). That is, listeners may create speaker models after having experience with non-native speakers and use these models when communicating with a novel speaker instead of processing the speakers' speech from scratch. If this is the case, it is likely that listeners' linguistic experience affects adaptation and generalization to novel speakers. However, it is less well-understood how different types of linguistic experience affects adaptation and generalization to novel non-native speakers. Two outcomes are possible regarding the effect of linguistic experience on generalization to a novel speaker. First, it is possible that adaptation and generalization may be scaffolded with listeners' previous linguistic experience with non-native speakers. That is, it is possible that listeners who have extended linguistic experience with non-native speakers are

familiar with common characteristics of non-native speakers and this knowledge may help listeners adapt and generalize their adaptation to a novel non-native speaker. On the other hand, extended linguistic experience may in fact disrupt generalization to a novel speaker. Previous studies demonstrate that listeners who have smaller social networks show stronger perceptual learning than listeners who have bigger social networks suggesting that extended experience could be harmful for speech perception (Lev-Ari, 2017). Similarly, it is possible that listeners with extended linguistic experience are less malleable for adapting and generalizing to non-native speech that the listeners are not familiar with. Overall, exposure to variability may have different effects on speech perception for listeners who have linguistic experience with non-native speakers than listeners who have no experience with non-native speakers. By examining how listeners' linguistic experience affect generalization to a novel speaker, it is possible to have a better understanding of how listeners process non-native speakers' speech and generalize to novel speakers.

1.4.2. Structure of the dissertation

The studies in this dissertation use an intelligibility task with three set of stimuli to examine how acoustic similarity and talker information, accentedness of non-native speech, and listeners' linguistic experience affect generalization of adaptation to a novel speaker. In Chapter 2, we examine the roles of acoustic similarity between non-native English speakers and talker information in generalization to a novel non-native speaker. Specifically, we investigate how training listeners with a Korean learner of English affects native English speakers' perception of a Korean learner of English who has similar acoustic characteristics but is perceived as a different speaker than the Korean learner of English that they are trained with. By examining the

effects of acoustic similarity between talker and talker information on generalization to a novel speaker, we test the hypothesis that both acoustic characteristics and talker information play a significant role in generalization to a novel talker and better understand how these factors affect speech perception.

In Chapter 3, we examine the effect of accentedness of non-native speech on generalization to a novel speaker. That is, we examine whether being exposed to more accented non-native speakers or less accented non-native speakers facilitates generalization to a novel non-native speaker from the same language background. By examining how accentedness of non-native speech affects generalization to a novel non-native speaker, we better understand how exposure to variability facilitates or disrupts speech perception.

In Chapter 4, we investigate the effect of listeners' linguistic experience on generalization of adaptation. Specifically, we examine whether native English listeners who have extended linguistic experience with multiple non-native accents or a single non-native accent are better at generalizing their adaptation to a novel non-native speaker than listeners who do not have linguistic experience with non-native English speakers. By examining the effect of linguistic experience on generalization to a novel speaker, it is possible to better understand the types of variability that are helpful for speech perception.

In Chapter 5, we present a summary of the findings and discuss the novel contributions to the field.

II. EFFECTS OF ACOUSTIC SIMILARITY AND TALKER INFORMATION ON GENERALIZATION OF ADAPTATION

2.1. Introduction

Listeners often have difficulty understanding speech that they are not familiar with. Specifically, listeners often have difficulty understanding non-native speech. For example, Mandarin accented English speech takes longer for native English listeners to understand than native English speech and is rated as less comprehensible (Munro & Derwing, 1995). While understanding non-native speech may be initially challenging for listeners, listeners may become better at understanding non-native speech as the listeners get exposure to non-native speech (Bradlow & Bent, 2008; Clarke & Garrett, 2004; Sidaras, Alexander, & Nygaard, 2009) and generalize their adaptation to novel speakers who have the same language background as the speakers the listeners are trained with (Bradlow & Bent, 2008; Sidaras, Alexander, & Nygaard, 2009; Xie et al., 2018). While previous studies demonstrate that listeners are able to adapt and generalize their adaptation to a novel speaker the underlying mechanisms of generalization of adaptation is less understood. Specifically, it is less well-understood how talkers' acoustic characteristics and talker information interact in generalization to a novel speaker. Thus, in Experiment 1, we investigate the effect of acoustic similarity between talkers and talker information on generalization of adaptation.

2.1.1. The roles of acoustic characteristics and talker information in speech perception

When listeners encounter phonetic categories that they are unfamiliar with, they are able to retune their phonetic categories. For example, Norris, McQueen, & Cutler (2003) demonstrate that listeners categorize an ambiguous sound differently depending on the contexts the

ambiguous sound is presented with. Specifically, listeners are asked to listen to words that end with an ambiguous sound that is on a [s] – [f] continuum (i.e., the sound can be categorized as either an [s] or an [f]) and then categorize the ambiguous sound. In the training phase, one group of listeners hear [f]-final words that end with the ambiguous sound and [s]-final words that end with an [s] and another group of listeners hear [s]-final words that end with the ambiguous sounds and [f]-final words that end with an [f]. Then, when listeners are asked to categorize sounds on an [s] – [f] continuum, listeners that are trained with ambiguous [f]-final words are more likely to categorize sounds on the continuum as an [f], suggesting that listeners are able to quickly retune their phonetic categories.

After retuning phonetic categories, listeners are able to generalize to novel speakers in certain conditions (e.g., Eisner & McQueen, 2005; Kraljic & Samuel, 2016; Reinisch & Holt, 2014; Xie & Myers, 2017). Specifically, generalization of phonetic category retuning is likely constrained by acoustic similarity between phonetic categories. For example, Eisner & McQueen (2005) demonstrate that while phonetic category retuning does not initially generalize to a novel speaker, the retuning is generalized to the novel speaker if the target phonetic category (i.e., ambiguous sound on an [s] – [f] continuum) is spliced into the novel talker's speech. This result suggests that acoustic similarity between the phonetic category in training and post-test plays a significant role in generalization of phonetic category retuning and talker identity is less important. Indeed, Xie & Myers (2017) demonstrate that phonetic category retuning generalizes when the target phonetic category has similar acoustic characteristics across training and post-test (i.e., similar stop characteristics).

While acoustic characteristics of speech are important in speech perception, previous studies point out the significant role of talker information in speech perception. That is, even

when listeners listen to the same speech sound, listeners may hear the sound differently depending on who they think the speaker is (e.g., Hay & Drager, 2010; Hay, Nolan, & Drager, 2006; Niedzielski, 1999). Further, Kraljic & Samuel (2011) demonstrate that listeners do not demonstrate perceptual retuning when listeners learn that the speaker has a pen in their mouth by watching a video. That is, phonetic category retuning does not automatically occur whenever listeners have exposure to ambiguous sounds. Rather, listeners' perception of the speaker affects whether the listeners learn the ambiguous sound or not. Thus, it is important to examine how acoustic similarity between speakers and talker information interact in order to better understand how listeners adapt and generalize this adaptation to a novel speaker.

2.1.2. Current study

In the current study, we examine how acoustic similarity between talkers in the training session and post-test and talker information affect generalization to a novel Korean learner of English. The current study consists of two experiments. Experiment 1A aims to replicate previous studies (e.g., Bradlow & Bent, 2008; Weil, 2001) that examine whether training listeners with the same speaker in training and post-test is more helpful for listeners for understanding a novel speaker from the same language background than training listeners with one speaker and testing with another speaker. Specifically, in Experiment 1A, we train listeners with a single Korean learner of English and examine whether listeners generalize their adaptation to a novel non-native speaker from the same language background. Two outcomes are possible. Given the findings of previous studies that demonstrate training listeners with multiple speakers facilitate generalization to a novel speaker (e.g., Bradlow & Bent, 2008; Sidaras, Alexander, & Nygaard, 2009) and training listeners with a single speaker does not help generalization

(Bradlow & Bent, 2008), it is possible that listeners do not demonstrate generalization to a novel Korean learner of English in post-test after having exposure to a single Korean learner of English in training. On the other hand, it is also possible that listeners generalize to a novel speaker even after having exposure to one non-native speakers, as shown in Weil (2001).

Experiment 1B aims to better understand how acoustic similarity between talkers and talker information interact in generalization to a novel speaker. In Experiment 1B, we investigate whether listeners generalize their adaptation to a novel speaker if the speaker in post-test have similar acoustic characteristics as the speaker in training but is perceived as a different speaker. Specifically, we compare the intelligibility scores of listeners who are trained and tested with the same Korean learner of English and listeners who are trained with a Korean learner of English and tested with a novel Korean learner of English who has similar acoustic characteristics but have a different median F0. Three outcomes are possible. First, it is possible that listeners who are trained and tested with the same speaker demonstrate better performance in the post-test than listeners who are trained and tested with acoustically similar speakers (i.e., perceived as different speakers). Previous studies suggest that listeners' perception of the talker plays a significant role in speech perception (Hay & Drager, 2010; Kraljic & Samuel, 2011; Hay, Nolan, & Drager, 2006; Niedzielski, 1999). Similarly, it is possible that generalization of adaptation is disrupted when listeners perceive a talker change even when the speakers in training and post-test have similar acoustic characteristics (i.e., similar acoustic characteristics but different median F0).

Second, it is possible that listeners in the two conditions discussed above demonstrate similar performance in the post-test. Previous studies on the generalization of phonetic category retuning suggest that generalization is constrained by acoustic similarity between the target items (e.g., Eisner & McQueen, 2015; Kraljic & Samuel, 2016; Reinisch & Holt, 2014; Xie & Myers,

2017) and talker information is orthogonal to generalization. It is possible that generalization of phonetic category retuning has similar underlying mechanisms as generalization of adaptation. That is, listeners may generalize their adaptation to a novel speaker if the novel speaker has similar acoustic characteristics as the speaker they are trained with even if the speakers are perceived as different talkers.

Third, it is also possible that listeners who are trained and tested with speakers that are acoustically similar but perceived as different demonstrate better performance in the post-test than listeners who are trained and tested with the same speaker. That is, a talker change between training and post-test may reorient listeners' attention and paying more attention may facilitate listeners' perception of the novel talker in post-test. As previous studies suggest that an introduction of a new talker disrupts speech perception (e.g., Bradlow & Bent, 2008; Mullennix, Pisoni, & Martin, 1989), it is not likely that reorienting listeners' attention necessarily facilitates speech perception. However, it may be the case that when a novel talker has similar acoustic characteristics as the talker that listeners are trained with, paying more attention to the speaker helps speech perception.

2.2. Experiment 1A

2.2.1. Methods

2.2.1.1. Participants

75 native English speakers between 18 and 40 (32 female, 41 male, 2 did not prefer to answer) years old participated in this experiment. Participants were recruited from two different platforms. First, participants were recruited from the University of Oregon Psychology and Linguistics subject pool and they received partial course credits for their participation.

Participants recruited from the University of Oregon Psychology and Linguistics subject pool were not screened for participation. However, the target population was native English speakers with no frequent interaction with non-native English speakers. Thus, participants were not included in the data analysis if: 1) participants were non-native English speakers, 2) participants had frequent interaction with non-native English speakers in their community and at school or work, 3) participants had frequent interaction with relatives that are non-native English speakers, 4) participants learned a second language before the age of 10 or earlier, and 5) participants lived in a non-English speaking country for an extended period

Participants were also recruited from Prolific, an online data collecting platform, and were paid $7.50 for their participation. Participants who did not meet the requirements of being a native English speaker with no frequent interaction with non-native English speakers were not invited to participate in the experiment. To ensure that participants met the criteria, the participants began the experiment with questions that checked participants' eligibility. The questions included four language background questions and five language environment questions. The language background questions asked participants to check all that are true among the four statements including: 1) I grew up hearing and speaking American English since I was born, 2) I grew up hearing and speaking a non-English language since before I was 10 years old, 3) I have lived abroad in a non-English speaking place for an extended period of time in my life (longer than a vacation), and 4) I have had family members or close community members who I regularly interacted with a non-native language for an extended period of time in my life. The language environment questions asked participants all that are true among the five statements including: 1) I often hear non-native English speakers at home, 2) I often hear non-native English speakers at work, 3) I often hear non-native English speakers at school, 4) I hear non-native

English speakers in my community pretty much every day, and 5) I don't often hear non-native English speakers around me. Participants who checked only "I grew up hearing and speaking American English since I was born" in the language background questions and only "I don't often hear non-native English speakers around me" were invited to the experiment. The language background questions and language environment questions took less than a couple of minutes to answer and participants who did not meet the requirements of the study were not allowed to continue the study.

In both platforms, participants were not included in the analysis or not invited to the experiment if they did not use headphones during the experiment as an attempt to control for the listening environment in an online experimentation set-up. Further, participants were asked whether they had a history of speech or hearing disorder, and participants who had a history of speech or hearing disorder were not included in the data analysis or invited to participate in the experiment.

2.2.1.2. Materials

Items used in the present study were drawn from the Online Speech/Corpora Archive and Analysis Resource (OSCAAR) (Bradlow, n.d.), which includes Bamford-Kowal-Bench (BKB) sentence lists (Bamford & Wilson, 1979; Bench & Bamford, 1979). BKB sentences were simple English declarative sentences with 3 or 4 keywords (e.g., *The thin dog was hungry*). The sentences were read by 10 (five female and five male speakers) Korean-English bilinguals whose L1 was Korean and L2 was English. The speakers' L2 proficiency was from 34 to 77 (mean: 55.2, SD: 11.55) on the Versant English test (Pearson, 2009). Further, the speakers were born in Korea and were educated up to their undergraduate degree in Korean. The speakers' length of

residence in English speaking countries ranged from 0.1 to 3.5 years (M = 1.6 years) (Bradlow, Blasingame, & Lee, 2018).

Among the 10 Korean-English bilinguals in the OSCAAR corpus, the recordings of one male and one female speaker were used in the present study. A male and a female speaker were chosen to ensure that listeners perceive the talker switch from the training session to the post-test in one of the experiment conditions. Specifically, as described in the section 2.2.1.3., Experiment 1A has one condition in which the Korean learner of English remains the same in the training session and the post-test (i.e., Same Speaker Condition) and another condition in which the Korean learner of English changes from the training session from the post-test (i.e., Different Speaker Condition). Thus, in the Different Speaker Condition, a female Korean learner of English was presented in the training session and a male Korean learner of English was presented in the post-test to ensure that listeners perceive the talker change from the training session to the post-test.

The OSCAAR corpus also included lists of BKB sentences read by 10 (five female and five male) native English speakers. On the contrary to the Korean-English bilinguals, language background information was not available for the 10 native English speakers. The criteria used for deciding the speaker to be included was the same as in the 10 Korean-English bilinguals. That is, the male speaker with the most sentences uploaded to the corpus was used as the speaker in the training session of the control condition.

In the present study, 120 BKB sentences were used in the training session and 16 BKB sentences that were not presented in the training session were used in the post-test. The sentences were read by the female and male Korean learners of English and the male native English speaker described above. Specifically, the sentences read by the Korean learners of English were

presented in the training sessions of the Same Speaker and Different Speaker Conditions and the sentences read by the native English speaker were presented in the training session of the Control Condition. Further, the 16 sentences in the post-test were read by the male Korean learner of English in all conditions.

The sentences were leveled to a fixed root-mean-square (RMS) amplitude of 73 dB. Then, the stimuli were mixed with speech shaped noise at a signal-to-noise ratio (SNR) of -5 dB to prevent ceiling effects. This SNR was determined based on the results of a pilot experiment Specifically, the pilot experiment aimed to determine the SNR that would prevent native English listeners from scoring over 70% so that participants who are trained are able to improve beyond this baseline level of performance. The results of the pilot experiment showed that at an SNR of -5 dB, native English listeners did not show ceiling effects (mean = 44.7, SD = 38.7) Thus, all sentences in Experiment 1 were mixed with speech-shaped noise at an SNR of -5 dB. The BKB sentences used in the training session and the post-test are provided in the Appendix.

2.2.1.3. Design

Participants completed an intelligibility task in which they were asked to listen sentences and transcribe what they heard. The design of this experiment largely followed Bradlow & Bent (2008) and Baese-Berk, Bradlow, & Wright (2013). However, the design of the present experiment differed in two aspects. First, while Bradlow & Bent and Baese-Berk, Bradlow & Wright were two-day studies, the present experiment was a single day study, given restrictions from COVID-19 and to avoid attrition (e.g., Stoycheff, 2016). Second, Bradlow & Bent and Baese-Berk, Bradlow, & Wright used 160 sentences in training sessions, and in the current study, we used 120 training sentences to avoid attrition and lack of attention to the task. Unlike

previous studies in this area, the present study was conducted using an online experimentation setting due to restrictions from COVID-19. Listeners participating in an online experiment are more likely to get distracted (e.g., having people in the same room or multitasking while participating in the experiment) than listeners participating in an experiment in the lab (Clifford & Jerit, 2014). Thus, the present study closely followed the design of the previous studies but presented fewer sentences during the training session than previous experiments conducted in the lab to reduce the chance of participants being distracted. In spite of using fewer sentences, we followed a similar blocking design as the previous studies. In these studies, listeners were exposed to five repetitions of 16 sentences in the first section (blocked such that 16 sentences were heard, and then a different randomization of those same 16 sentences were heard). In the second training session, listeners were exposed to five repetitions of a different set of 16 sentences blocked in the same way as the first training session. That is, listeners heard a total of 32 unique sentences, with each sentence repeated five times (5 X 16 = 80 sentences/day; 160 sentences total). Previous studies utilized multiple repetitions because it is likely that this repetition of the target sentences is important in listeners' adaptation to non-native English speakers since listeners are more likely to have access to the lexical information of the non-native speech if they hear the same set of sentences repeatedly. That is, if a speaker hears a sentence once and they are able to understand some of the words in the sentence, they may be able to use this information to scaffold their perception the next time they hear the same sentence, even if the acoustic properties of the sentence are not identical, which could facilitate adaptation to non-native English speakers. This hypothesis is consistent with work that demonstrates that perceptual adaptation can involve integration of both acoustic and lexical information (e.g., Norris, McQueen, & Cutler, 2003).

Therefore, in the present study, the training session included six blocks of 20 sentences per block. The first three blocks (i.e., Blocks 1-3) included the same 20 sentences, presented in a random order each block. The second three blocks (i.e., blocks 4-6) included a different set of 20 sentences, also presented in a random order. That is, listeners heard a total of 40 unique sentences, with each sentence repeated three times (3 X 40 = 120; 120 sentences total). Each sentence was repeated three times due to the potential for repetition improving adaptation, as described above. Following the training session was the post-test. As in Bradlow & Bent and Baese-Berk, Bradlow, & Wright, the post-test included 16 sentences that were not presented in the training session.

Using the basic paradigm described above, three conditions were created: Same Speaker, Different Speaker, and Control. Each of the three conditions included a training session and a post-test described above. The post-test was the same in all three conditions. That is, at test, participants were asked to transcribe 16 sentences read by the Korean learner of English described in the Materials section above. However, the training session was different in each condition. In the Same Speaker Condition, the sentences in the training session were read by the same Korean learner of English as the sentences in the post-test. Thus, all sentences (i.e., sentences in the training session and post-test) in the Same Speaker Condition were read by the male Korean learner of English described above.

In the Different Speaker Condition, the sentences in the training session were read by a female Korean learner of English. The sentences in the training session were the same as the sentences in the training session of the Same Speaker Condition except for two sentences because not all sentence recordings were available for all speakers in OSCAAR. That is, 38 sentences out of 40 sentences matched in the training sessions of the two conditions; however,

two sentences were not identical across the two conditions. The two sentences that differed consisted of the same number of words and the same number of target words (i.e., six words in each sentence and three target words in each sentence). The post-test was the same as the post-test of the Same Speaker Condition. That is, the same set of sentences read by the male Korean learner of English.

In the Control Condition, the sentences in the training session were the same set of sentences as the Same Speaker and Different Speaker Conditions. However, the sentences were read by the native English speaker described in the Materials section. The post-test was identical to Same Speaker and Different Speaker Conditions. The Control Condition was included to examine whether the native English listeners adapted and generalized to the Korean learner of English or whether the native English listeners solely adapted to the intelligibility task.

2.2.1.4. Procedure

The experiment was conducted online using Qualtrics (https:///www.qualtrics.com). Participants were asked to answer a short language experience questionnaire described in section 2.2.1.1. As described in section 2.2.1.1., the questionnaire was very short and was included to ensure participants met the selection criteria. Participants that did not meet the selection criteria were not invited to complete the experiment. After the language experience questionnaire, participants were asked to read and sign a consent form to participate in the experiment. After signing the consent form, participants were asked to wear their headphones and transcribe three repetitions of an English sentence to make sure participants could hear the items. The sentence that was repeated three times was a short declarative sentence ("This is her favorite sport") read by a native English speaker that was not presented in the main tasks of the experiment. The

sentence was leveled to a fixed RMS amplitude of 73 db. Participants were also asked to adjust the volume to a comfortable level during the sound check. After the sound check, participants were randomly assigned to one of the three conditions (i.e., Same Speaker, Different Speaker, or Control Conditions).

After finishing the sound check, participants were introduced to the main task for the study: an intelligibility task. In the intelligibility task, participants were instructed to listen to sentences using headphones and type what they heard on the keyboard. Participants first heard a practice sentence to familiarize themselves with the task. The practice sentence was from the BKB sentence lists but was read by a different native English speaker than the native English speaker in the Control Condition. The practice sentence was leveled to a fixed RMS amplitude of 73 dB but was not mixed with speech-shaped noise. The practice sentence was not mixed with noise to ensure participants understood the overall transcription task and types of sentences they would be hearing.

After transcribing the practice sentence, participants completed a training session followed by a post-test. Participants heard 120 sentences in the training session and 16 sentences in the post-test. During the training session, participants were exposed to six blocks of 20 sentences. Participants listened to the same set of sentences in the first three blocks and they listened to another set of sentences in the second three blocks. As a result, the participants heard each item three times in the training session. The participants could use as much time as needed to transcribe each sentence. In the post-test, participants were presented with 16 novel sentences that they were not exposed to during the training session. As in the training session, participants could listen to each item once and take as much time as needed to respond to the sentences.

After the experiment, participants were asked to fill out a second questionnaire about their language experience to ensure participants met the selection criteria (i.e., native English speakers who did not have frequent interaction with non-native English speakers) and to record participants' linguistic experience in detail. This language questionnaire that was presented after the intelligibility task served different purposes than the questionnaire that was presented at the beginning of the experiment conducted through Prolific. Specifically, the questionnaire presented at the beginning of the experiment were short multiple-choice questions that aimed to prevent participants who did not meet the selection criteria of the experiment from participating in the experiment. That questionnaire was designed to be short so that participants that did not meet the selection criteria of the experiment would not have to spend more than a couple of minutes on an experiment that they were not allowed to participate in. The questionnaire presented after the intelligibility task asked participants' language background information in detail to understand participants' language experience. This questionnaire asked participant's linguistic experience in detail (e.g., how frequently participants had interaction with family, in the community, at work, etc.) and was used to ensure participants indeed meet the selection criteria of the experiment. Specifically, it was important that the listeners who participated in the experiment did not have frequent interaction with non-native English speakers since it is possible that listeners' previous experience with non-native English speech affects listeners adaptation to non-native English speakers and its generalization to novel non-native English speakers. The questionnaire presented at the beginning of the experiment served the purpose of rejecting listeners who had frequent interaction with non-native English speakers, but it was still possible that listeners answered that they did not have frequent interaction with non-native English listeners in the first language questionnaire even if they had the experience, so that they could complete the

experiment and get paid for their participation. It's also possible that participants didn't fully understand the screen questions at the start of the experiment. For example, one of the experiments conducted in the Speech Perception and Production Lab used the same language experience questionnaire at the beginning and end of the task. In the experiment, there were participants who answered that they did not have frequent interaction with non-native English listeners in the language experience questionnaire that was presented at the beginning of the task but answered that they did have frequent interaction with non-native English speakers in the language experience questionnaire presented at the end. Thus, the data collected in the language experience questionnaire presented at the end of the experiment were used to ensure only native English listeners with no frequent interaction with non-native English speakers participated in the experiment.

2.2.1.5. Analysis

Participants' transcription from the intelligibility task was unnested (i.e., sentences were separated into words) using an R script, manually aligned in Microsoft Excel, and each target word was scored automatically as correct or incorrect using an autoscoring package (Borrie, Barrett, & Yoho, 2019) within the R computing program (R Core Team, 2021). Following previous work (Lee & Baese-Berk, 2020), obvious spelling mistakes and homophones were scored as correct, and words did not need to be transcribed in the order in which they were spoken. While most previous studies analyzed intelligibility task data with logistic mixed-effects regression models (e.g., Baese-Berk, Bradlow, & Wright 2013; Lee & Baese-Berk; 2020), the results of the present study were analyzed with a Bayesian mixed-effects logistic regression model within the R computing program. Results were analyzed with a Bayesian approach to

regression modeling because one of the possible results of the present experiment was a null result in which participants in the Same Speaker, Different Speaker, or Control Conditions showed similar performance in the intelligibility task. If this were the case, the result would be difficult to interpret because the null result does not provide evidence for the null hypothesis. That is, even if the listeners in the Same Speaker, Different Speaker, or Control Conditions show similar intelligibility scores and there is no significant difference between the two, it is not possible to make an interpretation that listeners in the two conditions demonstrated similar results using null-hypothesis significance testing (e.g., a logistic mixed-effects regression model). On the other hand, a Bayesian approach to regression modeling allows describing how likely it is that listeners in the Same Speaker, Different Speaker, Control Conditions have similar intelligibility scores instead of making a threshold-based decision of whether the intelligibility scores of the conditions are significantly different or not. Specifically, a threshold-based decision making is a type of decision making of which a certain threshold value (e.g., probability of the data coinciding with a null hypothesis) is set and the data has meaningful interpretation only when researchers obtain a value that is smaller than the threshold value. For example, if listeners in the Same and Different Speaker Conditions demonstrate similar intelligibility scores, the result is not informative for a threshold-based decision method (e.g., null-hypothesis significance testing) since this method only provides evidence to reject the null hypothesis (i.e., listeners in the Same and Different Speaker Conditions demonstrate similar intelligibility scores). However, a Bayesian approach estimates the probability of the results (e.g., how probable it is that listeners in the Same and Different Speaker Conditions demonstrate similar intelligibility scores). This approach allows a meaningful interpretation of the results especially for the present experiment where it is possible that the three conditions could have similar intelligibility scores.

We fitted a Bayesian logistic mixed model to predict the percent correct of keywords as a function of Condition (Same Speaker, Different Speaker, and Control Conditions) and the model included by-item random intercepts and slopes for condition and random intercepts for participants using the package *brms* (Buerkner, 2017). Condition was Helmert coded to compare the Same Speaker and Different Speaker Conditions to the Same Speaker Condition and the Different Speaker Condition to the Same Speaker Condition. We used weakly informative priors following common practice. Specifically, we used a Student-*t* prior distribution with a mean of 0, degree of freedom of 1, and a scale of 2.5 for the fixed effects. For random effects, we used a Cauchy distribution with a center of 0 and scale of 2, following Gelman, Jakulin, Pittau, & Su (2008).

2.2.2. Results

Figure 1 shows listeners' intelligibility scores (i.e., percent correct of the target words) in the post-test. As shown in Figure 1, listeners trained in the Different Speaker Condition and the Same Speaker Condition (box in the middle and on the right, respectively) demonstrate higher intelligibility scores in the post-test than the listeners trained in the Control Condition (box on the left). These findings suggest that listeners trained with a Korean learner of English in the training session are better at understanding a Korean learner of English in the post-test than listeners who are not trained with a Korean learner of English in the training session.

To examine the effect of training with a single Korean learner of English on the perception of a novel Korean learner of English, we investigated intelligibility scores for listeners in the Different Speaker Condition (box in the middle) and the Same Speaker Condition (box on the right), as shown in Figure 1. Listeners in the Different Speaker Condition demonstrate similar intelligibility scores as the listeners in the Same Speaker Condition. This

finding suggests that training with a single Korean learner of English may help generalization to a novel speaker. This result replicates previous findings that show training with a single non-native English speaker facilitates generalization of adaptation to a novel non-native English from the same language background (e.g., Weil, 2001).

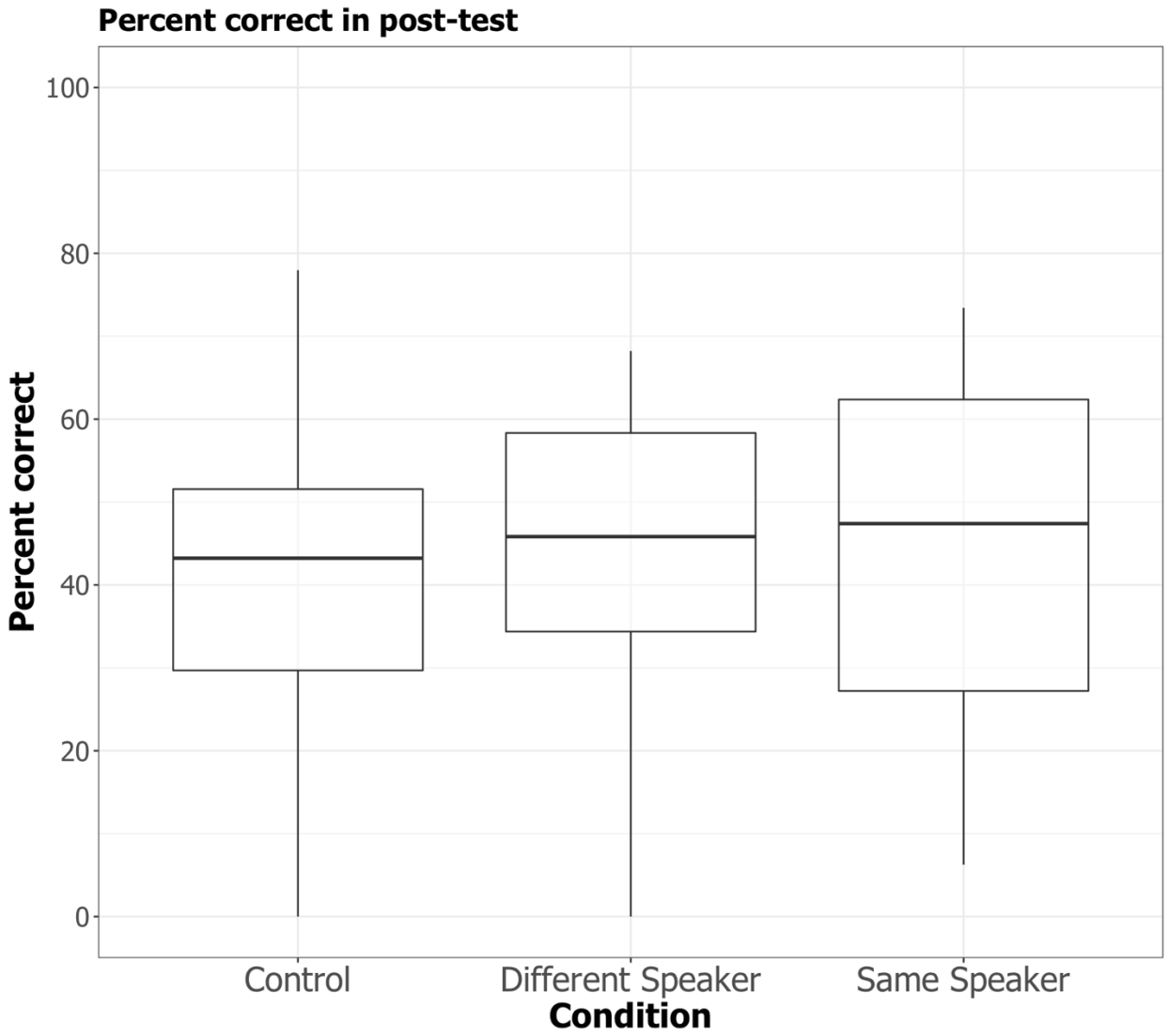

Figure 1. Box plot showing the percent correct on the post-test of the intelligibility task as a function of Condition (Control, Different Speaker, and Same Speaker Conditions). Listeners in the Same Speaker Condition demonstrate the highest intelligibility scores followed by the Different Speaker Condition and the Control Condition.

The Bayesian mixed-effect logistic regression model confirms this trend. Specifically, there is less than a 50% probability that the highest density interval of the mean intelligibility difference of listeners in the Same Speaker and Different Speaker Conditions is smaller than zero, suggesting listeners who listen to the same speaker in training and post-test and listeners who are trained with a non-native speaker and tested with a novel non-native speaker show similar performance in the post-test. Further, there is a 60% probability that the highest density interval of the mean intelligibility difference of listeners in the Same Speaker and Different Speaker Conditions and listeners in the Control Condition is smaller than zero, suggesting that listeners who are trained with a non-native speaker in training are better in the post-test than listeners who do not have training with a non-native speaker in training.

Figure 2 shows listeners' intelligibility scores in the training session. As shown in Figure 2, listeners in the Same Speaker and Different Speaker Conditions demonstrate a general improvement in intelligibility scores across the training session. Specifically, the listeners demonstrate better performance at the end of the training session (i.e., Block 6) than the beginning of the training session (i.e., Block 1). Further, listeners in the Same Speaker and Different Speaker Conditions demonstrate an improvement across the first three blocks (i.e., Blocks 1 – 3) and an improvement across the second three blocks (i.e., Blocks 4 – 6). This result is expected since listeners hear the same sentences in each block from the first block to the third block and hear another set of same sentences in each block from the fourth to the sixth block. That is, listeners are expected to demonstrate better performance as the sentences are repeated. Listeners in the Control Condition demonstrate a similar pattern as the listeners in the Same Speaker and Different Speaker Conditions. However, the improvement of intelligibility scores

across the blocks is smaller compared to that of the listeners in the Same Speaker and Different Speaker Conditions. Further, as in listeners in the Same Speaker and Different Speaker Conditions, listeners in the Control Condition demonstrate general improvement within the first three blocks and within the second three blocks. However, the improvement is not as clear as that of the listeners in the Same Speaker and Different Speaker Conditions, especially in the second three blocks.

Another pattern that listeners in all three conditions demonstrate is a decline in intelligibility scores from Block 3 to Block 4. The decline in Block 4 is likely caused by the introduction of a different set of sentences in Block 4. Specifically, while listeners hear the same set of sentences in the first three blocks, the listener hear a different set of sentences in the second three blocks.

These findings suggest that the listeners in the Same Speaker and Different Speaker Conditions demonstrate adaptation to the Korean learner of English in training and that the listeners do not simply adapt to the intelligibility task. Specifically, if it is the case that listeners adapted only to the task and not to the speech of the Korean learner of English, listeners would not show a decline in intelligibility scores in Block 4 where the listeners are introduced with a new set of sentences.

The results of this study suggest that listeners demonstrate adaptation to non-native English speech as demonstrated in previous studies (e.g., Clarke & Garrett, 2004; Bradlow & Bent, 2008; Xie *et al.*, 2018). That is, as described above, listeners in the Same Speaker Condition and the Different Speaker Conditions demonstrate improvement of intelligibility scores across the six blocks in the training session.

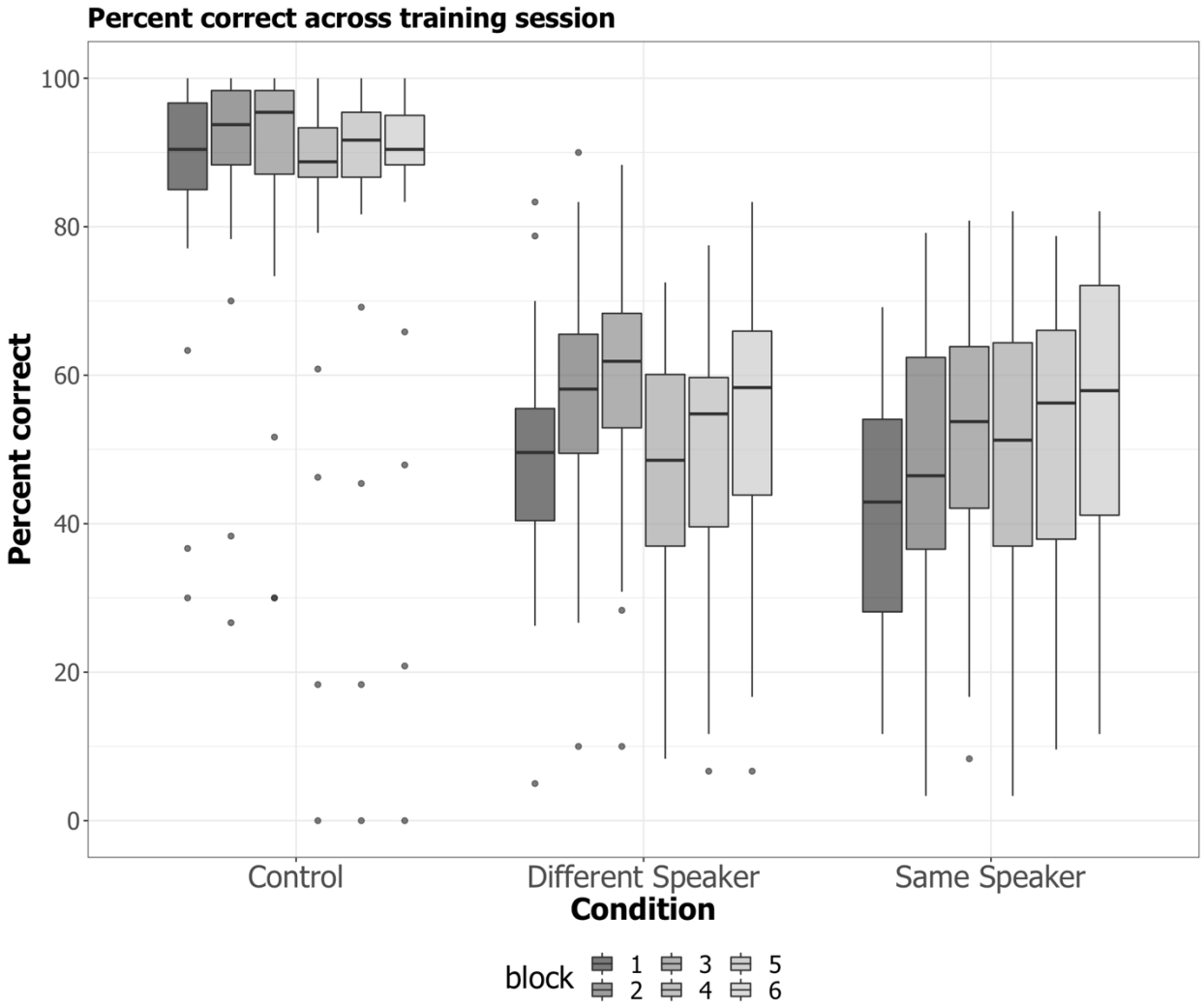

Figure 2. Box plot showing the percent correct on the training session of the intelligibility task as a function of condition (Control, Different Speaker, and Same Speaker Conditions) and block. Listeners in the Different Speaker and Same Speaker Conditions demonstrate a clear increasing pattern in the training session while the trend is weaker for listeners in the Control Condition.

Similarly, listeners demonstrate an increase in intelligibility scores within the first three blocks and the second three blocks. More importantly, the intelligibility scores decrease from Block 3 to Block 4 where listeners are introduced to a new set of sentences. If it were the case that listeners

solely adapt to the intelligibility task and not to the Korean learner of English, listeners would not demonstrate the decrease in intelligibility scores in Block 4.

The results also suggest that listeners are able to generalize to a novel speaker when they are trained with one non-native speaker and tested with another non-native speaker as shown in previous studies (e.g., Weil, 2001). While listeners in the Same Speaker and Different Speaker Conditions show similar performance in the post-test, it may be the case that listeners in the two conditions demonstrate similar performance for different reasons. That is, it is possible that listeners in the Same Speaker Condition demonstrate talker-specific adaptation, as listeners have exposure to the same speaker through training and post-test. On the other hand, for listeners in the Different Speaker Condition, it is possible that the talker change at the beginning of the post-test increased the listeners' attention to the speaker in the post-test. Therefore, Experiment 1B examines how acoustic similarity between speakers and talker information affect generalization of adaptation.

2.3. Experiment 1B

2.3.1. Methods

2.3.1.1. Participants

25 native English speakers between 18 and 38 (13 females, 10 males, 1 non-binary, and 1 transgender) years old participated in this experiment. As in Experiment 1A, participants were recruited from the University of Oregon Psychology and Linguistics subject pool and from Prolific. The inclusion criteria for Experiment 1B were the same as in Experiment 1A.

2.3.1.2. Materials

As in Experiment 1A, items were drawn from OSCAAR, which includes BKB sentence lists. 120 BKB sentences were presented in the training session and 16 BKB sentences that were not presented in the training session were presented in the post-test. Specifically, 40 unique sentences were repeated three times in the training session and 16 unique sentences were presented in the post-test in the same manner as in Experiment 1A. The 40 sentences presented in the training session of Experiment 1B were the same as the 40 sentences presented in the Same Speaker and Control Conditions of Experiment 1A. All sentences were leveled to a fixed RMS amplitude of 73 dB and the stimuli were mixed with speech shaped noise at an SNR of -5 dB (i.e., the same SNR as in Experiment 1A) to prevent ceiling effects. The BKB sentences used in the training session and the post-test are provided in the Appendix.

2.3.1.3. Design

Participants completed an intelligibility task in which they were asked to listen to sentences and transcribe what they heard. In the training session, listeners were presented with six blocks of 20 sentences per block. The first three blocks (i.e., Blocks 1-3) included the same 20 sentences and the second three blocks (i.e., Blocks 4-6) included a different set of 20 sentences. Within each block, all sentences were presented in a random order. Thus, listeners heard a total of 40 unique sentences, with each sentence repeated three times. Following the training session, listeners heard 16 sentences presented in a random order in the post-test.

Using the paradigm described above, a Different F0 Condition was created. In the Different F0 Condition, the sentences in the training session were read by the same Korean learner of English as in the Same Speaker Condition. However, the F0 of the sentences in the

training session was modified using the Change Gender function in Praat (Boersma & Weenink, 2021). Using the Change Gender function, the median F0 of the sentences in the training session was increased to 220Hz from 129.90 Hz and all other acoustic information including formant frequencies, center of gravity, and intensity remained the same. A pilot study was conducted to examine whether the non-native English speaker in the training session was perceived as a different speaker than the non-native English speaker in the post-test (i.e., the same speaker before the F0 modification). In the pilot study, four listeners listened to 10 of the original sentences and another set of 10 sentences that were modified in F0. After listening to the 20 sentences, all listeners responded that they did not think that the 10 original sentences were read by the same speaker as the 10 F0 modified sentences, confirming that listeners perceived the two speakers as distinct. The post-test was identical to that of the same speaker condition.

2.3.1.4. Procedure

The experiment was conducted online using Qualtrics. Participants participated in an intelligibility task and were asked to listen to English sentences and transcribe what they heard. Participants recruited via Prolific started the task with a short language experience questionnaire to ensure only participants that met the selection criteria of the experiment were invited to the experiment. On the other hand, there was no option for the University of Oregon Psychology and Linguistics subject pool to screen participants. Thus, participants recruited via the University of Oregon Psychology and Linguistics subject pool did not participate in the short language experience questionnaire. Instead, participants who answered to have frequent interaction with non-native English speakers in the language experience questionnaire were excluded. Participants recruited from the two platforms went through the same procedure except the short

language experience questionnaire that was included in the Prolific experiment. Listeners were asked to read and sign the consent form and wear headphones before starting the intelligibility task. Then, participants finished a sound check to ensure they could hear the items and to adjust the volume to a comfortable level. After the sound check, participants read the instructions of the intelligibility task and was presented with a practice sentence, as in Experiment 1A. The intelligibility task consisted of a single training session followed by a post-test. In the training session, participants transcribed 120 sentences as described in section 2.3.1.3. Within each block, the sentences were randomly presented and participants could listen to each sentence once. After listening to each sentence, participants could take as much time to transcribe the sentence. In the post-test, participants transcribed 16 sentences that were presented in a random order. Participants were allowed to take as much time to transcribe the sentences. After finishing the intelligibility task, participants were asked to fill out a language experience questionnaire. The task including the intelligibility task and the language experience questionnaire took approximately an hour.

2.3.1.5. Analysis

Participants' transcription from the intelligibility task was unnested using an R script in the R computing program and the transcription was aligned with the target words in Microsoft Excel. Then, each target word was automatically scored as correct or incorrect using an autoscoring script (Borrie, Barrett, & Yoho, 2019) to measure generalization of adaptation. Results were analyzed with a Bayesian mixed-effects logistic regression model within the R computing language as in Experiment 1A. Specifically, a Bayesian logistic mixed model was fitted to predict the performance on the post-test as a function of Condition (Different Speaker,

Same Speaker, Control Conditions from Experiment 1A and Different F0 Condition from Experiment 1B). Condition was Helmert coded to compare: 1) the Control Condition to the Different Speaker, Same Speaker, and Different F0 Conditions, 2) the Different Speaker Condition to the Same Speaker and Different F0 Condition, and 3) the Same Speaker Condition to the Different F0 Condition. The model included by-item random intercepts and slopes for condition and by-participant random intercepts and used the same weakly informative priors as in Experiment 1A.

2.3.2. Results

Figure 3 demonstrates listeners' intelligibility scores in the post-test. As shown in Figure 3, listeners in the Different F0 and Same Speaker Conditions demonstrate higher intelligibility scores than listeners in the Different Speaker Condition. This finding shows that distinct acoustic characteristics between the talkers in the training session and the post-test disrupt generalization of adaptation for listeners training with a single Korean learner of English suggesting that acoustic similarity between speakers plays a significant role in generalization to a novel speaker. Further, the results show that listeners in the Different F0 and Same Speaker Conditions demonstrate similar intelligibility scores in the post-test. The only difference between the Different F0 and Same Speaker Conditions is whether listeners perceive a talker change at the beginning of the post-test. Thus, this result suggests that it is not the case that listeners in the Different Speaker Condition in Experiment 1A demonstrate generalization to a novel speaker because of a perceived talker change at the beginning of the post-test.

The results also show that listeners in the Different, Same Speaker, and Different F0 Conditions (second, third, and fourth boxes) demonstrate higher intelligibility scores than the

listeners in the Control Condition (box on the left). If exposure to a Korean learner of English in the training session did not facilitate perception of a Korean learner of English in the post-test, listeners in the Control Condition would demonstrate similar intelligibility scores as listeners in the Different F0, Different Speaker, and Same Speaker Conditions. However, this is not the case suggesting that exposure to a Korean learner of English in the training session helps perception of a Korean learner of English in the post-test.

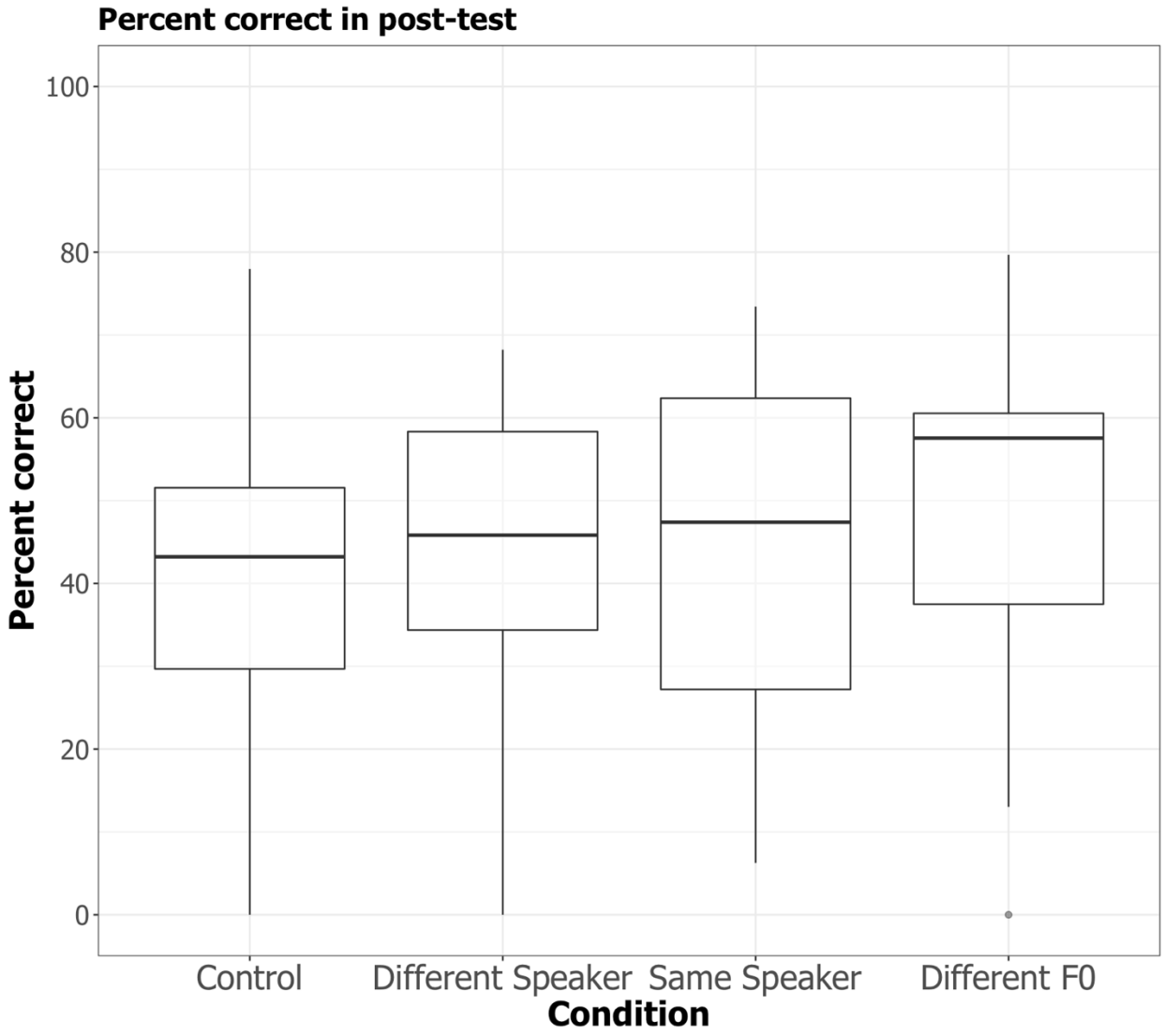

Figure 3. Box plot showing the percent correct on the post-test of the intelligibility task as a function of condition (Control, Different Speaker, Same Speaker, and Different F0 Conditions).

The Bayesian mixed-effect logistic regression model confirmed this trend. Specifically, there is a 54% probability that the highest density interval of the mean intelligibility score difference of listeners in the Same Speaker and Different F0 Conditions and listeners in the Different speaker condition does not include zero, suggesting that acoustic similarity between speakers in training and post-test facilitates speech perception. Further, there is less than a 50% probability that the highest density interval of the mean difference of intelligibility scores of listeners in the Same Speaker Condition and Different F0 Condition does not include zero, suggesting that a perceived talker change does not help perception of a novel speaker. Lastly, there is a 75% probability that the highest density interval of the mean difference of intelligibility scores of listeners in the Same Speaker, Different Speaker, and Different F0 Conditions and the listeners in the Control Condition does not include zero. This result suggests that training listeners with a non-native speaker is helpful for understanding another non-native speaker.

Figure 4 demonstrates listeners' intelligibility scores in the training session. The intelligibility scores of the listeners in the Control, Different Speaker, and Same Speaker Conditions are the same as the intelligibility scores reported in Experiment 1A. The listeners in the Different F0 Condition show similar patterns as listeners in the Different Speaker and Same Speaker Conditions. That is, listeners in the Different F0 Condition demonstrate a general improvement in intelligibility scores across the training session (i.e., higher intelligibility scores in Block 6 than Block 1). Further, the listeners show higher intelligibility scores in the third block than the first block and higher intelligibility scores in the sixth block than the first block. As in the listeners in the Different Speaker and Same Speaker Conditions, listeners in the Different F0 Condition demonstrate a slight decrease in intelligibility scores from Block 3 to

Block 4. As in Experiment 1A, this result is expected since listeners hear the same sentences from Block 1 to Block 3 and another set of sentences from Block 4 to Block 6.

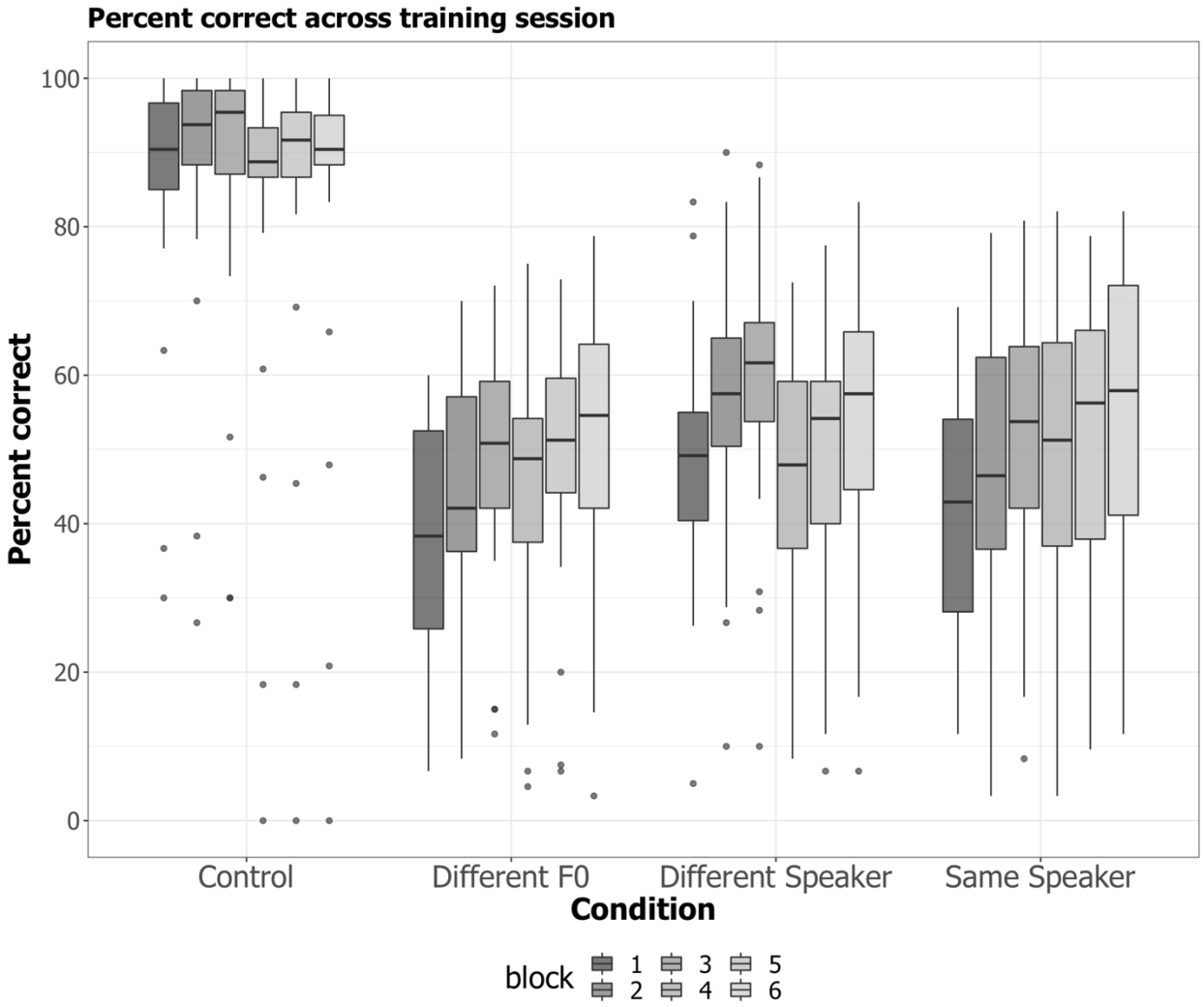

Figure 4. Box plot showing the percent correct on the training session of the intelligibility task as a function of condition (Control, Different F0, Different Speaker, and Same Speaker Conditions) and block. Listeners in the Different F0, Different Speaker, and Same Speaker Conditions show an increase patter in the training session. While the listeners in the Control Condition show a similar pattern, the listeners demonstrate a weaker trend than the listeners in the other three conditions.

The results of Experiment 1B suggest that listeners who have exposure to the Korean learner of English in the training session demonstrate adaptation to non-native English speech over and above adaptation to the intelligibility task, as shown in prior studies (e.g., Clarke & Garret, 2004; Bradlow & Bent, 2008; Xie *et al.*, 2018) and in Experiment 1A. Specifically, listeners demonstrate a slight decline in intelligibility scores from Block 3 to Block 4. As discussed in Experiment 1A, if it were the case that listeners solely adapted to the intelligibility task and not to the speaker, listeners would not demonstrate a decline in intelligibility scores from Block 3 to Block 4. Thus, the decline in Block 4 suggests that listeners adapt to the Korean learner of English in training and that it is not the case that the general increase in intelligibility throughout the six training blocks is driven solely by adapting to the intelligibility task.

2.4. Discussion

2.4.1. Summary of findings

In the present study, we explore how acoustic similarity between talkers and talker information affect generalization of adaptation to a novel speaker. Specifically, we examine whether the acoustic similarity between speakers and a perceived talker change between training and post-test affect listeners' perception of the talker in the post-test. In Experiment 1A, the results demonstrate that training listeners with a single non-native English speaker facilitates listeners' perception of a novel non-native English speaker whether the listeners are trained and tested with the same speaker or trained with one speaker and tested with another. Specifically, listeners who are trained with one non-native speaker and tested with a different non-native speaker demonstrate similar performance as listeners trained and tested with the same non-native

speaker. In order to understand why listeners in the two conditions demonstrate similar performance in the post-test, Experiment 1B investigates how acoustic similarity and talker information affect generalization of adaptation. In Experiment 1B, the results demonstrate that a perceived talker change does not facilitate generalization, as listeners in the Different F0 and Same Speaker Conditions demonstrate similar performance. Rather, it is likely that acoustic similarity between speakers facilitates generalization to a novel speaker, as listeners in the Different F0 and Same Speaker Condition demonstrate better performance in the post-test than listeners in the Different Speaker Condition.

2.4.2. Effect of acoustic characteristics and talker information on generalization of adaptation

The results of the present study suggest that acoustic similarity between speakers in the training session and post-test affects generalization of adaptation. Specifically, it is possible that listeners in the Different Speaker and Same Speaker Condition in Experiment 1A demonstrate similar performance because the speakers in training and post-test in the Different Speaker Condition have similar acoustic features that facilitate generalization to the speaker in post-test. However, as speakers in the training and post-test of the Different Speaker Condition are actual different speakers, it is likely that the speakers in training and post-test still have different acoustic features, which may explain why listeners in the Different F0 and Same Speaker Conditions together demonstrate better performance than listeners in the Different Speaker Condition.

These results also have implications for the role of talker information on generalization of adaptation. That is, the results suggest that the role of talker information may vary as a function of acoustic similarity between speakers. In the current study, listeners in the Different F0 and

Same Speaker Conditions demonstrate similar performance in the post-test even though there is a talker change between the training and post-test in the Different F0 Condition and no talker change between the training and post-test in the Same Speaker Condition. We do not suggest that talker information is not important on speech perception, as numerous studies show how talker information affects perception (e.g., Hay & Drager, 2010; Hay, Nolan, & Drager, 2006; Niedzielski, 1999). However, we suggest that the importance of talker information on generalization of adaptation may be down-weighted when speakers in training and post-test have very similar acoustic features.

Taken together, these results suggest that generalization of adaptation and generalization of phonetic category retuning may have similar underlying mechanisms. Specifically, previous studies show that generalization of phonetic category retuning is constrained by acoustic similarity (e.g., Eisner & McQueen, 2005; Kraljic & Samuel, 2005; Reinisch & Holt, 2014; Xie & Myers, 2017). For example, Eisner & McQueen (2005) show that phonetic category retuning generalizes to a novel speaker when the target phonetic category from training is spliced into the words that the novel talker produces, suggesting that acoustic similarity plays a significant role and different talker information does not disrupt generalization of phonetic category retuning. While it is widely assumed that phonetic category retuning involves similar processes as adaptation to unfamiliar speech (Kleinschmidt & Jaeger, 2015), the assumption has not been thoroughly tested. As the results of the present study suggest that acoustic similarity may be an important factor in generalization of adaptation and a perceived talker change does not disrupt generalization to a novel speaker, it is possible that phonetic category retuning and adaptation to unfamiliar speech involve similar underlying mechanisms.

2.4.3. Effect of Change Gender function on accentedness

One potential issue of the present study is the fact that using the Change Gender function in Praat may modify the accentedness of the Korean learner of English. It is possible that the original Korean learner of English and the F0-modified Korean learner of English have different accentedness and the difference in accentedness could affect listener' performance in the post-test. However, this is not the case in the present study. Specifically, in a pilot test, the accentedness of the original Korean learner of English, the F0-modified Korean learner of English, and the native English speaker received similar accentedness ratings. Native English listeners rated the accentedness of the speakers (i.e., original Korean learner of English, F0-modified Korean learner of English, and native English speaker). Listeners were asked to rate the speakers on the scale of 1 ("no accent") through 9 ("a strong foreign accent"). The mean accentedness rating of the original Korean learner of English was 4.58 (SD = 2.21) and the F0-modified Korean learner of English was 4.92 (SD = 1.96). The original and F0 modified Korean learners of English did not show difference in accentedness ratings (t = -1.56, df = 352.96, p = 0.12).

2.4.4. Conclusion

The current study aims to examine the effects of acoustic characteristics and talker information on generalization of adaptation. Specifically, we examine how acoustic similarities between talkers in the training session and post-test and a perceived talker change affect generalization of adaptation. The results demonstrate that acoustic similarity between speakers play an important role in generalization to a novel speaker. That is, even when listeners perceive a talker change, listeners generalize their adaptation if speakers in training and post-test are

acoustically similar. This result does not necessarily indicate that talker information does not matter in generalization of adaptation. Rather, we suggest that the importance of talker information in generalization may be down-weighted when speakers in training and post-test have similar acoustic characteristics.

III. EFFECT OF ACCENTEDNESS OF NON-NATIVE SPEECH ON GENERALIZATION OF ADAPTATION

3.1. Introduction

Previous studies demonstrate that variability has different effects on speech perception (Baese-Berk, Bradlow, & Wright, 2013; Bradlow & Bent, 2008; Lively, Logan, & Pisoni, 1993; Mullennix, Pisoni, & Martin, 1989). That is, while exposure to variability (e.g., exposure to more than one talker or exposure to different realizations of speech sounds) may initially be a challenge for speech perception, it may also be helpful for speech perception in that exposure to variability facilitates learning characteristics of unfamiliar speech (e.g., time-compressed speech, regional-accented speech, non-native speech) and help better understand unfamiliar speech. For example, listening to multiple talkers makes word identification more difficult than listening to a single talker (Mullennix, Pisoni, & Martin, 1989) but exposure to variability helps listeners understand the speech of a novel non-native speaker (Baese-Berk, Bradlow, & Wright, 2013; Bradlow & Bent, 2008). Although these studies suggest that exposure to variability is helpful for learning the characteristics of unfamiliar speech, it is less well-understood how exposure to variability is helpful. That is, exposure to variability is not uniformly helpful for speech perception (e.g., Perrachione, Lee, Ha, & Wong, 2011; Tzeng, Alexander, Sidaras, & Nygaard, 2016). Thus, the present study aims to better understand what aspects of variability is helpful for speech perception by examining the effect of accentedness of non-native speech on generalization of adaptation.

3.1.1. Effects of high-variability training on speech perception

Previous studies demonstrate the effects of high variability training on perceptual learning involving both non-native English speakers and native English speakers (Lively, Logan, & Pisoni, 1993; Wang, Spence, Jongman, & Sereno, 1999; Clopper & Pisoni, 2004; Bradlow & Bent, 2008; Baese-Berk, Bradlow, & Wright, 2013). For example, non-native English listeners trained with words produced by multiple native English speakers successfully learned a novel phonetic category and generalized to a novel speaker. On the other hand, listeners trained with a single speaker failed to generalize their learning to a novel speaker (Lively, Logan, & Pisoni, 1993). The benefits of high-variability perceptual training are also demonstrated in native listeners' adaptation to non-native speech. That is, native English listeners trained with multiple Mandarin accented English speakers demonstrated better understanding of a novel Mandarin accented English speaker than listeners trained with native English speakers. However, listeners trained with a single Mandarin accented English speaker did not show better understanding of a novel speaker than listeners trained with native English speakers (Bradlow & Bent, 2008). These results suggest that high-variability perceptual training may be helpful for speech perception, especially for generalizing perceptual learning to a novel speaker.

However, high-variability training does not uniformly facilitate generalization to a novel speaker. That is, it is not high-variability perceptual training per se that facilitates generalization of adaptation. For example, even when listeners have exposure to the same set of sentences produced by the same non-native speakers in a high-variability perceptual training, listeners may or may not generalize their adaptation to a novel non-native speaker. Specifically, while listeners who have exposure to training items blocked by speaker and sentence do not demonstrate generalization to a novel speaker, listeners who have exposure to training items randomized by

both speaker and sentence demonstrate generalization to a novel speaker (Tzeng, Alexander, Sidaras, & Nygaard, 2016). The results of the study suggest that exposure to variability does not necessarily facilitate generalization of adaptation. Further, these results have implications for how high-variability perceptual training is helpful for speech perception. That is, it is possible that exposure to variability helps listeners learn the common characteristics of the target speech as suggested in previous studies (e.g., Baese-Berk, Bradlow, & Wright, 2013; Laturnus, 2018; Laturnus, 2020). Indeed, previous studies suggest that non-native speakers may share common characteristics (Guion, Flege, Liu, & Yeni-Komshian, 2000; Laturnus, 2020; Munro & Derwing, 1995; Toivola, Lennes, & Aho, 2009). For example, Farsi and Italian learners of English have shorter VOT in stop sounds than native English speakers Farsi, Korean, and Thai learners of English often produce schwas as unreduced vowels while native English speakers tend to reduce schwas (Laturnus, 2020). Thus, it is possible that non-native English speakers share common characteristics and other features of speech that highlight these common characteristics of non-native speech may facilitate generalization of adaptation to a novel non-native speaker.

One feature that may affect generalization of adaptation to a novel non-native speaker is accentedness of non-native speech. Accentedness refers to a perceived degree of a speaker's non-native accent (Munro & Derwing, 1995). Previous studies show that the deviance in acoustic characteristics between native and non-native speech is an important predictor of accentedness (Munro, 1993; Porretta, Kyröläinen, & Tucker, 2015). For example, Porretta, Kyröläinen, & Tucker (2015) demonstrate that the distance of F1 and F2 values between native and non-native speakers have a positive correlation with accentedness ratings. Further, Munro (1993) demonstrates similar results with a different group of non-native speakers and suggests that non-native speakers' production of their non-native speech may deviate from native speakers because

their first language affects the production of their non-native speech. Indeed, Laturnus (2020) demonstrates that the speech of non-native speakers who share the same native language systematically differs from that of native speakers. For example, Farsi and Italian learners of English demonstrate significantly shorter voiced VOT than native English speakers. These results suggest that exposure to non-native speech with different degrees of accentedness may affect how listeners learn characteristics of non-native speech. That is, if non-native speech indeed has common characteristics, more accented speech may highlight the characteristics as more accented speech is likely to have more distinct characteristics than less accented speech. However, it is unclear how accentedness of non-native speech affects speech perception. Examining how accentedness of non-native accents affects generalization to a novel speaker would allow us to broaden our understanding of exposure to variability could be beneficial for speech perception.

3.1.2. Current study

In the current study, we examine the effect of accentedness of non-native speech on generalization of adaptation to a novel non-native English speaker. The study consists of an intelligibility task and acoustic analyses of the Korean learners of English in the training session and the post-test. The intelligibility task aims to examine how accentedness of non-native English speech affects generalization of adaptation to a novel Korean learner of English. Two outcomes are possible for the effect of accentedness of a non-native speech on generalization to a novel speaker. First, it is possible that more accented non-native speech facilitates generalization to a novel speaker than less accented non-native speech. Previous studies suggest that non-native speakers who share the same first language may share common characteristics (e.g., Guion,

Flege, Liu, & Yeni-Komshian, 2000; Laturnus, 2020; Munro & Derwing, 1995; Toivola, Lennes, & Aho, 2009). Then, exposure to more accented non-native speech may highlight these common characteristics. That is, as more accented non-native speech deviates more from native speech than less accented non-native speech (Munro, 1993; Porretta, Kyröläinen, & Tucker, 2015), exposure to more accented non-native speech may help listeners learn characteristics of non-native speech and generalize to a novel non-native speaker from the same language background.

On the other hand, it is also possible that exposure to more accented non-native speech disrupts generalization to a novel non-native speaker. Previous studies demonstrate that learning in easy environments is transferred to novel items while learning in difficult environments is item specific in visual perceptual learning (Ahissar & Hochstein, 2004). Specifically, participants focus on more general patterns when learning occurs in easy environments and participants focus on the specific details when learning occurs in difficult environments. It is possible that learning occurs in a similar manner in the speech domain. That is, listeners may learn the acoustic details of non-native accented speech when they are trained in difficult environments and learn the common characteristics of non-native accented speech when they are trained in easy environments. More accented non-native speech would likely be more difficult for listeners to process than less accented non-native speech since more accented non-native speech is likely to have more distinct characteristics than less accented speech non-native speech (Munro, 1993; Porretta, Kyröläinen, & Tucker, 2015). If this is the case, having exposure to more accented non-native speech would disrupt generalization to a novel non-native speaker.

The acoustic analyses examine the acoustic similarity between each of the talkers in the training session and the talker at post-test. One potential factor that may affect generalization of adaptation is the acoustic similarity between the talkers in the training and post-test as suggested

in Experiment 1. If this is the case, generalization to the talker in post-test may occur regardless of the training condition (i.e., accentedness of non-native speech) if speakers in the training condition and post-test have similar acoustic characteristics. That is, accentedness of non-native speech may, in fact, have no effect on generalization and generalization of adaptation may be constrained by acoustic similarity between speakers in training and post-test. For example, regardless of the accentedness of speakers, listeners in one condition may generalize to the test talker if the talkers in training and post-test have similar acoustic features. Similarly, listeners in another condition may not demonstrate generalization if talkers in training and post-test have distinct acoustic features. Thus, the acoustic analyses aim to address this potential issue and examine whether speech rate, median F0, and F0 range between speakers in training and post-test are more similar in either one of the conditions presented in the current study. Specifically, we examine speech rate because slower speech rate is one of the characteristics of non-native speakers (Munro & Derwing, 1995; Guion, Flege, Liu, & Yeni-Komshian, 2000) and it is possible that listeners attend to this feature in adapting to non-native speakers and generalizing to a novel speaker. Further, we examine median F0 and F0 range as previous studies demonstrate that F0 affects whether listeners perceive different speakers as similar or dissimilar (Perrachione, Furbeck, & Thurston, 2019; Roark, Fend, & Chandrasekaran, 2022). As the results of Experiment 1 suggests that similarity between speakers may affect generalization, it is possible that similar F0 and F0 range affect generalization.

3.2. Methods

3.2.1. Participants

75 native English speakers between 18 and 39 years old participated in this experiment (37 female, 38 male). Participants recruited from the University of Oregon Psychology and Linguistics subject pool received partial course credits for their participation. As in Experiment 1, participants recruited from the University of Oregon Psychology and Linguistics subject pool who had frequent exposure to non-native English speakers and who had not used headphones were excluded from data analysis (i.e., the same criteria used in Experiment 1).

Participants recruited from Prolific were paid $7.50 for their participation. As in Experiment 1, participants on Prolific began the experiment with a language experience questionnaire that checked participants' eligibility. Then, the participants were not allowed to continue the experiment if they had frequent interaction with non-native English speakers and did not have headphones. No participant reported a history of speech or hearing disorder.

3.2.2. Materials

Items were drawn from the Online Speech/Corpora Archive and Analysis Resource (OSCAAR) (Bradlow, n.d.), which includes Bamford-Kowal-Bench (BKB) sentence lists (Bamford & Wilson, 1979; Bench & Bamford, 1979) read by 10 (five females and five males) Korean-English bilinguals, 10 (five females and five males) native English speakers. Further, two female Korean learners of English were recruited and recorded in the present study. The two speakers read the same BKB sentence lists as the 10 Korean-English bilinguals and the 10 native English speakers in OSCAAR. As in Experiment 1, the 10 Korean learners of English from OSCAAR were all born in Korea and spoke Seoul dialect as their first language. They were

educated in Korean from elementary school to university and their length of residence in English-speaking countries ranged from 0.1 to 3.5 years (M = 1.6 years). Further, the speakers' L2 proficiency ranged from 34 to 77 (mean: 55.2, SD: 11.55) on the Versant English test (Bradlow, Blasingame, & Lee, 2018). Among the five female and five male Korean learners of English from OSCAAR, only the five female speakers were selected in the present study and two additional female speakers were recruited for recording. Female speakers were chosen in the present study instead of a mix of female and male speakers or only male speakers for two reasons. Sex could be a confounding factor. That is, the design of the present study required seven Korean learners of English. Since the ALLSSTAR corpus contains only five female and five male speakers of Korean, it was possible to use a mix of the female and male speakers, but it was not possible to present speakers of the same sex in each condition. If listeners were presented with speakers consisting of both female and male speakers, some listeners would experience a switch in sex of talkers from the training session to the post-test which could affect the results. That is, more salient changes in talker characteristics may affect listeners' perception of the talker in the post-test. Thus, only female speakers were included and two additional female Korean learners of English were recruited for recording. These speakers were chosen because of challenges due to the COVID-19 pandemic. Therefore, in the present study, only speech from female Korean learners of English was used as stimuli.

The two Korean learners of English recruited to record the BKB sentences were born in Korea and spoke Seoul dialect as their first language. One talker was educated in Korean from elementary school to university and the other talker was educated in Korean from elementary school to university except a year of high school that the speaker spent in the United States. The two speakers did not have any official test scores of their English skills that can be compared to

the Versant English test which the 10 Korean-English bilinguals from OSCAAR had taken; however, the accentedness of the Korean learners of English from OSCAAR and the two Korean learners of English recruited for the present experiment was rated in a pilot experiment and the Korean learners of English were allocated to one of the two conditions (described in section 3.2.3.) based on their accentedness ratings. The two Korean learners of English were asked to read the same BKB sentence lists from OSCAAR that the 10 Korean-English bilinguals and 10 native English speakers read. The two talkers were recorded while they read the BKB sentences and from the recordings, 40 sentences were used in the training session and 16 sentences in the post-test. The sentences in the training session and post-test were leveled to a fixed RMS amplitude of 73 dB.

As in Experiment 1, a pilot experiment was conducted to determine the speech-to-noise ratio (SNR) between the speech-shaped sound and the sentences that would prevent ceiling effects. The pilot experiment was designed to determine the SNR that would prevent native English listeners from scoring over 70% in the intelligibility task so that participants trained were able to improve beyond the baseline level of performance. The results showed that at an SNR of -5 dB, native English listeners did not show ceiling effects. Thus, all sentences were mixed with a speech-shaped noise at an SNR of -5 dB. The list of BKB sentences presented in Experiment 2 are provided in the Appendix.

3.2.3. Design

As in Experiment 1, participants completed an intelligibility task in which participants listened to short declarative sentences and transcribed what they heard. The design of the task largely followed Experiment 1. However, Experiment 2 differed in the number of speakers in the

training session. Specifically, in Experiment 1, all six blocks were read by the same speaker within each condition. In Experiment 2, on the other hand, each of the first three blocks was read by a different speaker in a randomized order (i.e., three speakers in total) and each block in the second three blocks was read by the same three speakers as in the first three blocks in a randomized order. At the end of the training session, native English listeners heard 120 English sentences (i.e., 40 sentences read by three speakers).

Participants were randomly assigned to one of three conditions: More Accented, Less Accented, and Control Conditions. Each condition included a training session and a post-test. The post-test was the same in all three conditions. That is, the listeners were asked to transcribe 16 sentences read by a Korean learner of English. However, the training session was different in each condition. That is, the sentences in the training session were read by speakers of accentedness in the three conditions. To determine the speakers that would be included in the More Accented and Less Accented Conditions, a pilot experiment was conducted. In the pilot experiment, eight native English listeners heard eight Korean learners of English reading eight sentences from the BKB sentence lists (i.e., each speaker read a different set of eight sentences). The native English listeners were asked to transcribe what they heard and rate the accentedness of the sentences on the scale of 1 (“no accent”) through 9 (“a strong foreign accent”). Based on the results of the pilot experiment, two groups that each consisted of three Korean learners of English were created. Specifically, the two groups were rated as having significantly different accentedness ratings. The three Korean learners of English in the group that was rated as more accented were included in the training session of the More Accented Condition and the three Korean learners of English in the group that was rated as less accented were included in the training session of the Less Accented Condition. Further, the mean accentedness ratings of

speakers in the More Accented and Less Accented Conditions did not overlap with each other. That is, all Korean learners of English in the More Accented Condition were rated as more accented than all of the Korean learners of English in the Less Accented Condition. The post-test was the same in both conditions. That is, in the post-test of the More Accented and Less Accented Conditions, a Korean learner of English that was rated as more accented than the speakers in the two conditions read the 16 sentences that were not presented in the training session. The accentedness ratings of the Korean learners of English are shown in Table 1.

Speaker	Condition	Accentedness
MA01	More Accented Condition	4.86 (2.22)
MA02	More Accented Condition	5.41 (1.97)
MA03	More Accented Condition	4.35 (1.97)
LA01	Less Accented Condition	4.12 (1.92)
LA02	Less Accented Condition	4.10 (2.05)
LA03	Less Accented Condition	4.26 (1.97)
PT01	Post-test	5.56 (2.02)

Table 1. Mean accentedness rating and standard deviation (in parentheses) of the Korean learners of English in the More Accented and Less Accented Conditions and the Korean learner of English in the post-test.

As in Experiment 1, the Control Condition aimed to examine whether the participants in the More Accented and Less Accented Conditions showed generalization of adaptation to the

Korean learners of English or adaptation to the intelligibility task. Note: the Control Condition was different than the Control Condition in Experiment 1 because of the different designs of the experiments. Specifically, while listeners heard one speaker in the training sessions of the experimental conditions (i.e., Different F0, Different Speaker, and Same Speaker Conditions) in Experiment 1, listeners heard three speakers in the training session of the experimental conditions (More Accented and Less Accented Conditions) in Experiment 2. To match the number of speakers listeners heard in the experimental conditions, listeners heard one native English speaker in the Control Condition in Experiment 1 and three native English speakers in the Control Condition in Experiment 2.

3.2.4. Procedure

As in Experiment 1, the experiment was conducted online using Qualtrics (https:///www.qualtrics.com). The description of the experiment explained that participants were to listen to English sentences and transcribe what they heard. As in Experiment 1, participants recruited through Prolific were first asked to answer a short language experience questionnaire. The questionnaire was included to ensure only participants that meet the selection criteria of the experiment were invited to the experiment. After finishing the questionnaire, participants were asked to read and sign the consent form. Then, participants were asked to wear headphones before starting the intelligibility task. After participants were asked to wear headphones, the same sentence used in the sound checking process of Experiment 1 was presented. Specifically, participants were asked to listen to three repetitions of a short English sentence and transcribe what they heard to ensure they could hear the items. Participants were also asked to adjust the volume to a comfortable level.

As in Experiment 1, after the sound check was over, participants were given a description of the intelligibility task and was presented with a practice sentence (i.e., the same sentence as in Experiment 1). In the intelligibility task, participants went through a single training session followed by a post-test. During the training session, participants transcribed 120 sentences. As described in the section 3.2.3., the sentences were presented in six blocks and within each block, participants could listen to each item once. However, they could take as much time to transcribe each sentence. After finishing the training session, listeners participated in the post-test. In the post-test, participants were presented with 16 sentences that they did not hear during the training session. As in the training session, participants could listen to each sentence once and take as much to transcribe the sentences.

After the intelligibility task, participants were asked to fill out a language experience questionnaire. As in Experiment 1, the language experience questionnaire asked participants' language background information in detail to understand participants' language experience. The experiment including the main task and the language experience questionnaire lasted approximately an hour.

3.2.5. Analysis

As in Experiment 1, participants' transcription from the intelligibility task was unnested using an R script. Further, each target word was scored automatically as correct or incorrect using an autoscoring script (Borrie, Barrett, & Yoho, 2019) to measure generalization after a manual alignment using Microsoft Excel. Following previous work (Lee & Baese-Berk, 2020), obvious spelling mistakes and homophones were scored as correct, and words did not need to be transcribed in the order in which they were spoken. Results were analyzed with a Bayesian

mixed-effects logistic regression model within the R computing language. The results were analyzed using a Bayesian approach because one of the possible results of Experiment 2 was participants in the More Accented and Less Accented Conditions showing similar performance in the post-test. If this were the case, interpreting the results would be difficult because the null result does not provide evidence for the null hypothesis. However, a Bayesian approach would allow a meaningful interpretation as it is possible to calculate the probability of two different conditions having similar results. Specifically, we fitted a Bayesian logistic mixed model to predict listeners' performance on the post-test with Condition (More Accented, Less Accented, and Control Conditions). Condition was Helmert coded [(i) Control vs Less Accented and More Accented Conditions, (ii) Less Accented vs More Accented Condition] to compare the Control Condition to the Less Accented and More Accented Conditions and the Less Accented Condition to the More Accented Condition. The model included by-item random intercepts and slopes and random intercepts for participants and weakly informative priors were used following common practice. That is, the prior included a Student-*t* prior distribution with a mean of 0, degree of freedom of 1, and a scale of 2.5 for the fixed effects and a Cauchy distribution with a center of 0 and a scale of 2 (Gelman, Jakulin, Pittau, & Su, 2008).

To complement the results of the intelligibility task, acoustic analyses were conducted and reported. Specifically, the acoustic analyses aimed to examine whether the Korean learners of English in the training session had similar acoustic characteristics as the Korean learner of English in the post-test. These acoustic analyses were crucial since the findings of Experiment 1 suggested that acoustic similarities between the talkers in the training session and the post-test affects generalization of adaptation to novel Korean learners of English. Therefore, we needed to

examine whether acoustic analyses alone could account for any patterns in our adaptation experiment for this experiment.

The acoustic analyses included measures of speaking rate, median F0, and F0 range. Speaking rate was measured using a script within the R computing program (R Core Team, 2018). The script measured the duration of each item and speaking rate was calculated by dividing the duration of the duration of item by number of syllables within the item. Median F0 and F0 range were measured using a Praat (Boersma & Weenink, 2021) script. Specifically, the Praat script measured median F0, 75th, and 25th quantiles of F0. Median F0 was measured because the F0 information was measured automatically using a Praat script and the median is more resistant to outliers than the mean. That is, if the Praat script makes measurement errors (e.g., erroneously high F0), it is likely that F0 mean is influenced by the error than median F0. Similarly, F0 range was measured by subtracting the 25th quantile from the 75th quantile instead of subtracting minimum F0 from maximum F0 to reduce the possibility of erroneous measurements affecting the F0 range.

3.3. Results

The results section of Experiment 2 reports the analyses of the intelligibility task and the acoustic analyses of the Korean learners of English presented in the training session and the post-test of the intelligibility task of Experiment 2. The acoustic analyses are conducted to examine whether patterns of intelligibility scores could be driven by the acoustic similarity between the Korean learners of English in the training sessions and the post-test. In the acoustic analyses, speech rate (syllables per second), median F0, and F0 range of the Korean learners of English are reported.

3.3.1. Intelligibility task

Figure 5 shows listeners' intelligibility scores in the post-test. As shown in Figure 5, listeners trained in the Less Accented Condition (box in the middle) demonstrate higher intelligibility scores in the post-test than listeners trained in the Control Condition (box on the left), but the listeners in the More Accented Condition (box on the right) demonstrate similar intelligibility scores as the listeners in the Control Condition. That is, while listeners in the More Accented and Less Accented Conditions hear multiple Korean learners of English, only listeners in the Less Accented Condition demonstrate generalization of adaptation to a novel Korean learner of English. This finding suggests that training with multiple non-native English speakers does not necessarily facilitate generalization of adaptation to a novel non-native English speaker. Further, the listeners trained in the Less Accented Condition demonstrate higher intelligibility scores in the post-test than the listeners trained in the More Accented Condition. This finding demonstrates that exposure to a more accented non-native speech disrupts generalization to a novel speaker.

The Bayesian mixed-effect logistic regression model confirms this trend. That is, there is a 65% probability that the highest density interval of the mean intelligibility difference of listeners in the More Accented and Less Accented Conditions does not include zero, suggesting that exposure to more accented non-native speech disrupts generalization. Further, there is less than a 50% probability that the highest density interval of the mean intelligibility difference of the listeners in the More Accented and Less Accented Conditions and the listeners in the Control Condition does not include zero, suggesting that exposure to multiple non-native speakers may not always be helpful for understanding a novel speaker.

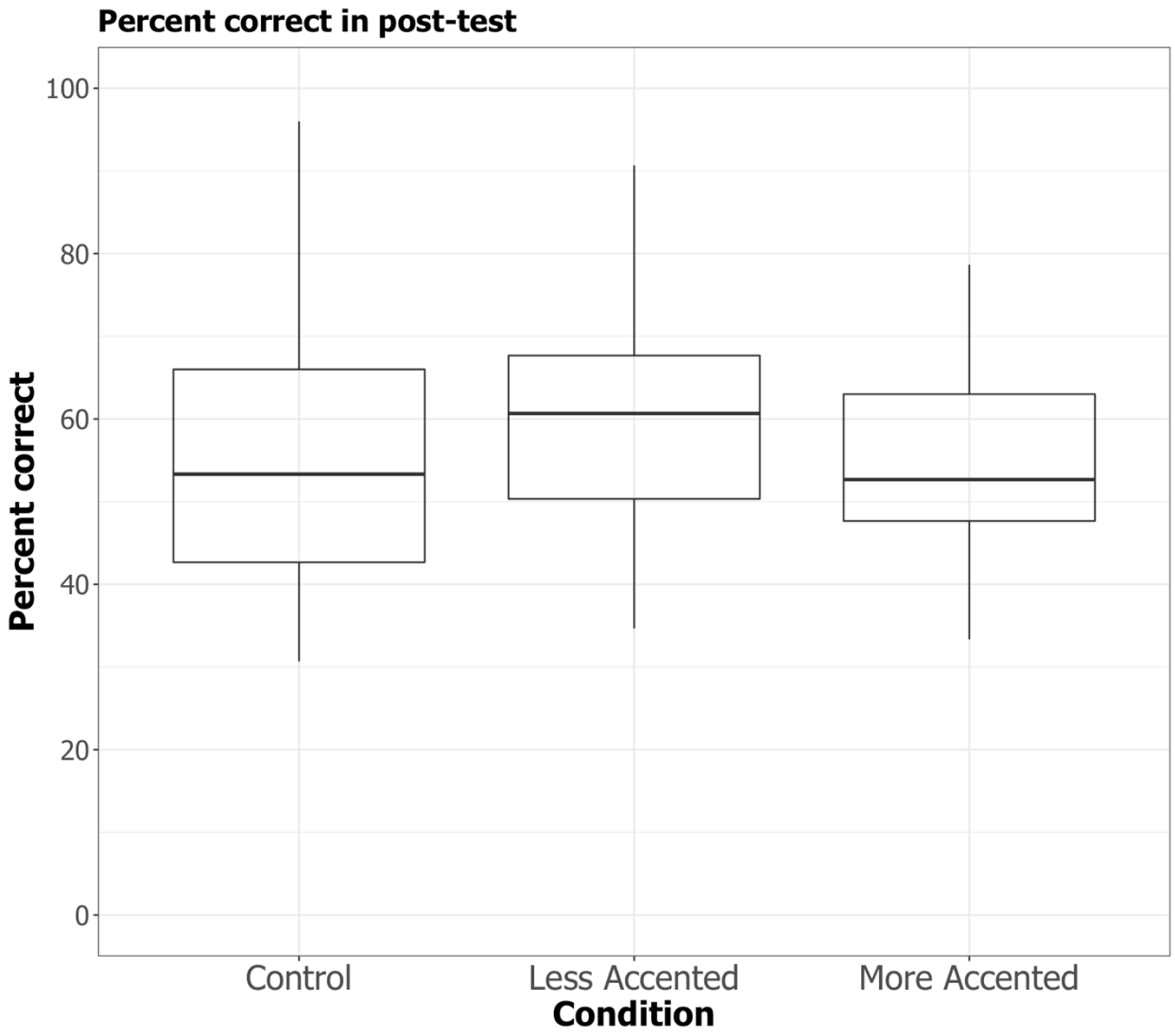

Figure 5. Box plot showing the percent correct on the post-test of the intelligibility task as a function of condition (Control, Less Accented, and More Accented Conditions). Listeners in the Less Accented Condition demonstrate better performance than the listeners in the Control and More Accented Conditions. However, listeners in the More Accented Condition show similar performance as the listeners in the Control Condition.

Further, a post-hoc analysis compared the Less Accented Condition and Control Condition to examine whether listeners in the Less Accented Condition demonstrate generalization. The

results show that there is a 65% probability that the highest density interval of the mean intelligibility difference of listeners in the Less Accented and Control Condition does not include zero, suggesting that listeners in the Less Accented Condition demonstrate generalization to the novel speaker.

Figure 6 shows listeners' intelligibility scores in the training session. As shown in Figure 6, listeners in all three conditions (i.e., More Accented, Less Accented, and Control Conditions) demonstrate a general improvement in intelligibility scores across the six blocks of the training session. Specifically, the listeners in the More Accented and Less Accented Conditions demonstrate higher intelligibility scores at the end of the training session (i.e., Block 6) than the beginning of the training session (i.e., Block 1). While the listeners in the Control Condition demonstrate a similar pattern as the listeners in the More Accented and Less Accented Conditions, the improvement of intelligibility scores across the training session is smaller than that of the listeners in the More Accented and Less Accented Conditions.

The results of this study suggest that accentedness of non-native speech affects generalization of adaptation to a novel Korean learner of English. Specifically, exposure to more accented non-native speech disrupts generalization of adaptation to a novel Korean learner of English. As discussed in the introduction, if accentedness of non-native speech did not affect generalization of adaptation, listeners in the More Accented and Less Accented Conditions would not demonstrate different intelligibility scores in the post-test since the difference between the two conditions is the accentedness of the Korean learners of English in the training session. The implications of these findings are explained in section 3.4. below.

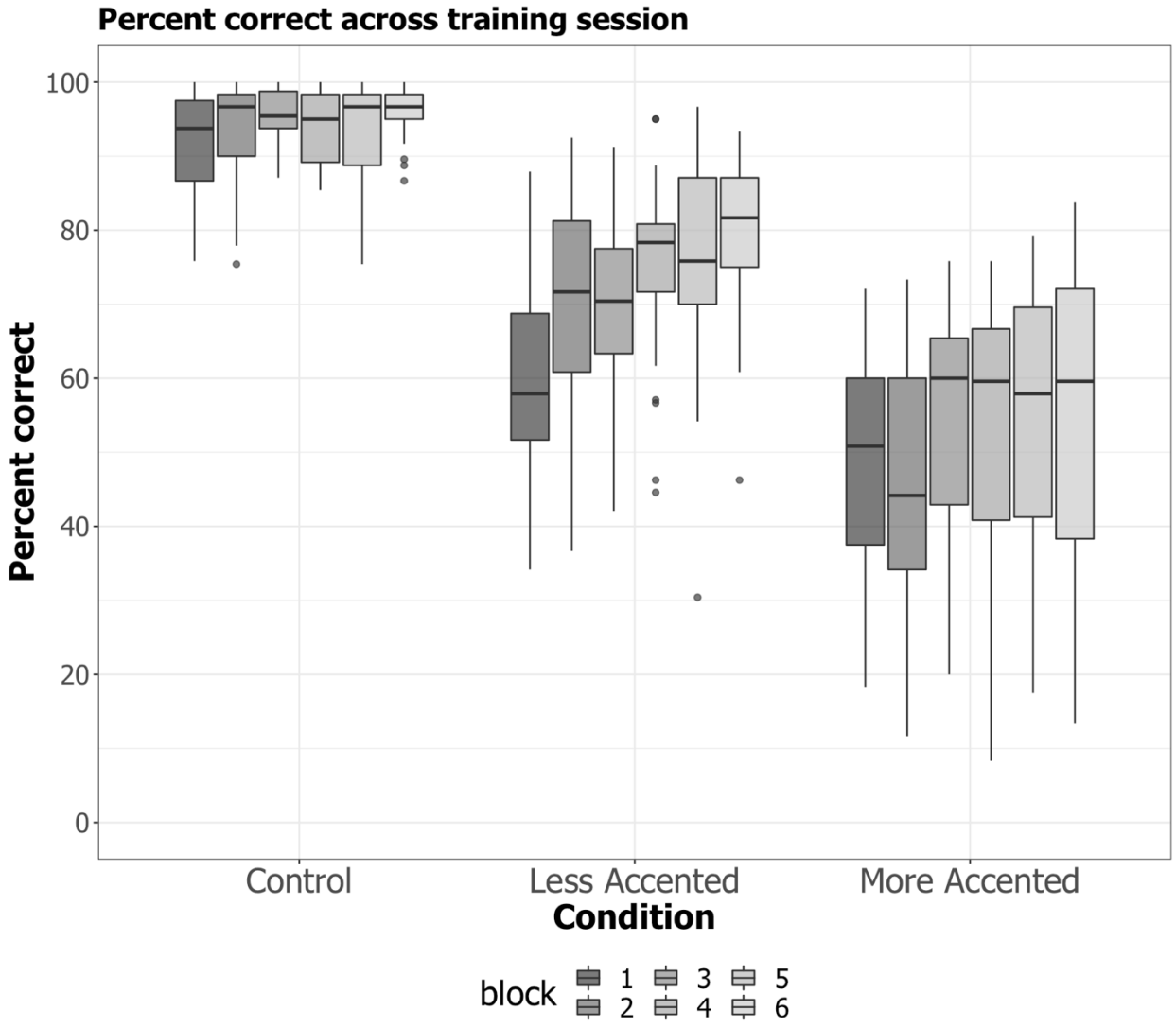

Figure 6. Box plot showing the percent correct on the training session of the intelligibility task as a function of condition (Control, Less Accented, and More Accented Conditions) and block.

Further, exposure to multiple non-native English speakers does not necessarily facilitate generalization of adaptation to a novel non-native English speaker from the same language background. If hearing multiple non-native English speakers always facilitated generalization of adaptation to a novel Korean learner of English, listeners in the More Accented Condition would demonstrate higher intelligibility scores in the post-test than listeners who hear native English

speakers in the training session (i.e., listeners in the Control Condition). That is, while listeners in the More Accented Condition hear multiple Korean learners of English in the training session, listeners in the Control Condition hear native English speakers in the training session and the listeners in the two conditions demonstrate similar performances in the post-test. Thus, the listeners in the More Accented Condition do not demonstrate generalization adaptation beyond generalizing their adaptation to the intelligibility task. These findings suggest that exposure to multiple non-native English listeners does not guarantee generalization of adaptation to a novel non-native English speaker and the implications of these findings is explained in section 3.4.

The results also suggest that in some cases, listeners demonstrate generalization of adaptation to a novel non-native English speaker after a short training session. Most previous studies that examine generalization of adaptation to novel non-native speakers used more items (e.g., Bradlow & Bent, 2008; Baese-Berk, Bradlow, & Bent, 2013; Laturnus, 2018) in the training session than the present experiment and trained listeners for two consecutive days (e.g., Bradlow & Bent, 2008; Baese-Berk, Bradlow, & Wright, 2013). On the other hand, listeners in the present experiment hear three repetitions of 40 sentences in the training session and participate in a single training session which takes less than 60 minutes on average. Even though the training session is shorter and fewer items are presented in the present experiment than previous experiments, listeners in the present experiment demonstrate generalization of adaptation to a novel Korean learner of English. This conclusion is supported by the result that listeners in the Less Accented Condition demonstrate higher intelligibility scores than the listeners in the Control Condition after a single training session that took less than 60 minutes on average. Specifically, in a post-hoc analysis, we compare the intelligibility scores in the post-test of listeners in the Less Accented and Control Conditions. The results show that there is a 60%

probability that the difference in intelligibility scores between the two conditions do not include zero, suggesting that listeners in the Less Accented Condition demonstrate generalization to a novel speaker after a short training session.

In summary, accentedness of non-native speech affects generalization of adaptation to a novel non-native speaker. That is, listeners who hear a less accented non-native speech in training show better perception of a novel non-native speaker in the post-test than listeners who hear a more accented non-native speech in training. However, there is a potential issue that may impact generalization of adaptation to a novel Korean learner of English. Specifically, as shown in Experiment 1, the acoustic similarity between the Korean learners of English in the training session and post-test may impact generalization of adaptation. For example, if there were one Korean learner of English in the training session that had similar acoustic characteristics as the Korean learner of English in the post-test, listeners who hear the Korean learner of English in the training session would likely demonstrate higher intelligibility scores than listeners who do not hear this talker in the training session. Thus, we address the potential issue in the next section.

3.3.2. Acoustic analyses

In the present section, the results of the acoustic analyses are reported to examine whether the patterns of intelligibility scores reported in section 3.3.1. are driven by the acoustic similarity between the Korean learners of English presented in the training sessions and the post-test. In this section, the speech rate, median F0, and F0 range of the three Korean learners of English in the training session of the More Accented Condition, the three Korean learners of English in the training session of the Less Accented Condition, and the Korean learner of English in the post-test (i.e., the same speaker across the More Accented, Less Accented, and Control

Conditions) are reported. Specifically, we compare the acoustic similarity between the Korean learners of English in the training session and the post-test to address the possible issue that generalization is driven by acoustic similarity between these speakers. Speech rate is analyzed in the present analysis since slower speech rate is one of the characteristics of non-native English speech (Munro & Derwing, 1995; Guion, Flege, Liu, & Yeni-Komshian, 2000). Thus, it is possible that listeners utilize speech rate in adapting to Korean learners of English and generalizing their adaptation to novel Korean learners of English. Further, Median F0 and F0 range are analyzed since the results of Experiment 1 suggest that similarity between speakers in training and post-test may play a significant role in generalization. Specifically, previous studies demonstrate that fundamental frequency affects whether listeners perceive speakers as similar or dissimilar (e.g., Perrachione, Furbeck, & Thurston, 2019; Roark, Fend, & Chandrasekaran, 2022). Thus, it is possible that F0 and F0 range affects generalization to a novel speaker.

In general, the three Korean learners of English in the More Accented Condition demonstrate similar acoustic characteristics to the three Korean learners of English in the Less Accented Condition. As the speaker in the post-test is the same in both conditions, the acoustic similarity between speakers in training and post-test is not closer in one condition than the other. Thus, these results suggest that listeners in the Less Accented Condition demonstrate better performance in the post-test than listeners in the More Accented Condition because of having exposure to less accented non-native speech rather than a closer acoustic similarity between speakers in training and post-test than that of the More Accented Condition. Specifically, as shown in Figure 7, the mean speaking rates of the Korean learners of English in the training session of the Less Accented and the More Accented Conditions are the same (3.75 syllables per second). As the mean speaking rate of speakers in the Less Accented Condition is the same as

the speakers in the More Accented Condition, it is not likely that the acoustic similarity between speakers in training and post-test is closer in one condition than another.

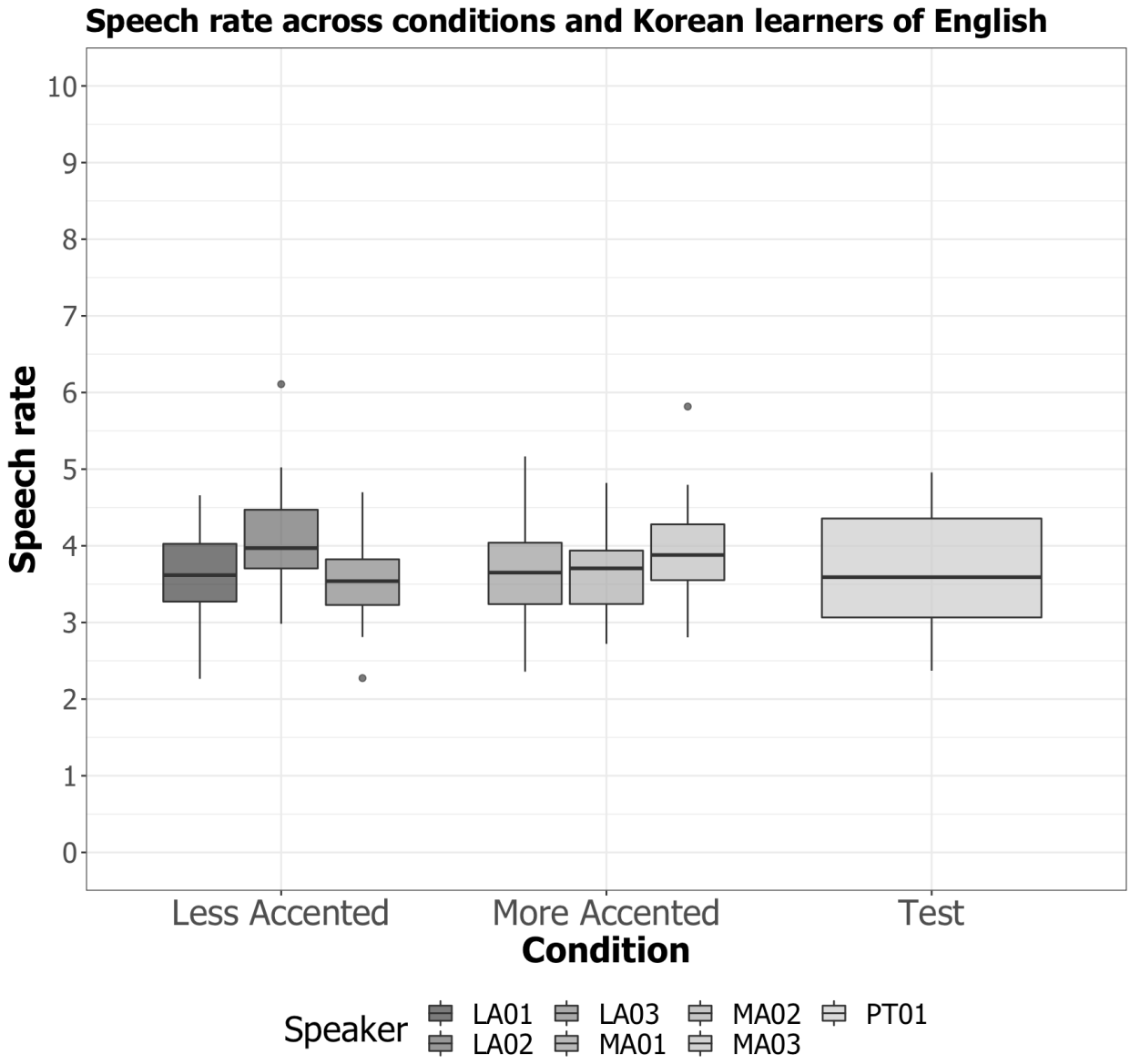

Figure 7. Box plot demonstrating the median speech rate (syllables per second) across conditions and participants.

The speaking rate of individual speakers demonstrates similar patterns. That is, the speaking rate of speakers in the Less Accented Condition are in similar range as the speaking rates of speakers in the More Accented Condition. Taken together, these results suggest that listeners in the Less Accented Condition do not demonstrate better performance than listeners in the More Accented Condition because of acoustic similarity. That is, if the better performance in the Less Accented Condition is driven by acoustic similarity, speakers in the training and post-test of the Less Accented Condition should have closer acoustic similarity than speakers in the More Accented Condition. However, this is not the case, suggesting that the better performance in the Less Accented Condition is driven by the accentedness of non-native speakers.

As shown in Figure 8, the median F0s of the Korean learners of English in the training session of the More Accented Condition and Less Accented Condition are 235.59 Hz and 211.22 Hz, respectively. The median F0 of the Korean learner of English in the post-test of both conditions is 223.75 Hz. As this is an intermediate value between the speakers in training of the More Accented and Less Accented Condition, the differences in median F0 between the speakers in training and post-test are similar between the More Accented and Less Accented Condition.

While the median F0 difference between speakers in training and post-test in the More Accented Condition is similar as the median F0 difference between speakers in training and post-test in the Less Accented Condition, individual speakers in the training session of the More Accented Condition demonstrate a larger variance in median F0. Specifically, the median F0s of the speakers in the training session of the More Accented Condition are 250.18 Hz, 202.77 Hz, and 253.82 Hz, respectively. As the median F0 of the speaker in the post-test is 223 Hz, each speaker in the training session demonstrates 20-30 Hz median F0 difference than the speaker in post-test. On the other hand, the median F0s of the speakers in the training session of the Less

Accented Condition are 199.98 Hz, 212.97 Hz, and 220.71 Hz, respectively and one speaker (LA03) has similar median F0 as the speaker in the post-test.

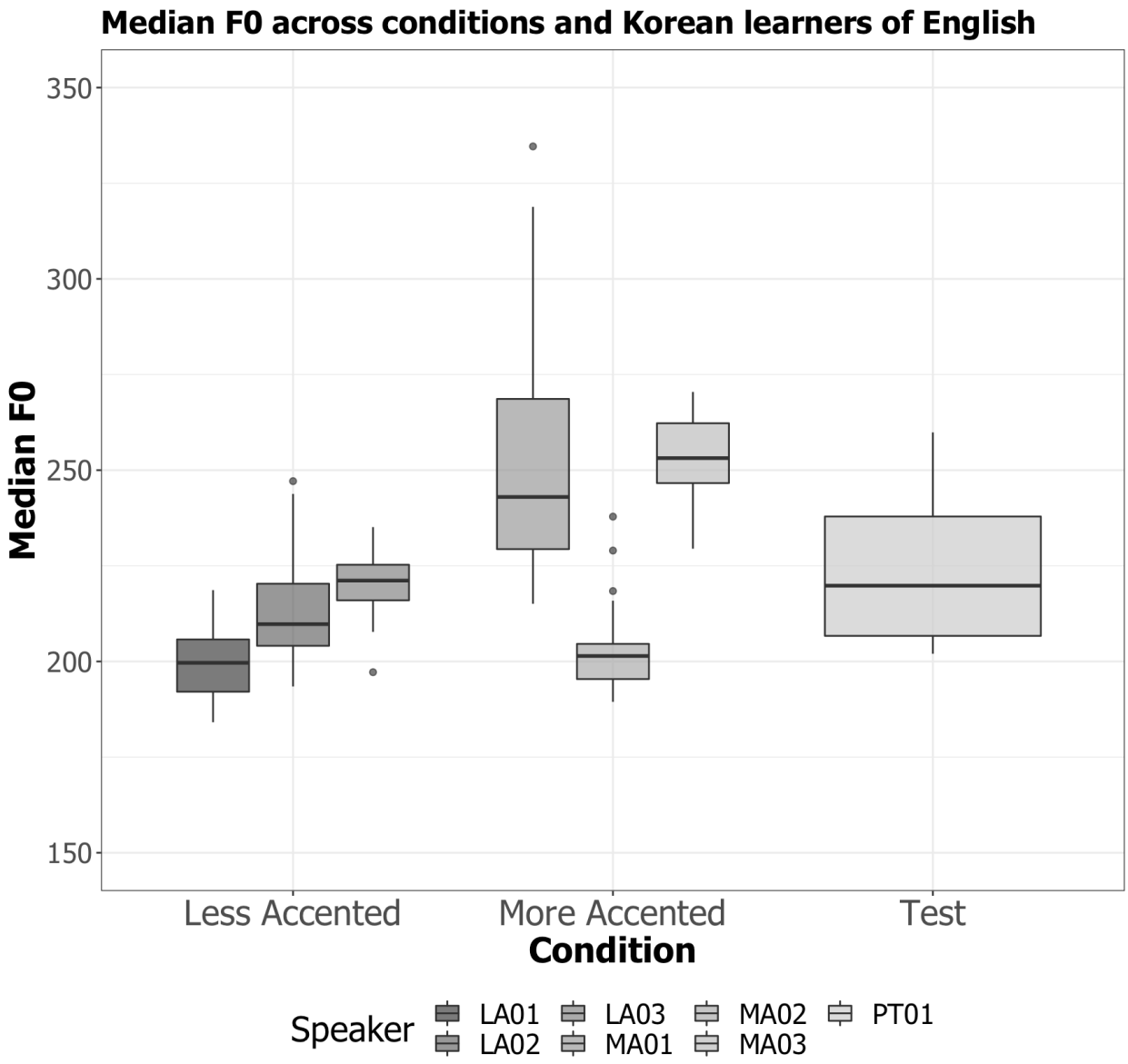

Figure 8. Box plot demonstrating the median F0 across conditions and Korean learners of English.

However, as shown in Figure 9 below, the speaker demonstrates different F0 range than the speaker in post-test. Thus, it is not likely that listeners in the Less Accented Condition demonstrate better performance than listeners in the More Accented Condition because of F0 similarity between speakers in training and post-test.

The mean F0 ranges of the Korean learners of English in the More Accented and Less Accented Conditions also suggest that listeners in the Less Accented Condition do not demonstrate better performance in the post-test than listeners in the More Accented Condition because of acoustic similarity between speakers in training and post-test. Figure 9 shows that the mean F0 range of the Korean learners of English in the More Accented Condition is 46.39 Hz and the mean F0 range of the Korean learners of English in the Less Accented Condition is 34.01 Hz. Further, the mean F0 range of the Korean learners of English in the post-test is 48.87 Hz. Thus, the mean F0 range difference between speakers in training and post-test is closer in the More Accented Condition than Less Accented Condition. The F0 range of each Korean learner of English shows a similar pattern. That is, while all Korean learners of English in the Less Accented Condition demonstrate lower F0 range than the Korean learner of English in the post-test, one talker in the More Accented Condition (MA03) has a similar F0 range as the Korean learner of English in the post-test. Taken together, the F0 range analysis suggests that listeners in the Less Accented Condition do not demonstrate higher intelligibility scores in the post-test than listeners in the More Accented Condition because of similar F0 ranges between speakers in the training session and the post-test. Specifically, if listeners' better performance in the Less Accented Condition were driven by the similar F0 ranges between the speakers in the training session and the post-test, speakers in the training session and post-test in the Less Accented Condition would demonstrate closer F0 ranges than speakers in the More Accented Condition.

However, the F0 range shows the opposite pattern, suggesting that acoustic similarity is not the main factor that facilitates generalization to a novel speaker in the Less Accented Condition..

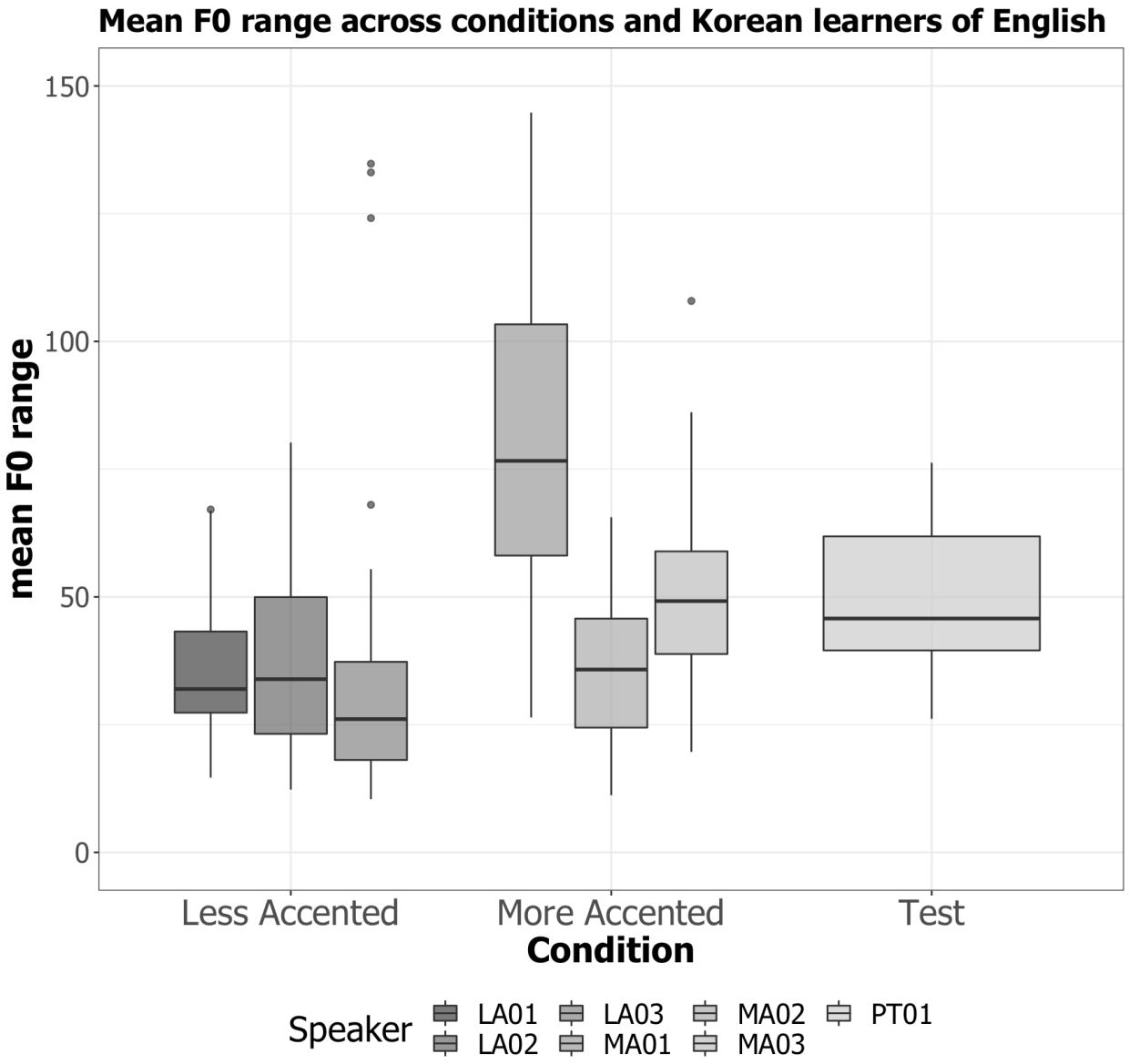

Figure 9. Box plot showing the mean F0 range across conditions and Korean learners of English.

In summary, the acoustic analyses (i.e., speech rate, median F0, and F0 range) of the Korean learners of English suggest that acoustic similarity between speakers in training and post-

test does not drive the better performance in the post-test of the Less Accented Condition than the More Accented Condition. That is, if it is the case that this effect was caused by acoustic similarity between speakers in training and post-test, the acoustic similarity should be closer in the Less Accented Condition than the More Accented Condition. However, this is not the case, suggesting that exposure to different accentedness of non-native speech affects generalization of adaptation.

3.4. Discussion

3.4.1. Summary of findings

The present study examines the effect of accentedness of non-native speech on generalization of adaptation to a novel non-native English speaker. Experiment 2 consists of an intelligibility task and acoustic analyses of the Korean learners of English. The results of the intelligibility task demonstrate that exposure to less accented non-native speech is more helpful for understanding a novel speaker from the same language background than exposure to more accented non-native speech. Specifically, listeners trained with less accented non-native speech perform better in the post-test than listeners trained with more accented non-native speech. Further, the acoustic analyses of the Korean learners of English show that listeners who are trained with less accented non-native speakers do not show higher intelligibility scores in the post-test than listeners who are trained with more accented non-native speakers because of the acoustic similarity between talkers in the training session and post-test. Below, we discuss these results and the implications for our understanding of generalization of adaptation to novel non-native speakers.

3.4.2. The effects of accentedness of non-native speech on generalization of adaptation

The results of the present study show that accentedness of non-native speech affects generalization of adaptation. Specifically, listeners who listen to less accented non-native speech demonstrate higher intelligibility scores in the post-test than listeners who listen to more accented non-native speech. This result suggests that exposure to less accented non-native speech facilitates generalization of adaptation. The results of the present study extend previous findings that learning in easy environments transfer to different items while learning in difficult environments is item specific in visual perceptual learning (Ahissar & Hochstein, 2004). Specifically, the results of the present study extend the findings from visual modality to the speech modality. It is likely that more accented non-native speech is more difficult to process than less accented non-native speech as more accented non-native speech is distinct from the type of speech that listeners are familiar with and less accented non-native speech is more similar to the speech listeners are familiar with. Thus, as in visual perceptual learning, it may be the case that when listeners have exposure to more accented non-native speech, they focus on the acoustic details of speech that do not help generalization to novel speaker.

Interestingly, the results of the present study contrast with some previous findings that show that exposure to items produced in accented speech helps phonetic category retuning. For example, words produced in accented speech facilitate adaptation for native German listeners than word produced in non-accented speech (Grohe & Weber, 2016). It is assumed that phonetic category retuning and adaptation to non-native speech involve similar processes (Kleinschmidt & Jaeger, 2015) and Experiment 1 of the present study demonstrates that generalization of both type of learning indeed share similar underlying mechanisms (i.e., acoustic similarity facilitates generalization). Thus, one would expect a similar effect of accentedness on phonetic category

retuning and adaptation to non-native speech. However, the results of the present demonstrate show that this is not necessarily the case. That is, it is possible that while phonetic category retuning is one of the underlying processes of adaptation to non-native speech, it is not the sole factor that drives adaptation to non-native speech.

3.4.3. The effect of exposure to multiple non-native English speakers on generalization of adaptation

Previous studies demonstrate the benefits of high-variability perceptual training on speech perception (Baese-Berk, Bradlow, & Wright, 2013; Bradlow & Bent, 2008; Lively, Logan, & Pisoni, 1993; Mullennix, Pisoni, & Martin, 1989). For example, training listeners with multiple Mandarin learners of English helps listeners better understand a novel Mandarin learner of English than training listeners with a single Mandarin learner of English. However, while high-variability training may be an important factor for generalization of adaptation, the results of the present study suggest that training listeners with multiple non-native English speakers does not necessarily facilitate generalization of adaptation. Specifically, if exposure to multiple non-native English speakers facilitated generalization of adaptation in the present study, listeners who hear Korean learners of English in the training session (i.e., listeners in the More Accented and Less Accented Conditions) would demonstrate higher intelligibility scores in the post-test than listeners who hear native English speakers in the training session. However, listeners in the More Accented Condition who hear sentences read by three Korean learners of English demonstrate similar intelligibility scores in the post-test as listeners in the Control Condition who hear sentences read by three native English speakers.

It is possible that listeners in the More Accented Condition do not demonstrate generalization of adaptation because the training session is shorter than previous studies (e.g., Bradlow & Bent, 2008; Baese-Berk, Bradlow, & Wright, 2013). That is, while previous studies train listeners for two days and present 160 sentences total, the present study consists of only one training session and presents 120 sentences in total. However, it is not likely that the shorter training session disrupts generalization of adaptation since listeners in the Less Accented Condition are trained in the same training session as listeners in the More Accented Condition and demonstrate better performance in the post-test than listeners in the More Accented Condition. Thus, it is likely that high-variability training is not always effective for generalization of adaptation to novel non-native speakers.

Indeed, previous studies have shown that high-variability training is not uniformly helpful for all listeners in perceptual learning. For example, learners with weak perceptual abilities for pitch are disrupted with learning a novel phonological contrast based on pitch when trained in a high-variability condition than when they are trained in a low-variability training condition (Perrachione, Lee, Ha, & Wong, 2011). The authors suggest that in high-variability training, listeners with weak perceptual abilities are not able to attend to the cues that are most informative for perceptual learning. Similarly, it is possible that listeners who are trained with more accented non-native speech have difficulty attending to the cues that are helpful for generalization to a novel speaker than listeners who are trained with less accented non-native speech.

3.4.3. Alternative explanation

One alternative explanation for the results of the present study that listeners in the Less Accented Condition demonstrate higher intelligibility scores in the post-test than listeners in the More Accented Condition is the acoustic similarities between the non-native speakers in the training session and post-test. As suggested in Experiment 1, acoustic similarity between the speakers in the training session and post-test may play a significant role in generalization of adaptation. Thus, it is possible that listeners in the Less Accented Condition show higher intelligibility scores in the post-test than listeners in the More Accented Condition because acoustic similarity between the speakers in training and post-test is closer in the Less Accented Condition than the More Accented Condition.

However, the acoustic analyses of the Korean learners of English suggest that this is not the case. Specifically, the acoustic characteristics (i.e., speech rate, median F0, and mean F0 range) of the speakers in the training session and post-test are not more similar in the Less Accented Condition than the More Accented Condition suggesting that the results of Experiment 2 are not driven by acoustic similarity between speakers in the training session and post-test. While the present study includes acoustic analyses of the Korean learners of English, the acoustic features that are analyzed in the present study do not include all possible acoustic features that may affect generalization of adaptation. Thus, there is a possibility that the features analyzed in the present study do not capture similarity between speakers.

3.4.4. Conclusion

The present study examines how accentedness of non-native speech affects generalization of adaptation. Specifically, the study consists of an intelligibility task and acoustic analyses of

the Korean learners of English. The intelligibility task demonstrates that training listeners with non-native speakers with less accented non-native speech facilitates generalization of adaptation. Further, the acoustic analyses show that this result is not driven by the acoustic similarities between the Korean learners of English in the training session and the post-test. The results of the present study suggest that exposure to multiple non-native English speakers does not uniformly facilitate generalization of adaptation and generalization of adaptation is likely to be driven by specific characteristics of non-native speech (i.e., accentedness of non-native speech).

IV. EFFECT OF LINGUISTIC EXPERIENCE ON GENERALIZATION OF ADAPTATION

4.1. Introduction

Most previous studies focus on the effects of speaker characteristics on perceptual learning. For example, acoustic similarity (i.e., duration of vowel preceding a stop, closure duration of stop, and length of burst and aspiration of stop) between talkers in training and post-test and accentedness of non-native speech are factors that constrain generalization of phonetic category retuning (Grohe & Weber, 2016; Xie & Myers, 2017). On the other hand, listener characteristics receive less attention compared to speaker characteristics. That is, most previous studies on generalization of adaptation recruit participants that have do not have significant exposure to non-native speech in general (e.g., Baese-Berk, Bradlow, & Wright, 2013; Bradlow & Bent, 2008) or speakers that are not fluent with the target non-native accented language (e.g., Sidaras, Alexander, & Nygaard, 2009).

However, considering previous studies that demonstrate training listeners with non-native English speakers in the lab facilitates generalization of adaptation (e.g., Baese-Berk, Bradlow, & Wright, 2013; Bradlow & Bent, 2008; Sidaras, Alexander, & Nygaard, 2009), it is likely that listeners' linguistic experience with non-native speakers affects generalization of adaptation to novel non-native speakers. Indeed, previous studies demonstrate that listeners that have frequent interaction with non-native English speakers are better at understanding a non-native English speaker than listeners who do not have frequent interaction with non-native English speakers (Laturnus, 2018). Thus, in order to better understand the factors that affect generalization to a novel speaker, the present study examines how listeners' linguistic experience affect

generalization of adaptation, with a focus on how extended linguistic experience with different numbers of non-native accents affects generalization of adaptation.

Further, it is important to investigate the effect of listeners' lifetime linguistic experience with non-native speakers because there is a growing population that do not speak English as their first language. Specifically, according to a United States Census Bureau report, over 60 million people speak a language other than English at home. Among the 60 million people, around 25 million reported that they are not able to speak English "very well" (U.S. Census Bureau, 2015). Thus, it is likely that there are a number of native English listeners that have frequent interaction with non-native English speakers and it is important to include these populations in studies that examine generalization of adaptation to non-native English speakers for ecological validity as well. That is, while examining native English listeners with no frequent interaction with non-native English speakers provides meaningful information about the underlying mechanisms of generalization of adaptation, it is important to examine listeners with frequent interaction with non-native English speakers to have an accurate picture of factors that affect generalization to a novel speaker.

4.1.1. Effects of lifetime experience with non-native English speakers

As discussed above, training native English listeners in the lab with non-native English speakers' speech helps listeners better understand the speaker they are trained with (Clarke & Garrett; 2004; Bradlow & Bent, 2008; Xie et al., 2018) and even a novel non-native English speaker (e.g., Baese-Berk, Bradlow, & Bent, 2013; Bradlow & Bent, 2008; Sidaras, Alexander, & Nygaard, 2009;). Further, sleep between training sessions facilitates generalization of adaptation to a novel speaker (e.g., Xie, Earle, & Myers, 2017). If this is the case, it is likely that

listeners' lifetime experience with non-native English speakers helps generalization of adaptation to a novel non-native English speaker. Specifically, listeners' lifetime experience involves interacting with multiple non-native English speakers and sleep that would likely consolidate listeners adaptation to non-native English speech and its generalization.

However, only a few previous studies on generalization of adaptation examine how listeners' lifetime linguistic experience with non-native English speakers affects generalization of adaptation to a novel non-native English speaker. For example, Laturnus (2018) demonstrate that native English listeners with frequent interaction with non-native English speakers are better at understanding a non-native English speaker than listeners with no frequent interaction with non-native English speakers. Specifically, without having any training with non-native English speakers in the lab, listeners who have lifetime interaction with non-native English speakers perform better in an intelligibility task than listeners who do not have lifetime experience with non-native English speakers. This result suggests that listeners' lifetime linguistic experience with non-native English speakers helps listeners' comprehension of a novel non-native English speaker. However, it is less clear whether previous experience with non-native English speakers affects adaptation to non-native English speakers from a novel language background and its generalization.

Further, different types of extended linguistic experience may have different effects on generalization of adaptation to a novel non-native English speaker. Previous studies demonstrate that while training listeners with multiple non-native English speakers facilitates generalization of adaptation, training listeners with a single non-native accent does not help generalization of adaptation (e.g., Bradlow & Bent, 2008; Sidaras, Alexander, & Nygaard, 2009). These findings suggest that different types of linguistic exposure have different consequences on generalization

of adaptation to a novel non-native English speaker. Specifically, Laturnus (2020) suggests that exposure to multiple non-native English speakers helps listeners learn the common characteristics of non-native speech, which facilitates generalization of adaptation as a result. If this is the case, it is possible that different types of lifetime linguistic experience with non-native English speakers also have different effects on generalization of adaptation to a novel non-native English speaker. Therefore, the present study asks whether different types of extended experience with non-native English speakers have different effects on generalization of adaptation to a novel non-native English speaker.

4.1.2. Current study

In the current study, we examine whether listeners' extended experience with non-native English speakers affects generalization of adaptation to a novel non-native English speaker. Specifically, we ask whether listeners' extended experience with non-native English speakers facilitates generalization of adaptation and whether different types of extended experiences have different effects on generalization of adaptation. To answer the questions, we recruit native English speakers from three different populations. Specifically, we recruit native English listeners who have extended experience with multiple non-native English accents, listeners who have extended experience with a single non-native English accent, and listeners who do not have frequent interaction with non-native English speakers. The listeners with extended experience with multiple non-native English speakers have family members that are non-native English speakers and had frequent interaction with non-native English speakers at school or in their community.

For the listeners with extended experience with a single non-native accent, Spanish heritage speakers are recruited. Spanish heritage speakers are recruited for this population because we assume that Spanish heritage speakers have extended linguistic experience with at least one group of non-native English speakers (i.e., English speakers whose first language is Spanish). The other reason we recruited Spanish heritage speakers is to control for the non-native accent that listeners have frequent interaction with. Specifically, if acoustic similarities between non-native English speakers in the training session and post-test affect generalization of adaptation, the degree of generalization of adaptation may differ depending on the non-native accent that listeners have frequent interaction with. For example, if it is the case that Korean-accented English have similar characteristics as Japanese-accented English and distinct characteristics than Mandarin-accented English, listeners who have interaction with Japanese-accented English are likely to better understand Korean learners of English than listeners who have interaction with Mandarin-accented English would. Thus, to control for the effect of experience with different non-native English accents on generalization of adaptation, listeners with extended experience with a specific single non-native accent (i.e., Spanish-accented English) are recruited in the present study. Further, Spanish heritage speakers are recruited because there is a large population of speakers who speak Spanish at home in the U.S. Specifically, over 34 million U.S. residents over the age of five and older speak Spanish at home (U.S. Census Bureau, 2010). Thus, recruiting participants are less challenging than other populations that is smaller in size (e.g., U.S. residents who use Japanese at home). The last group of listeners are listeners who do not have frequent interaction with non-native English speakers. For this group, listeners who do not have family members that are non-native English speakers

and who do not have frequent interaction with non-native English speakers at school or in their community are recruited.

With regards to the effect of extended experience with non-native English speakers on generalization of adaptation, it is possible that extended linguistic experience facilitates generalization of adaptation. Specifically, if training native English listeners with non-native English speech in the lab and sleep facilitate generalization of adaptation, it is likely that listeners' lifetime experience with non-native English speakers have a similar effect because lifetime experience with non-native English speakers involve an extended exposure to non-native speech and sleep. If this is the case, listeners who have extended experience with non-native English speakers would demonstrate higher intelligibility scores in the intelligibility task than listeners with no frequent interaction with non-native English speakers.

On the other hand, it is also possible that listeners' lifetime experience with non-native English speakers disrupts generalization to a novel speaker. Listeners may have speaker models of certain speaker groups and this model may become less malleable as the model becomes larger (cf. Lev-Ari, 2017). That is, it may be the case that native listeners who do not have prior experience with non-native English speakers demonstrate rapid adaptation and generalization to novel non-native speakers (e.g., Bradlow & Bent, 2008; Clarke & Garrett, 2004; Sidaras, Alexander, & Nygaard, 2008) since the listeners have small and malleable model of non-native speakers. On the contrary, listeners who have extended experience with non-native speakers may have a larger and less malleable model of non-native speakers than listeners who do not have linguistic experience with non-native speakers. If this is the case, listeners who have extended experience with non-native speakers would demonstrate lower intelligibility scores in the post-test than listeners who do not have experience with non-native speakers.

Further, the types of extended experience with non-native English speakers may have different effects on generalization of adaptation. It is possible that extended experience with multiple non-native accents is more helpful for generalization of adaptation than extended experience with a single non-native accent. Specifically, previous studies that train listeners with non-native speakers in the lab show that exposure to multiple non-native speakers facilitates generalization of adaptation (e.g., Bradlow & Bent, 2008; Sidaras, Alexander, & Nygaard, 2009). Laturnus (2020) suggests that exposure to multiple non-native English speakers helps listeners learn the common characteristics of non-native speech and facilitates generalization of adaptation as a result. If listeners' lifetime experience with non-native English speakers has similar effects on generalization of adaptation, listeners who have extended experience with multiple non-native accents would demonstrate better performance in the intelligibility task than listeners who have extended experience with a single non-native accent. However, it is also possible that extended lifetime experience with non-native English speakers have different effects on generalization of adaptation than short training with non-native English speakers in the lab. That is, it is possible that extended exposure to a single non-native English accent provides the listeners with enough variability to learn the characteristics of non-native English speech. Then, listeners who have extended experience with a single non-native English accent would demonstrate similar intelligibility scores in the intelligibility task as listeners who have extended experience with multiple non-native English accents.

4.2. Methods

4.2.1. Participants

75 native English speakers between 18 and 40 years old (29 female, 46 male) participated in this experiment. Participants were recruited from the University of Oregon Psychology and Linguistics subject pool and from Prolific. Participants recruited from the University of Oregon Psychology and Linguistics subject pool were paid partial course credits for their participation and participants recruited from Prolific were paid $7.50 for their participation. Experiment 3 was different than Experiments 1 and 2 in terms of the selection criteria. Specifically, while Experiments 1 and 2 recruited native English speakers with no frequent interaction with non-native English speakers, Experiment 3 recruited native English speakers with different linguistic experience. In Experiment 3, there were three target populations including: 1) native English speakers who had extensive experience with multiple non-native accents, 2) native English speakers who had extensive experience with a single non-native accent, and 3) native English speakers who had very limited experience with non-native accents.

As in Experiments 1 and 2, participants recruited from the University of Oregon Psychology and Linguistics subject pool were not screened for participation. However, participants were not included in the data analysis if participants were non-native English speakers, participants reported a history of speech or hearing disorder, and participants did not use headphones during the experiment. Participants recruited from Prolific were invited to participate in the experiment if they met the requirement of the experiment. That is, in the description of the study, it was described that participants were allowed to participate in the study if they were native English speakers with no history of speech or hearing disorder and if they could use headphones for the experiment.

4.2.2. Materials

As in Experiments 1 and 2, items were BKB sentences read by Korean learners of English and the sentences were drawn from OSCAAR. While participants in Experiments 1 and 2 heard sentences read by different speakers depending on the experimental condition, participants in Experiment 3 were all exposed to the same training session and post-test regardless of the condition they were assigned to. That is, all participants in Experiment 3 participated in the Less Accented Condition of Experiment 2. The Less Accented Condition of Experiment 2 was chosen to ensure that participants show generalization of adaptation after the training session, since participants in the Less Accented Condition of Experiment 2 demonstrated a strong generalization effect. 120 BKB sentences were used in the training session and 16 BKB sentences were used in the post-test. The sentences in the training session and post-test were leveled to a fixed RMS amplitude of 73 dB. All sentences were mixed with speech-shaped noise at a signal-to-noise ratio of -5dB.

4.2.3. Design

The experiment aims to examine whether different types of linguistic experience affect generalization of adaptation. Previous studies have shown that different types of exposure in the lab (i.e., exposure to a multiple non-native accents or exposure to a single non-native accent) have different effects on generalization of adaptation (e.g., Baese-Berk, Bradlow, & Wright, 2013). In the present study, we ask whether listeners' extended experience with multiple non-native accents or extended experience with a single non-native accent affects generalization of adaptation differently. Therefore, participants from three different populations were recruited.

The first group (i.e., Multiple-accent Exposure Condition) consisted of listeners who have extended experience with multiple non-native accents. The second group (i.e., Single-accent Exposure Condition) consisted of listeners who have extended experience with a single non-native accent. Specifically, Spanish heritage speakers were recruited. Spanish heritage speakers were recruited in the Single-accent Exposure Condition for two reasons. First, we assumed that that Spanish heritage speakers have frequent interaction with at least one group of non-native English speakers (i.e., English speakers whose first language is Spanish). Specifically, we aimed to control for the non-native accent that listeners had frequent with to avoid the type of non-native accent being a confounding factor. The second reason was to overcome the challenge of recruiting participants among the COVID-19 pandemic. Since the experiment was conducted during the COVID-19 pandemic, recruiting native English listeners who had extended experience with any type of non-native accent was a challenge. As we aimed to recruit listeners who had experience with a single non-native accent to avoid the type of non-native accent being a confounding factor, we recruited from a population that has a large number in the United States (over 40 million U.S. resident over the age of five and older speak Spanish at home; U.S. census, 2015). The third group (i.e., No Exposure Condition) consisted of listeners that did not have frequent interaction with non-native English speakers. The No Exposure condition served as a control condition.

To ensure participants met the linguistic experience requirements, participants completed a language experience questionnaire and an intelligibility task. Participants were asked to answer questions about their language experience that were used to determine participants' linguistic experience. The questions asked to determine the linguistic experience conditions included: (1) whether participants had frequent interaction with family members that are non-native English

speakers, (2) whether participants had frequent interaction with non-native English speakers in high school, (3) whether participants had frequent interaction with non-native English speakers in elementary school, and (4) whether participants had frequent interaction with non-native English speakers over the past year. Further, if participants answered that they had frequent interaction with non-native speakers in questions 1 to 4, they were asked to answer the first languages of the speakers they interacted with. Based on the results of the language experience questionnaire, participants were assigned to one of the three linguistic experience conditions: 1) Multiple-accent Exposure Condition, 2) Single-accent Exposure Condition, and 3) No Exposure Condition.

Participants were assigned to Multiple-accent Exposure condition if: (1) participants interacted frequently with family members that were non-native English speakers and (2) participants had frequent interaction with non-native English speakers in elementary school and high school.

For the Single-accent Exposure condition, Spanish heritage speakers were recruited. It was assumed that Spanish heritage speakers have extended linguistic experience with at least a single non-native English accent (i.e., Spanish accented English). To ensure participants indeed had an extended experience with only a single non-native accented speech, participants were assigned to the condition if: (1) participants had frequent interaction with family members that were non-native English speakers and (2) did not have frequent interaction with non-native English speakers other than Spanish learners of English in elementary and high school.

Participants were assigned to the No Exposure Condition if: (1) participants did not have family members that are non-native English speakers, (2) participants had limited or no

interaction with non-native English speakers in elementary and high schools, (3) participants did not frequently interact with non-native English speakers over the past year.

Participants in the three conditions participated in the same task. That is, participants in the three conditions heard and transcribed the same set of sentences read by the same Korean learner of English in the training session. Further, the sentences that participants heard and transcribed in the post-test were same in all conditions.

4.2.4. Procedure

As in Experiments 1 and 2, the experiment was conducted online using Qualtrics (https://www.qualtrics.com). Since Experiment 3 assigned participants to different experimental conditions based on their language experience, two separate studies were run on Prolific. Specifically, the first study was designed to assign participants to one of the three experimental conditions based on participants' linguistic experience. In the first study, native English speakers who did not have a history of speech or hearing disorder and who were able to use headphones during the experiment were invited to participate in the study. The participants were asked to read and sign a consent form. Then, the participants filled out the language experience questionnaire described in section 4.2.3. Participants were also asked to provide their Prolific ID so that the eligible participants could be invited to the second study (i.e., the main task including the training session and the post-test). The first study took approximately 10 minutes.

Based on the results of the first study, three lists of participants who met the linguistic experience conditions (i.e., extensive experience with multiple non-native accents, extensive experience with a single non-native accent, and limited experience with non-native accents) were created. Then, participants in the three lists were invited to participate in the second study. To

ensure only participants that met the selection criteria were invited to participate in the second study, the second study was only visible on Prolific to participants that were included in the three lists of eligible participants.

The second study consisted of a training session and a post-test and did not include the linguistic experience questionnaire since participants already filled out the linguistic experience questionnaire in the first study. Participants in the second study were asked to read and sign a consent form to participate in the study. Then, they were asked to wear their headphones and transcribe three repetitions of a short English sentence to ensure that participants could hear the items. The sentence was the same sentence that was used in Experiments 1 and 2. Participants were also asked to adjust the volume to a comfortable level. After finishing the sound check, participants were given a description of the intelligibility task and transcribed a practice sentence. The practice sentence was the same sentence used in Experiments 1 and 2. The sentence was a BKB sentence that was not presented in the main task and the sentence was read by a native English speaker. The main task (i.e., training session and post-test) was the same as the Less Accented Condition of Experiment 2. That is, participants transcribed 120 BKB sentences in the training session and 16 BKB sentences in the post test.

4.2.5. Analysis

As in Experiments 1 and 2, participants' transcription from the intelligibility task were unnested using a script within the R computing program and each target word was scored automatically as correct or incorrect using Autoscore (Borrie, Barrett, & Yoho, 2019) to measure generalization of adaptation after manually being aligned in Microsoft Excel. Obvious spelling mistakes and homophones were scored as correct, and target words did not need to be transcribed

in the order in which they were spoken to be scored as correct. Results were analyzed with a Bayesian mixed-effects logistic regression model within the R computing program. As in Experiments 1 and 2, a Bayesian approach to data analysis was used because it was possible that participants in different conditions have similar performance in the post-test of the intelligibility task (i.e., a null result). Since a null result does not provide evidence for the null hypothesis, this result would be difficult to interpret. However, it is possible to have a meaningful interpretation of the null results with a Bayesian approach to regression modeling. As in Experiments 1 and 2, we fitted a Bayesian logistic mixed model to predict participants' performance on the post-test as a function of language experience condition (extensive experience with multiple non-native accents, extensive experience with a single non-native accent, and limited experience with non-native accents). Condition was Helmert coded to compare: (1) No Exposure Condition vs Multi-accent and Single-accent Exposure Conditions and (2) Multi-accent Exposure vs Single-accent Exposure Conditions. The model included by-item random intercepts and slopes for Condition and random intercepts for participants and used weakly informative priors. That is, we used a Student-*t* prior distribution with a mean of 0, degree of freedom of 1, and a scale of 2.5 for the fixed effects and a Cauchy distribution with a center of 0 and scale of 2 for the random effects (Gelman, Jakulin, Pittau, & Su, 2008).

4.3. Results

Figure 10 shows listeners' intelligibility scores in the post-test. As shown in Figure 10, listeners in the Multiple-accent Exposure (box on the right) and Single-accent Exposure Conditions (box in the middle) demonstrate lower intelligibility scores in the post-test than listeners in the No Exposure Condition (box on the left). This finding suggests that having

extensive experience with non-native English speakers disrupts generalization of adaptation to novel non-native English speakers.

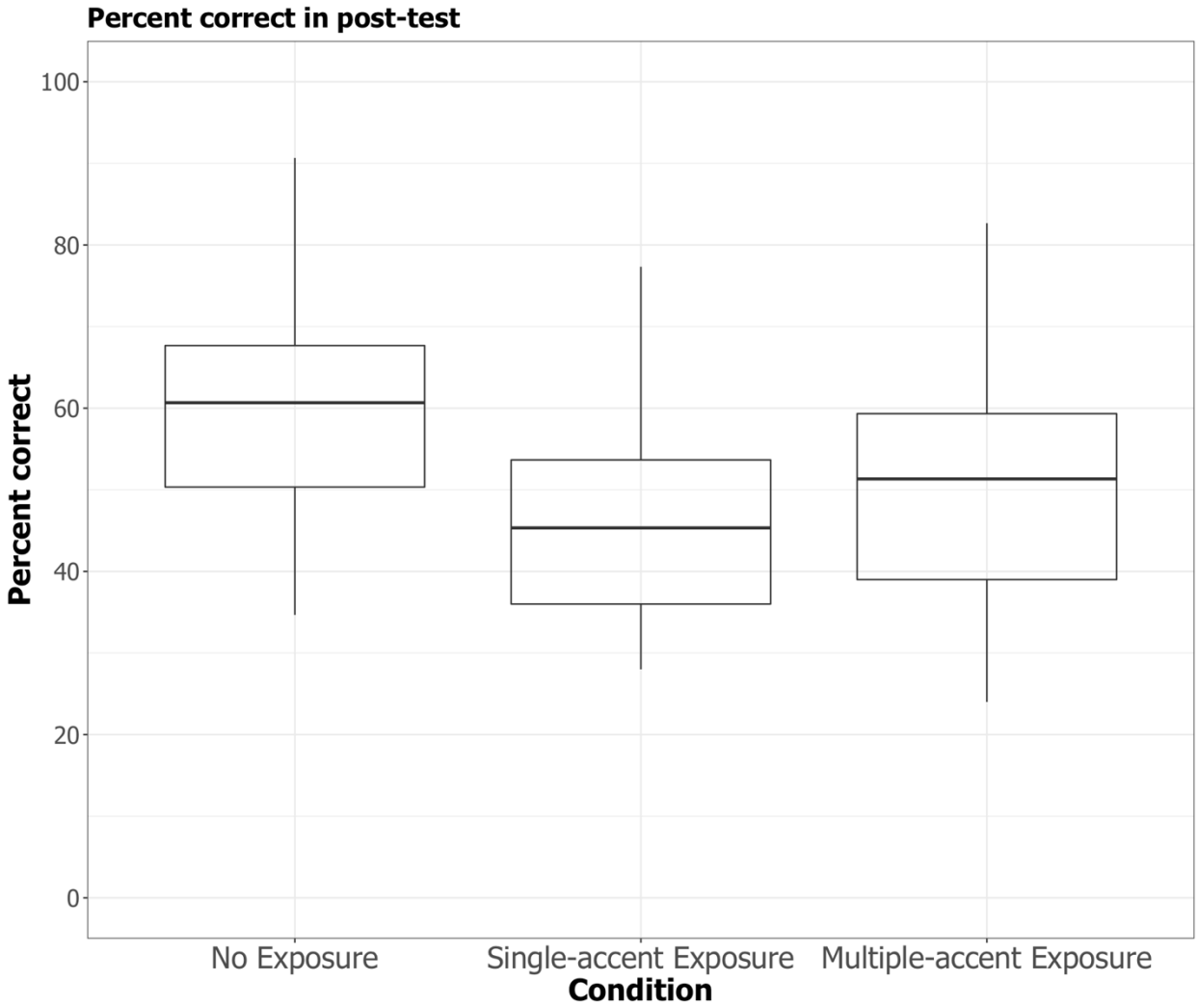

Figure 10. Box plot showing the percent correct on the post-test of the intelligibility task as a function of condition (No Exposure, Single-accent Exposure, and Multiple-accent Exposure Conditions). Listeners in the Single-accent and Multiple-accent Exposure Conditions demonstrate lower intelligibility scores than listeners in the No Exposure Condition.

Specifically, if it is the case that extended experience with non-native English speakers facilitated listeners' generalization of adaptation to a novel Korean learner of English, listeners in the Multiple-accent Exposure and Single-accent Exposure Conditions would demonstrate better intelligibility scores in the post-test than listeners in the No Exposure Condition.

Further, listeners in the Multiple-accent Exposure Condition and the Single-accent Exposure Condition demonstrate similar intelligibility scores in the post-test. This finding suggests that extensive experience with non-native English speakers may disrupt native English listeners generalization of adaptation to a novel non-native English speaker regardless of the number of non-native English accents native English listeners experienced. That is, if different types of linguistic experience had different effects on generalization of adaptation, listeners in the Multiple-accent Exposure and Single-accent Exposure Conditions would demonstrate different intelligibility scores in the post-test.

The Bayesian mixed-effect logistic regression model confirms this trend. Specifically, there is a 95% probability that the highest density interval of the mean intelligibility difference of listeners in the Multiple-accent and Single-accent Exposure Conditions and listeners in the No Exposure Condition is larger than zero, suggesting that linguistic experience with non-native speakers may disrupt generalization to a novel speaker. Further, there is less than a 50% probability that the highest density interval of the mean intelligibility difference of listeners in the Multiple- and Single-accent Exposure Conditions does not include zero, suggesting that the type of exposure does not affect generalization to a novel speaker.

Figure 11 shows listeners' intelligibility scores in the training session. As shown in Figure 11, listeners in all three conditions (i.e., No Exposure, Multiple-accent Exposure, and Single-accent Exposure Conditions) demonstrate improvements in intelligibility scores across

the six blocks of the training session. Specifically, the listeners demonstrate higher intelligibility scores at the end of the training session (i.e., Block 6) than the beginning of the training session (i.e., Block 1).

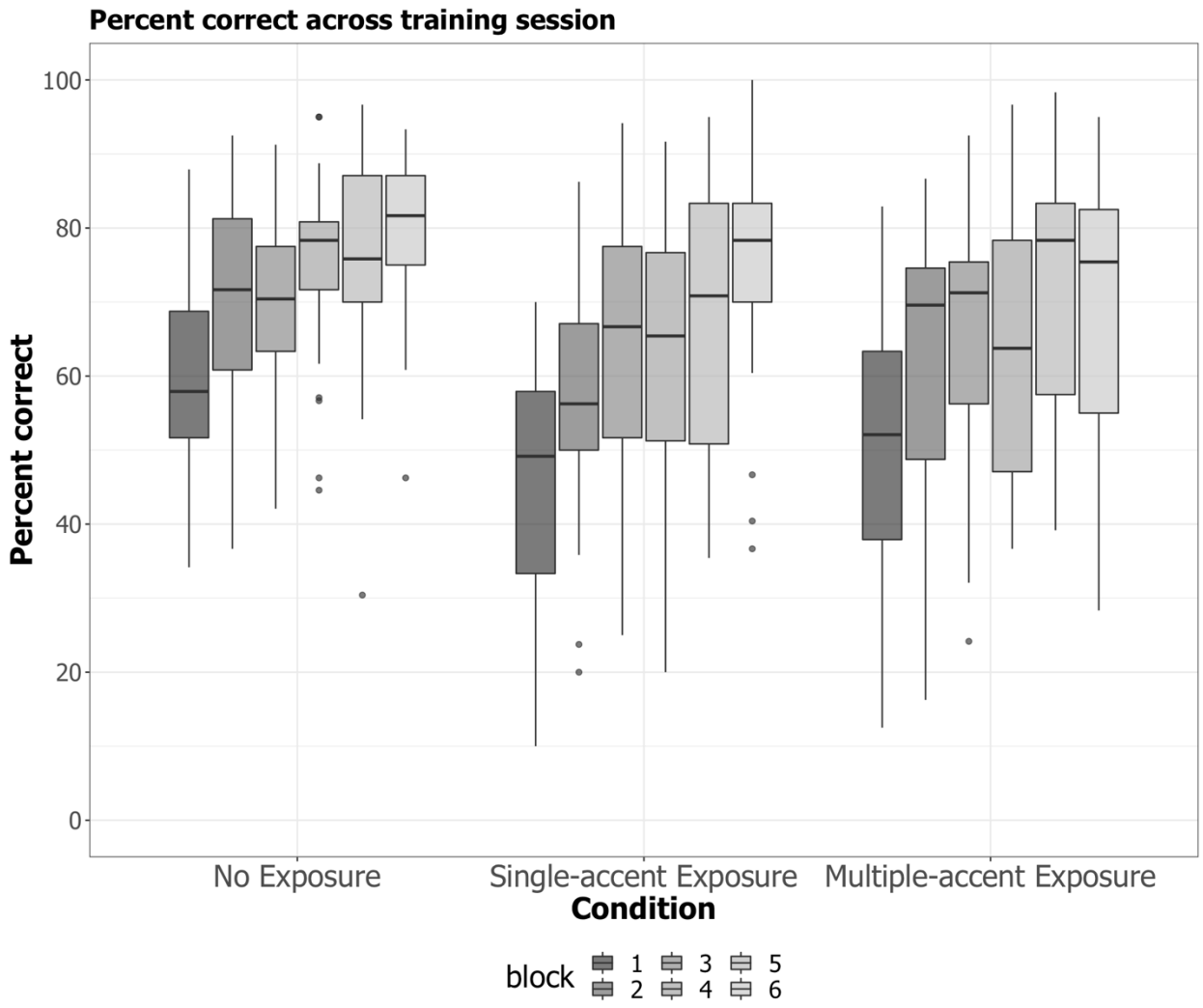

Figure 11. Box plot showing the percent correct on the training session of the intelligibility task as a function of condition (No exposure, Single-accent Exposure, and Multiple-accent Exposure Conditions) and block. Listeners in all three conditions demonstrate increase in intelligibility scores across blocks.

In summary, the results of this study suggest that listeners' linguistic experience with non-native English speakers disrupts generalization to a novel speaker. That is, listeners in the Multiple-accent and Single-accent Exposure conditions demonstrate lower performance in the post-test than listeners in the No exposure condition. Moreover, the results also suggest that the type of extended experience with non-native accents does not affect generalization to a novel speaker. Specifically, listeners in the Multiple-accent and Single-accent Exposure conditions show similar intelligibility scores in the post-test. The implications of these findings are explained in section 4.4.

4.4. Discussion

4.4.1. Summary of findings

The present study examines the effect of native English listeners' lifetime linguistic experience on generalization of adaptation to a Korean learner of English. Specifically, the study asks whether different types of linguistic experience (i.e., lifetime experience with multiple non-native English accents, lifetime experience with a single non-native English accent, and no experience with non-native English accents) affect generalization of adaptation. The results of the present study demonstrate that listeners' extended experience with non-native English speakers disrupts generalization of adaptation to a novel Korean learner of English. That is, listeners who have extended experience with non-native English speakers demonstrate lower intelligibility scores in the post-test than listeners who do not have frequent interaction with non-native English speakers. Further, the results also suggest that listeners who have extended experience with non-native English speakers, the type of linguistic experience does not affect generalization of adaptation. Specifically, listeners who have extended experience with multiple

non-native English accents demonstrate similar intelligibility scores in the post-test as listeners who have extended experience with a single non-native English accent. Below, we discuss the results and the implications for our understanding of generalization of adaptation to novel non-native speakers.

4.4.2. The effect of extended experience on generalization of adaptation

Previous studies demonstrate that native English listeners become better at understanding non-native English speech after short training sessions in the lab (e.g., Clarke & Garrett, 2004; Bradlow & Bent, 2008; Baese-Berk, Bradlow, & Wright, 2013; Xie et al, 2018). More importantly, listeners generalize their adaptation to novel non-native English speakers after listening to multiple non-native English speakers (Bradlow & Bent, 2008) and listeners generalize to novel non-native English speakers from novel language backgrounds after listening to multiple non-native English speakers from different language backgrounds (Baese-Berk, Bradlow, & Wright, 2013).

Since short training sessions in the lab facilitate generalization of adaptation, it is possible that listeners with extensive experience with non-native English speakers would be better at generalizing their adaptation to a novel non-native English speaker than listeners with no frequent interaction with non-native English speakers. However, the results of the present study suggest that this is not the case. Specifically, native English listeners who have frequent interaction with non-native English speakers (i.e., Multiple-accent Exposure and Single-accent Exposure Conditions) demonstrate lower intelligibility scores in the post-test than listeners who do not have frequent interaction with non-native English speakers (i.e., No Exposure Condition)

suggesting that extended experience with non-native English speakers disrupts generalization of adaptation to a novel non-native English speaker.

While the results of the present study may seem contradictory to previous studies that demonstrate training listeners with multiple non-native English speakers in the lab facilitates generalization of adaptation to novel non-native English speakers (e.g., Bradlow & Bent, 2008; Sidaras, Alexander, & Nygaard, 2009), the results of the present study do not necessarily contradict with previous findings. Participants in the previous studies (e.g., Bradlow & Bent, 2008; Baese-Berk, Bradlow, & Wright, 2013) and the present study have different language backgrounds and the different language backgrounds may affect generalization of adaptation. Specifically, while participants in the previous studies are native English listeners with no frequent interaction with non-native English speakers, participants in the Multiple-accent Exposure and Single-accent Exposure Conditions of the present study are listeners who have extended experience with multiple non-native English speakers from different language backgrounds and multiple non-native English speakers from the same language background, respectively. The difference in linguistic experience may have different consequences for generalization of adaptation. For example, previous studies have suggested that listeners generate models of speaker groups and use the model in speech perception (Kleinschmidt & Jaeger, 2015). Specifically, the Ideal Adaptor Framework (Kleinschmidt & Jaeger, 2015) posits that listeners generate speaker models, use the speaker models in speech perception, and update speech models as they interact with speakers. Within this model, each new speech input would contribute less to the model as the model becomes larger. If this is the case, listeners' models of non-native English speakers would be more malleable for listeners who do not have frequent interaction with non-native English speakers than listeners who have extended experience with

non-native English speakers. That is, listeners that have no frequent interaction with non-native English speakers would be better at adapting and generalizing their adaptation to a novel non-native speaker than listeners who have frequent interaction with non-native English speakers. Indeed, previous studies demonstrate that listeners with smaller social networks have more malleable linguistic representations than listeners with bigger social networks (e.g., Lev-Ari, 2017). Similarly, we suggest that listeners with extended experience with non-native English speakers have less malleable speaker models of non-native English speakers than listeners with no frequent interaction with non-native English speakers, and more robust models of non-native speakers may initially disrupt generalization of adaptation to novel non-native English speakers.

Further, previous studies demonstrate that native English listeners that have prior exposure to numerous non-native English speakers are better at understanding a novel non-native English speaker than listeners that have limited exposure to non-native English speakers (Laturnus, 2018). While this result may seem contradictory to the results of the present study, it is not necessarily the case. Specifically, while listeners in the present study are trained with non-native English speakers in the training session and tested with a novel non-native English speaker in the post-test, the listeners in Laturnus (2018) do not have a training session. Thus, the present study and Laturnus (2018) use slightly different paradigms. Further, in the present study, listeners who have extended experience with non-native English speakers do not have frequent interaction with Korean learners of English. Thus, listeners in the present study do not have experience with the non-native accent that they are trained and tested with in the training session and post-test. On the other hand, it is possible that listeners in Laturnus (2018) have exposure to the non-native accent that they are tested with. If this is the case, listeners would indeed demonstrate better performance at understanding a novel non-native English speaker than

listeners that have limited exposure to non-native English speakers. Thus, it is not the case that extended experience with non-native English speakers uniformly disrupts perception of a novel non-native English speaker. Rather, it is likely that extended experience with non-native English speakers facilitates perception of a novel non-native English speaker if the non-native English speaker shares language backgrounds with the non-native English accents the listeners have experience with.

4.4.3. The effect of type of lifetime experience with non-native English speakers on generalization of adaptation

The results of the present study demonstrate that different types of linguistic experience (i.e., extended experience with multiple non-native English accents and extended experience with a single non-native English accent) do not affect generalization of adaptation to a Korean learner of English. Specifically, while listeners in the Multiple-accent Exposure Condition have extended experience with multiple non-native accents and listeners in the Single-accent Exposure Condition have extended experience with a single non-native accent, listeners in both conditions demonstrate similar intelligibility scores in the post-test.

This result is interesting as previous studies on generalization of adaptation suggest that exposure to multiple non-native English accents facilitates generalization of adaptation (e.g., Baese-Berk, Bradlow, & Bent, 2013). If short training sessions with multiple non-native accented speakers facilitate generalization to a novel non-native speaker, one would expect extended experience with multiple non-native accents help generalization to a novel non-native speaker as well. However, the present study demonstrate that this is not the case. We suggest that exposure to multiple non-native accents does not uniformly facilitate adaptation and its

generalization to a novel non-native speaker. Specifically, the effect of exposure to multiple non-native English accents on generalization may interact with the length of linguistic experience that listeners had with non-native speakers.

4.4.4. Alternative explanation

One alternative explanation for the results of the present study that listeners who have extended experience with non-native English demonstrate lower intelligibility scores in the post-test than listeners who have limited experience with non-native English speakers is that factors other than linguistic experience may affect generalization to a novel speaker. For example, it is possible that in our recruitment, we recruited participants who differ from one another in more ways than just language background. For example, socioeconomic status, cognitive skills, or myriad other factors may differ across the groups. However, the present study did not ask participants for information other than linguistic experience and follow up studies are required to examine this possibility.

4.4.5. Conclusion

The current study examines how listeners' linguistic experience affects generalization of adaptation to a novel non-native English speaker. Specifically, the current study examines whether listeners' lifetime experience with non-native English speakers facilitates generalization of adaptation and whether the types of lifetime experience with non-native English speakers affect generalization of adaptation. The results of the present study show that listeners' lifetime experience with non-native English speakers disrupts generalization to a novel non-native speaker. Further, the types of lifetime linguistic experience do not have an effect on

generalization of adaptation for listeners who had extended linguistic experience with non-native speakers. The results of the study suggest that exposure to multiple non-native English speakers does not necessarily facilitate generalization to a novel speaker and the effect of exposure to multiple non-native English speakers on generalization of adaptation is affected by length of experience.

V. CONCLUSION

This dissertation sought to better understand the mechanisms underlying speech perception by examining the factors that affect generalization of adaptation to novel non-native speakers. Specifically, the dissertation aims to investigate how acoustic characteristics and talker information interact and when exposure to variability is beneficial for speech perception. We examine how acoustic similarity between speakers and their talker information affect generalization of adaptation. We also examine how accentedness of non-native speech affects generalization to a novel speaker. Further, we investigate how extended linguistic experience with non-native English speakers affects generalization of adaptation. In this chapter, we summarize the major findings of each of the three studies and discuss the novel contributions of the study. Further, we present implications of the current study for communication involving non-native English speakers and discuss directions for future work.

5.1. Summary of the current research

5.1.1. Main findings of the three studies

The first study examines whether acoustic similarity between non-native English speakers affects generalization of adaptation and what role talker information plays in generalization to a novel speaker. The results of the first study suggest that acoustic similarity between speakers in training and post-test may be an important factor in generalization of adaptation. Specifically, the results of the present study show that if speakers in training and post-test have very similar acoustic characteristics, listeners who perceive a talker change between training and post-test demonstrate similar performance in the post-test as listeners who are trained with the same speaker in training and post-test. Further, listeners in these two

conditions together demonstrate better performance in the post-test than listeners who hear an actual different speaker in the post-test than the training session. Taken together, these results suggest that acoustic similarity between speakers may play a significant role in generalization of adaptation.

The second study examines how accentedness of non-native speech affects generalization of adaptation to a novel non-native English speaker. We find that exposure to more accented non-native speech disrupts generalization of adaptation. That is, listeners who are trained with Korean learners of English with less accented non-native speakers are better at transcribing a novel Korean learner of English than listeners who are trained with more accented non-native speakers. Further, while listeners who are trained with more accented and less accented non-native speakers demonstrate similar performance in the post-test as listeners who are trained with native English speakers, we show in a post-hoc analysis that listeners who are trained with less accented non-native speakers are better in the post-test than listeners who are trained with native English speakers.

While the first two experiments examine how speaker characteristics affect generalization to a novel speaker, the third experiment focuses on how characteristics of listeners affect generalization of adaptation. In the third experiment, we investigate how native English listeners' lifetime linguistic experience affects generalization of adaptation to a novel non-native speaker. The third experiment demonstrates that native listeners' linguistic experience with non-native English speakers disrupts generalization of adaptation to a novel talker. Specifically, listeners who have extended exposure to non-native speakers demonstrate lower intelligibility scores in the post-test than listeners who do not have linguistic experience with non-native speakers. Further, the results also shows that the type of linguistic experience does not affect

generalization of adaptation for listeners who have extended linguistic experience. That is, Listeners show similar intelligibility scores in the post-test whether they have exposure to multiple non-native English accents or a single non-native English accent.

5.1.2. Novel contributions of the current research

The current work provides novel contributions that inform how acoustic characteristics and talker information interact in speech perception and what role variability plays in speech perception. Specifically, we suggest that acoustic similarity may play an important role in generalization of adaptation, at least in the early stages. Further, we suggest that being exposed to non-native speech that is too distinct from the speech listeners are familiar with may disrupt generalization. We also suggest that representation of non-native speakers may become less malleable with more experience with non-native speakers.

5.1.2.1. Effects of acoustic characteristics and talker information on generalization of adaptation

These results provide a better understanding of how acoustic characteristics and talker information interact in speech perception. Previous studies often assume that adaptation involves similar processes as phonetic category retuning (Kleinschmidt & Jaeger, 2015). Specifically, it is assumed that the underlying processes of adaptation consist of the underlying processes of phonetic category retuning. The current study suggests that generalization of adaptation and phonetic category may involve similar processes. Specifically, previous studies on phonetic category retuning suggest that acoustic similarity between speakers or items facilitates generalization of perceptual learning (e.g., Eisner & McQueen, 2005; Reinisch & Holt, 2014; Xie & Myers, 2017). For example, Eisner & McQueen (2005) demonstrate that listeners do not

generalize phonetic category retuning to a novel speaker. However, if the phonetic category from training is spliced into the novel speaker's speech, listeners demonstrate generalization of phonetic category retuning. These results suggest that acoustic similarity is a crucial factor in generalization of phonetic category retuning and talker information may play a less important role if the target phonetic categories are acoustically similar. As the results of the present study suggest that acoustic similarity between speakers may play an important role in generalization of adaptation, at least in the early stages, it is possible that generalization of adaptation and phonetic category retuning involve similar processes.

As the results of the present study suggest that generalization of adaptation may be driven by acoustic similarity in the early stages, it is possible that generalization of adaptation to a non-native speaker utilizes mechanisms that are general to speech perception rather than specific to non-native speech. Specifically, previous studies demonstrate that generalization of phonetic category retuning is constrained by similarity between speakers or items (e.g., Eisner & McQueen, 2005; Reinisch & Holt, 2014; Xie & Myers, 2017) and that generalization occurs to both native and non-native accents after training (e.g., Eisner, Melinger, & Weber, 2013; Kraljic & Samuel, 2016), suggesting that the underlying mechanisms of phonetic category retuning are general to speech perception. Similarly, we suggest that generalization of adaptation generalization of adaptation may occur not only for non-native speech but other types of speech that listeners are not familiar with, as long as talkers that listeners are exposed to are acoustically similar to a novel talker. Therefore, it is possible that generalization of adaptation to other types of speech that listeners may be unfamiliar with (e.g., regional-accented, dysarthric, noise-vocoded, time-compressed speech) is constrained by acoustic similarity, at least in the early stages (i.e., immediately after being exposed to unfamiliar speech).

As the present study investigates both acoustic characteristics and talker information on speech perception, the results have implications for how talker information is utilized in speech perception. We show that listeners become better at transcribing a novel non-native speaker after being exposed to a single non-native speaker even when there is a talker change between training and post-test, as long as the acoustic characteristics are similar between the talkers in training and post-test. This finding has implications for the argument that talker information is tightly connected to a talker's speech and talker information is used for speech perception instead of being discarded (Goldinger, 1996; Levi, Winter, & Pisoni, 2011; Nygaard & Pisoni, 1998; Nygaard, Sommers, & Pisoni, 1994). For example, listeners who learn talker identity (i.e., learn names that are associated with talkers' voices) are better at word identification than listeners who do not learn talker identity (Nygaard & Pisoni, 1998), suggesting that certain aspects of talker information are connected with talkers' linguistic properties. The results of the current study provide insight into how talker information is intertwined with linguistic properties and suggest that listeners' reliance on talker information may be down-weighted when speakers in training and post-test have similar acoustic characteristics. If it is the case that aspects of talker identity (e.g., whether the talker is a female or a male) are tightly connected to linguistic properties of a speaker, it is likely that listeners would not generalize to a novel speaker who has similar acoustic characteristics as the speaker in training but is perceived as a different speaker. That is, even if the speakers have similar acoustic characteristics, the perceived change in gender may disrupt generalization to the novel speaker if gender information is utilized for speech perception regardless of acoustic characteristics of speakers. However, current results show that listeners generalize to a novel speaker even when there is a talker change, as long as the talkers are acoustically similar. Thus, it is possible that listener may rely less on talker information in

generalization to a novel speaker if acoustic characteristics of talkers give sufficient information for speech perception (i.e., speakers in training and post-test have very similar acoustic characteristics).

5.1.2.2. Effect of accentedness of non-native speech on generalization of adaptation

The current study suggests that acoustic similarity between speakers may be a driving factor of generalization of adaptation. However, it is not likely that the acoustic similarity between non-native speakers in naturally produced speaking situations is as close as the acoustic similarity between the non-native speakers in the present study; speakers in the present study have very similar acoustic characteristics because one of the speakers is artificially created using Praat with the aim of being acoustically similar to the other speaker. Even though speakers in general are not likely to have the same degree of acoustic similarity as the speakers in the present study, listeners demonstrate generalization to a novel non-native speaker (Bradlow & Bent, 2008; Sidaras, Alexander, & Nygaard, 2009).

Thus, while generalization to a novel speaker occurs between speakers who have similar acoustic characteristics, an exact, or very close, acoustic match between speakers may not be necessary for generalization of adaptation. Specifically, acoustic similarity between speakers may not be the only factor that facilitates generalization of adaptation, and it is possible that listeners also learn general patterns of non-native speech and utilize the information to better understand a novel speaker, as suggested in previous studies (e.g., Baese-Berk, Bradlow, & Wright, 2013; Laturnus, 2020). If this is the case, being exposed to multiple non-native speakers would facilitate generalization to a novel speaker, as being exposed to multiple speakers would help listeners learn general patterns of non-native speech. Further, if generalization is facilitated

by learning general patterns of non-native speech, being exposed to more accented non-native speech could facilitate generalization as more accented non-native speech would highlight general characteristics of non-native speech. The results of the current work provide a novel contribution to this hypothesis by demonstrating that the effect of exposure to multiple speakers is affected by the type of exposure. Specifically, listeners exposed to less accented non-native speakers demonstrate better performance in transcribing a novel non-native speaker than listeners exposed to more accented non-native speakers. Thus, it is possible that listeners who have exposure to more accented non-native speech do not perform well in the post-test because more accented non-native speech has more distinct characteristics than less accented non-native speech. These distinct characteristics of more accented non-native speech may be more harmful than beneficial for generalization as processing speech that have distinct characteristics from the speech listeners are familiar with could be difficult, which may make it difficult for listeners to generalize to a novel speaker. This result is consistent with the argument that learning in easy environments is generalized to other items while learning in difficult environments is item specific (Reverse Hierarchy Theory; Ahissar & Hochstein, 2004; Ahissar et al., 2009). Specifically, the Reverse Hierarchy Theory posits that when listeners have difficulty processing input, listeners focus on low-level information and search for the most informative input. While focusing on low-level information of the input may help listeners process the input, they may lose access to high-level information. As a result, listeners may correctly understand the input, but they may not be able to generalize what they learned to a novel talker. The present study shows that this is indeed the case in perception of non-native speech. That is, more accented non-native speech has more distinct characteristics than less accented non-native speech. As listeners are not familiar with these distinct characteristics, processing more accented non-native speech

could be more difficult for listeners than processing less accented non-native speech. Thus, listeners who are exposed to more accented non-native speech are likely to focus on low-level information of speech rather than high-level information of speech that is likely more helpful for generalization of adaptation.

5.1.2.3. Effect of linguistic experience on generalization of adaptation

The results of the third experiment of the present study investigates how extended linguistic experience with non-native speakers affects generalization of adaptation and demonstrate that linguistic experience indeed affects generalization to a novel speaker. Specifically, the experiment shows that listeners who have extended experience with a single non-native accent and multiple non-native accents demonstrate poorer performance in transcribing a novel non-native speaker than listeners who do not have linguistic experience with non-native speakers. This result initially seems to contradict previous studies on generalization of adaptation that demonstrate that short training sessions in the lab help listeners adapt and generalize their adaptation to a novel non-native speaker (Baese-Berk, Bradlow, & Wright, 2013; Bradlow & Bent, 2008; Clarke & Garrett, 2004; Sidaras, Alexander, & Nygaard, 2009; Xie et al., 2018). That is, if short training sessions in the lab help listeners learn the characteristics of non-native speech and help listeners understand a novel non-native speaker, one would expect that lifetime experience with non-native speakers outside of the lab helps learning the characteristics of non-native speech and generalizing to a novel non-native speaker.

However, most previous studies on adaptation and generalization to non-native speech examines listeners who do not have extended experience with non-native speakers (e.g., Baese-Berk, Bradlow, & Wright, 2013; Sidaras, Alexander, & Nygaard, 2009; Xie & Myers, 2017), and

it is possible that generalization of adaptation has an inverse relationship with linguistic experience. Specifically, listeners' perception of non-native speech may be more malleable when they have little or no exposure to non-native speakers, as shown in previous studies that demonstrate that listeners who have no experience with non-native speakers successfully adapt and generalize to a novel non-native speaker (e.g., Clarke & Garrett, 2004; Bradlow & Bent, 2008; Sidaras, Alexander, & Nygaard, 2009; Baese-Berk, Bradlow, & Wright, 2013; Xie et al*.,* 2018). However, as listeners get more experience with non-native speakers, their perception of non-native speech may become less malleable. As a result, it would be less likely for listeners to adapt and generalize to novel talkers, as shown in the current study. These results provide support for the argument that listeners have speaker models that are updated as listeners interact with other speakers (e.g., Kleinschmidt & Jaeger, 2015; Lev-Ari, 2017). That is, listeners may have a speaker model for non-native speakers in general and this model may be updated with each interaction with other non-native speakers. For example, Lev-Ari (2017) suggests that interaction with listeners has different effects on speaker models depending on the size of the speaker model (i.e., the amount of interaction listeners have with other speakers). For listeners who do not have much linguistic experience with non-native speakers, the model will be updated with each new interaction with a non-native speaker since each new input has more weight for a smaller model (i.e., listeners with no linguistic experience with non-native speakers) than for a larger model (i.e., listeners with extended experience with non-native speakers). If this is the case, listeners who do not have extended linguistic experience with non-native speakers may adapt and generalize their adaptation to non-native speakers. On the other hand, for listeners who have extended experience with non-native speakers, the model may be less malleable than the speaker model of listeners who have no experience with non-native speakers. Therefore,

extended linguistic experience with non-native speakers could disrupt adaptation and its generalization to a novel non-native speaker.

Overall, the present study provides contributions that inform how speaker and listener factors affect generalization to a novel non-native speaker. In terms of speaker factors, we suggest that acoustic similarity between speakers constrains generalization of adaptation a novel speaker, at least in the early stages of generalization. Further, we suggest that exposure to non-native speech that are too different from non-native speech that listeners are familiar with may disrupt generalization to a novel speaker. In terms of listener factors, we show that extended experience with non-native speakers disrupts generalization and suggest that listeners with extended linguistic experience may have a less malleable representation of non-native speakers than listeners who have no experience with non-native speakers.

5.2. Future directions

5.2.1. Does linguistic experience have a gradual effect on generalization of adaptation?

The results of the present study show that listeners' linguistic experience with non-native speakers disrupts generalization to a novel speaker. However, it is not clear whether linguistic experience disrupts generalization regardless of the length of linguistic experience or if linguistic experience has a gradual effect on generalization. If it is the case that listeners have speaker models of non-native speakers that are updated with new input, as suggested above, the length of linguistic experience with non-native speakers may have a gradual effect on generalization of adaptation. That is, it may not be the case that linguistic experience with non-native speakers is uniformly harmful for adaptation and its generalization. For example, listeners who have relatively shorter linguistic experience with non-native speakers than the listeners in the current

study (i.e., listeners who have extended experience with a single and multiple non-native accents) may demonstrate better performance in adapting and generalizing to a novel non-native speaker than listeners in the current study.

Thus, a future study may investigate how different amount of linguistic experience affects generalization of adaptation to a non-native speaker. The study may include amount of linguistic experience as a factor and examine how different amount of linguistic experience affects generalization to a novel non-native speaker. Specifically, the study could investigate the effect of participants' length of experience with non-native accents on generalization by including linguistic experience as a continuous variable based on the information provided in a linguistic experience questionnaire. Investigating how the length of linguistic experience affects listeners' adaptation and its generalization to a novel speaker would help us better understand the mechanisms underlying listeners' perception of non-native speech. Specifically, the results would help explain the seemingly contrasting results of previous studies that demonstrate that short exposure to non-native speakers in the lab facilitates generalization to a novel speaker (e.g., Bradlow & Bent, 2008; Sidaras, Alexander, & Nygaard, 2009) and the results of the current study that show that short exposure to non-native speakers disrupts generalization for listeners who have extended experience with non-native speakers. If listeners indeed have speaker models for non-native speakers as discussed above, the amount of linguistic experience is likely to have a gradual effect on generalization of adaptation. That is, speaker models of non-native accents will be more malleable for novel non-native accents as listeners have less experience with non-native accents than more experience. If this is the case, previous linguistic experience with non-native speakers will not uniformly disrupt generalization to a novel speaker and it is possible for

listeners who have prior experience with non-native speakers to demonstrate generalization of adaptation.

5.2.2. Does linguistic experience uniformly disrupt generalization of adaptation?

While the current study demonstrates that extended linguistic experience disrupts generalization to a novel speaker, it should be noted that extended linguistic experience with non-native speakers is not necessarily detrimental for perception of non-native speech. Specifically, previous studies show that native listeners benefit from extended linguistic experience with non-native speakers. For example, native English listeners who have greater lifetime experience with non-native English speakers are better at understanding novel non-native English speakers than listeners who have less experience with non-native English speakers (Laturnus, 2018). Thus, it is possible that listeners learn common characteristics of non-native speech that facilitate listeners' perception of a novel non-native speaker, as suggested in previous studies (e.g., Baese-Berk, Bradlow, & Wright, 2013; Laturnus, 2018, 2020). If this is the case, listeners who have extended linguistic experience with non-native speakers may benefit from their experience. That is, if the target non-native accent shares similar characteristics as the non-native accents the listeners are familiar with, having linguistic experience may, in fact, facilitate generalization to a novel speaker. Thus, a future study may investigate how the similarity between non-native accents listeners are familiar with and a novel accent affect generalization to a novel speaker to have a better understanding how listeners benefit from exposure to variable non-native speakers. Specifically, the study could manipulate the similarity between the first languages of the non-native speaker in training and post-test. For example, one group of listeners could be trained with non-native speakers whose first language is similar to the first language of

the non-native speaker in post-test. Further, another group of listeners could be trained with non-native speakers whose first language has distinct characteristics than the first language of the non-native speaker in post-test. Then, the study could examine how similarity between non-native accents listeners are familiar with and a novel accent affect generalization to a novel speaker. The results of this study would shed light on how extended experience with non-native speakers affects speech perception. That is, it is possible that listeners have speaker models of non-native speakers that become less malleable as the models become larger, as discussed above. If this is the case, the model will disrupt adaptation and generalization to a novel non-native accent if the novel non-native accent is dissimilar to the non-native accent listeners are familiar with. However, if the non-native accent that listeners are familiar with and the novel accent are similar, extended experience with the non-native accent will not necessarily disrupt generalization of adaptation as listeners are able to utilize the model for processing the novel non-native accent.

5.2.3. How does sleep affect listeners with linguistic experience?

In the methods sections of the current study, we note that the current study differs from previous studies (e.g., Baese-Berk, Bradlow, & Wright, 2013; Bradlow & Bent, 2008; Sidaras, Alexander, & Nygaard, 2009) in terms of the number of training sessions. Specifically, listeners in the current study are trained for one day while most previous studies train listeners for multiple days. Thus, the participants in the present study do not sleep overnight between sessions while participants in most other studies do. Previous studies demonstrate that sleep facilitates generalization of learning (Earle & Myers, 2015a; Earle & Myers, 2015b; Xie, Earle, & Myers, 2017). For example, Xie, Earle, & Myers (2017) show that listeners who sleep overnight show

better performance in categorizing unfamiliar speech sounds produced by a novel talker than listeners who do not sleep. The authors suggest that sleep may help listeners store the salient features of non-native speech and facilitate generalization to a novel speaker. Similarly, it is possible that sleep facilitates generalization of adaptation to a novel talker for listeners who have extended linguistic experience. That is, listeners who have extended linguistic experience with non-native speakers do not immediately benefit from training with non-native speakers, as shown in the current study. However, sleep may help the listeners learn more abstracted characteristics of the target non-native accent. To test this hypothesis, a future study may use a similar paradigm used in the present study but could separate training and testing by 24 hours so that participants could sleep overnight before participating in the post-test.

5.3. Conclusion

In the three studies of this dissertation, we examine speaker-related and listener-related factors that affect adaptation and its generalization to a novel non-native English speaker. In terms of speaker-related factors, we suggest that acoustic similarity between speakers may affect adaptation and generalization to a novel non-native speaker. Specifically, even when there is a perceived talker change between training and post-test, listeners demonstrate generalization of adaptation if non-native speakers in training and post-test are acoustically similar (i.e., same acoustic characteristics other than median F0). Further, we show that listeners demonstrate generalization of adaptation when they are trained with less accented non-native speakers than more accented non-native speakers. In terms of listener-related factors, we demonstrate that native listeners' lifetime linguistic experience with non-native English speakers disrupts generalization of adaptation to a novel non-native speaker. This effect remains consistent

whether listeners have extended linguistic experience with a single non-native English accent or multiple non-native English accents.

The current research provides a unique contribution to the research on adaptation and its generalization. Specifically, the findings of the current study suggest that acoustic similarity between speakers may play a significant role in generalization to a novel speaker and the role of talker information may be down-weighted when the speakers in training and post-test have very similar acoustic features. The results also suggest that exposure to multiple non-native speakers does not necessarily facilitate generalization to a novel speaker and that speech that is distinct from the type of speech listeners are familiar with may disrupt generalization to a novel speaker. Further, the results suggest that extended linguistic experience may be harmful for generalization of adaptation. As a whole, this dissertation provides insight into how acoustic characteristics and talker information interact in speech perception and the types of variability that may be helpful for speech perception.

APPENDICES

APPENDIX A

List of 40 training and 16 testing BKB sentences (marked as 'training' and 'test') used in Chapter 2 (Experiments 1A and 1B). Keywords were used for intelligibility scoring and are underlined.

Type	Sentence	Type	Sentence
training	They are buying some bread.	training	The girl lost her doll.
training	He played with his train.	training	The cook is making a cake.
training	The mailman shut the gate.	training	The dogs went for a walk.
training	The bag fell to the ground.	training	The lady stayed for lunch.
training	The rain came down.	training	The driver waited by the corner.
training	The ice cream was pink.	training	They finished the dinner.
training	He cut his finger.	training	The policeman knows the way.
training	She is taking her coat.	training	The little girl was happy.
training	The police chased the car.	training	The cow gave some milk.
training	The lady is making a toy.	training	The boy got into bed.
training	The glass bowl broke.	training	The two farmers are talking.
training	They say some silly things.	training	A fish swam in the pond.
training	The lady wore a coat.	test	Potatoes grow on the ground.
training	The children are walking home.	test	He is cleaning his car.
training	He needed his vacation.	test	They waited for an hour.
training	Milk comes in a carton.	test	The plant is hanging above the door.
training	The man cleaned his shoes.	test	The mother heard the baby.
training	The boy is running away.	test	The truck climbed the hill.
training	The room is getting cold.	test	They are drinking tea.
training	The wife helped her husband.	test	An old woman was at home.
training	The old man is worried.	test	They broke all the eggs.
training	A boy ran down the path.	test	The kitchen window was clean.
training	She spoke to her son.	test	The big fish got away.
training	Lemons grow on trees.	test	She is helping her friend.
training	He found his brother.	test	The children washed the plates.
training	Some animals sleep on straw.	test	The mailman comes early.
training	The jelly jar was full.	test	The sign showed the way.
training	They are kneeling down.	test	The grass is getting long.

APPENDIX B

List of 40 training and 16 testing BKB sentences (marked as 'training' and 'test') used in Chapters 3 and 4 (Experiments 2 and 3). Keywords were used for intelligibility scoring and are underlined.

Type	Sentence	Type	Sentence
training	The car engine is running.	training	They are crossing the street.
training	They are looking at the clock.	training	Some animals sleep on straw.
training	The bag fell to the ground.	training	The jelly jar was full.
training	The boy did a handstand.	training	They are kneeling down.
training	The truck carried fruit.	training	The cook is making a cake.
training	The ladder is near the door.	training	The child grabbed the toy.
training	They had a lovely day.	training	The mud stuck on his shoe.
training	The ball went into the goal.	training	The candy shop was empty.
training	The old gloves are dirty.	training	She is washing her dress.
training	The thin dog was hungry.	training	The driver waited by the corner.
training	She is taking her coat.	training	They finished the dinner.
training	The police chased the car.	training	He wore his yellow shirt.
training	A mouse ran down the hole.	test	The fruit came in a box.
training	The little baby is sleeping.	test	The husband brought some flowers.
training	They are watching the train.	test	They are playing in the park.
training	The glass bowl broke.	test	The mouse found the cheese.
training	They say some silly things.	test	They waited for one hour.
training	The children are walking home.	test	The big dog was dangerous.
training	The man cleaned his shoes.	test	The strawberry jam was sweet.
training	They ate the lemon pie.	test	The plant is hanging above the door.
training	The boy is running away.	test	The children are all eating.
training	She drinks from her cup.	test	The boy has black hair.
training	The room is getting cold.	test	The mother heard the baby.
training	The wife helped her husband.	test	The truck climbed the hill.
training	The old man is worried.	test	The angry man shouted.
training	A boy ran down the path.	test	They are drinking tea.
training	The house had a nice garden.	test	Mother opened the drawer.
training	She spoke to her son.	test	An old woman was at home.

REFERENECES CITED

Adank, P., & Janse, E. (2010). Comprehension of a novel accent by young and older listeners. *Psychology and Aging*, *25*(3), 736.

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, *8*(10), 457–464.

Ahissar, M., Nahum, M., Nelken, I., & Hochstein, S. (2009). Reverse hierarchies and sensory learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1515), 285–299.

Baese-Berk, M. M., Bradlow, A. R., & Wright, B. A. (2013). Accent-independent adaptation to foreign accented speech. *The Journal of the Acoustical Society of America*, *133*(3), EL174–EL180.

Baker, W., Trofimovich, P., Flege, J. E., Mack, M., & Halter, R. (2008). Child—Adult differences in second-language phonological learning: The role of cross-language similarity. *Language and Speech*, *51*(4), 317–342.

Bamford, J., & Wilson, I. (1979). Methodological considerations and practical aspects of the BKB sentence lists. *Speech-Hearing Tests and the Spoken Language of Hearing-Impaired Children*, 148–187.

Banks, B., Gowen, E., Munro, K. J., & Adank, P. (2015). Cognitive predictors of perceptual adaptation to accented speech. *The Journal of the Acoustical Society of America*, *137*(4), 2015–2024.

Bench, J., Kowal, Å., & Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British Journal of Audiology*, *13*(3), 108–112.

Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America*, *114*(3), 1600–1610.

Best, C. T., Shaw, J. A., Mulak, K. E., Docherty, G., Evans, B. G., Foulkes, P., Hay, J., Al-Tamimi, J., Mair, K., & Wood, S. (2015). *Perceiving and adapting to regional accent differences among vowel subsystems.* ICPhS.

Biersack, S., Kempe, V., & Knapton, L. (2005). *Fine-tuning speech registers: A comparison of the prosodic features of child-directed and foreigner-directed speech*. Ninth European Conference on Speech Communication and Technology.

Blumstein, S. E., & Stevens, K. N. (1979). Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *The Journal of the Acoustical Society of America*, *66*(4), 1001–1017.

Blumstein, S. E., & Stevens, K. N. (1980). Perceptual invariance and onset spectra for stop consonants in different vowel environments. *The Journal of the Acoustical Society of America*, *67*(2), 648–662.

Boersma, P., & Weenink, D. (2022). *Praat: doing phonetics by computer* [Computer program]. Version 6.2.17, retrieved 23 August 2022 from http://www.praat.org/

Borrie, S. A., Barrett, T. S., & Yoho, S. E. (2019). Autoscore: An open-source automated tool for scoring listener perception of speech. *The Journal of the Acoustical Society of America*, *145*(1), 392–399.

Borrie, S. A., Lansford, K. L., & Barrett, T. S. (2017). Generalized adaptation to dysarthric speech. *Journal of Speech, Language, and Hearing Research*, *60*(11), 3110–3117.

Borrie, S. A., McAuliffe, M. J., Liss, J. M., Kirk, C., O'Beirne, G. A., & Anderson, T. (2012). Familiarisation conditions and the mechanisms that underlie improved recognition of dysarthric speech. *Language and Cognitive Processes*, *27*(7–8), 1039–1055.

Borrie, S. A., McAuliffe, M. J., Liss, J. M., O'Beirne, G. A., & Anderson, T. J. (2012). A follow-up investigation into the mechanisms that underlie improved recognition of dysarthric speech. *The Journal of the Acoustical Society of America*, *132*(2), EL102–EL108.

Borrie, S. A., & Schäfer, M. C. (2015). The role of somatosensory information in speech perception: Imitation improves recognition of disordered speech. *Journal of Speech, Language, and Hearing Research*, *58*(6), 1708–1716.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English/r/and/l: Long-term retention of learning in perception and production. *Perception & Psychophysics*, *61*(5), 977–985.

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, *106*(2), 707–729.

Bradlow, A. R., Blasingame, M., & Lee, K. (2018). Language-independent talker-specificity in bilingual speech intelligibility: Individual traits persist across first-language and second-language speech. *Laboratory Phonology*, *9*(1).

Bradlow, A. R., Kim, M., & Blasingame, M. (2017). Language-independent talker-specificity in first-language and second-language speech production by bilingual talkers: L1 speaking rate predicts L2 speaking rate. *The Journal of the Acoustical Society of America*, *141*(2), 886–899.

Bradlow, A. R., Nygaard, L. C., & Pisoni, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception & Psychophysics*, *61*(2), 206–219.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English/r/and/l: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America*, *101*(4), 2299–2310.

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software*, *80*, 1-28.

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *The Journal of the Acoustical Society of America*, *116*(6), 3647–3658.

Clifford, S., & Jerit, J. (2014). Is there a cost to convenience? An experimental comparison of data quality in laboratory and online studies. *Journal of Experimental Political Science*, *1*(2), 120–131.

Clopper, C. G., & Pisoni, D. B. (2004). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics*, *32*(1), 111–140.

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., & McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. *Journal of Experimental Psychology: General*, *134*(2), 222.

Dupoux, E., & Green, K. (1997). Perceptual adjustment to highly compressed speech: Effects of talker and rate changes. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(3), 914.

Earle, F. S., & Myers, E. B. (2015a). Overnight consolidation promotes generalization across talkers in the identification of nonnative speech sounds. *The Journal of the Acoustical Society of America*, *137*(1), EL91–EL97.

Earle, F. S., & Myers, E. B. (2015b). Sleep and native language interference affect non-native speech sound learning. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(6), 1680.

Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception & Psychophysics*, *67*(2), 224–238.

Ferguson, S. H., Jongman, A., Sereno, J. A., & Keum, K. A. (2010). Intelligibility of foreign-accented speech for older adults with and without hearing loss. *Journal of the American Academy of Audiology*, *21*(03), 153–162.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*(1), 47–65.

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *The Journal of the Acoustical Society of America*, *89*(1), 395–411.

Flege, J. E., & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, *15*(1), 67–83.

Flege, J. E., Schirru, C., & MacKay, I. R. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, *40*(4), 467–491.

Flege, J. E., Takagi, N., & Mann, V. (1995). Japanese adults can learn to produce English/I/and/l/accurately. *Language and Speech*, *38*(1), 25–55.

Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance*, *32*(5), 1276.

Gass, S., & Varonis, E. M. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning*, *34*(1), 65–87.

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y. S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The annals of applied statistics*, *2*(4), 1360-1383.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1166.

Gordon-Salant, S., Yeni-Komshian, G. H., & Fitzgibbons, P. J. (2010). Recognition of accented English in quiet and noise by younger and older listeners. *The Journal of the Acoustical Society of America*, *128*(5), 3152–3160.

Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., & Schurman, J. (2010). Short-term adaptation to accented English by younger and older adults. *The Journal of the Acoustical Society of America*, *128*(4), EL200–EL204.

Greenspan, S. L., Nusbaum, H. C., & Pisoni, D. B. (1988). Perceptual learning of synthetic speech produced by rule. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 421.

Grohe, A.-K., & Weber, A. (2016). The penefit of salience: Salient accented, but not unaccented words reveal accent adaptation effects. *Frontiers in Psychology*, *7*, 864.

Guion, S. G., Flege, J. E., Liu, S. H., & Yeni-Komshian, G. H. (2000). Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics*, *21*(2), 205–228.

Hay, J., & Drager, K. (2010). *Stuffed toys and speech perception*.

Hay, J., Nolan, A., & Drager, K. (2006). *From fush to feesh: Exemplar priming in speech perception*.

Hayes-Harb, R., Smith, B. L., Bent, T., & Bradlow, A. R. (2008). The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts. *Journal of Phonetics*, *36*(4), 664–679.

Jacewicz, E., Fox, R. A., & Wei, L. (2010). Between-speaker and within-speaker variation in speech tempo of American English. *The Journal of the Acoustical Society of America*, *128*(2), 839–850.

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*(10), 1995–2004.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America*, *108*(3), 1252–1263.

Kang, K.-H., & Guion, S. G. (2006). Phonological systems in bilinguals: Age of learning effects on the stop consonant systems of Korean-English bilinguals. *The Journal of the Acoustical Society of America*, *119*(3), 1672–1683.

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, *122*(2), 148.

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*, *51*(2), 141–178.

Kraljic, T., & Samuel, A. G. (2011). Perceptual learning evidence for contextually-specific representations. *Cognition*, *121*(3), 459–465.

Krause, J. C., & Braida, L. D. (2002). Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility. *The Journal of the Acoustical Society of America*, *112*(5), 2165–2172.

Krause, J. C., & Braida, L. D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *The Journal of the Acoustical Society of America*, *115*(1), 362–378.

Laturnus, R. (2018). Perceptual adaptation to non-native speech: The effects of bias, exposure, and input variation. *New York University, New York, NY*.

Laturnus, R. (2020). Comparative Acoustic Analyses of L2 English: The Search for Systematic Variation. *Phonetica*, *77*(6), 441–479.

Lee, D., & Baese-Berk, M. M. (2020). The maintenance of clear speech in naturalistic conversations. *The Journal of the Acoustical Society of America*, *147*(5), 3702–3711.

Lev-Ari, S. (2017). Talking to fewer people leads to having more malleable linguistic representations. *PloS One*, *12*(8), e0183593.

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. *The Journal of the Acoustical Society of America*, *94*(3), 1242–1255.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *The Journal of the Acoustical Society of America*, *89*(2), 874–886.

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*(2), 173–190.

Maniwa, K., Jongman, A., & Wade, T. (2008). Perception of clear fricatives by normal-hearing and simulated hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *123*(2), 1114–1125.

Maniwa, K., Jongman, A., & Wade, T. (2009). Acoustic characteristics of clearly spoken English fricatives. *The Journal of the Acoustical Society of America*, *125*(6), 3962–3973.

Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2008). The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science*, *32*(3), 543–562.

Mitterer, H., & McQueen, J. M. (2009). Foreign subtitles help but native-language subtitles harm foreign speech perception. *PloS One*, *4*(11), e7785.

Mok, P., & Dellwo, V. (2008). *Comparing native and non-native speech rhythm using acoustic rhythmic measures: Cantonese, Beijing Mandarin and English*.

Mullennix, J. W., & Pisoni, D. B. (1990). Stimulus variability and processing dependencies in speech perception. *Perception & Psychophysics*, *47*(4), 379–390.

Mullennix, J. W., Pisoni, D. B., & Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *The Journal of the Acoustical Society of America*, *85*(1), 365–378.

Munro, M. J. (1993). Productions of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings. *Language and Speech*, *36*(1), 39-66.

Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, *38*(3), 289–306.

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *49*, 285–310.

Myers, E., Johns, A. R., Earle, F. S., & Xie, X. (2017). The invariance problem in the acquisition of non-native phonetic contrasts: From instances to categories. *The Speech Processing Lexicon: Neurocognitive and Behavioural Approaches*, *22*(1), 52–84.

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, *18*(1), 62–85.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*(2), 204–238.

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics*, *60*(3), 355–376.

Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, *5*(1), 42–46.

Oh, G. E., Guion-Anderson, S., Aoyama, K., Flege, J. E., Akahane-Yamada, R., & Yamada, T. (2011). A one-year longitudinal study of English and Japanese vowel production by Japanese adults and children in an English-speaking setting. *Journal of Phonetics*, *39*(2), 156–167.

Pallier, C., Sebastian-Gallés, N., Dupoux, E., Christophe, A., & Mehler, J. (1998). Perceptual adjustment to time-compressed speech: A cross-linguistic study. *Memory & Cognition*, *26*(4), 844–851.

Pearson. (2009). Official guide to Pearson Test of English Academic.

Peelle, J. E., & Wingfield, A. (2005). Dissociations in perceptual learning revealed by adult age differences in adaptation to time-compressed speech. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(6), 1315.

Perrachione, T. K., Furbeck, K. T., & Thurston, E. J. (2019). Acoustic and linguistic factors affecting perceptual dissimilarity judgments of voices. *The Journal of the Acoustical Society of America*, *146*(5), 3384-3399.

Perrachione, T. K., Lee, J., Ha, L. Y., & Wong, P. C. (2011). Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *The Journal of the Acoustical Society of America*, *130*(1), 461–472.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, *28*(1), 96–103.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, *29*(4), 434–446.

Pinet, M., Iverson, P., & Evans, B. G. (2011). *Perceptual Adaptation for L1 and L2 Accents in Noise by Monolingual British English Listeners.* 1602–1605.

Porretta, V., Kyröläinen, A. J., & Tucker, B. V. (2015). Perceived foreign accentedness: Acoustic distances and lexical properties. *Attention, Perception, & Psychophysics*, *77*(7), 2438-2451.

R Core Team. (2021). *R: a language and environment for statistical computing* [Computer Program]. Version 4.1.2, retrieved 1 November 2021 from https://www.R-project.org/.

Reinisch, E., & Holt, L. L. (2014). Lexically guided phonetic retuning of foreign-accented speech and its generalization. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(2), 539.

Roark, C. L., Feng, G., & Chandrasekaran, B. (2022). Talker identification as a categorization problem. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 44, No. 44).

RStudio Team. (2021). *RStudio: integrated development for R* [Computer Program]. Version 2021.09.1, retrieved 8 November 2021 from http://www.rstudio.com/.

Sidaras, S. K., Alexander, J. E., & Nygaard, L. C. (2009). Perceptual learning of systematic variation in Spanish-accented speech. *The Journal of the Acoustical Society of America*, *125*(5), 3306–3316.

Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. *The Journal of the Acoustical Society of America*, *64*(5), 1358–1368.

Stoycheff, E. (2016). Please participate in Part 2: Maximizing response rates in longitudinal MTurk designs. *Methodological Innovations*, *9*, 2059799116672879.

Toivola, M., Lennes, M., & Aho, E. (2009). *Speech rate and pauses in non-native Finnish.* 1707–1710.

Tzeng, C. Y., Alexander, J. E., Sidaras, S. K., & Nygaard, L. C. (2016). The role of training structure in perceptual learning of accented speech. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(11), 1793.

U.S. Census Bureau. (2010). *Detailed language spoken at home and ability to speak English for the population 5 years and older by states: 2006-2008*. Retrieved from https://www.census.gov/data/tables/2008/demo/2006-2008-lang-tables.html

U.S. Census Bureau. (2015). *Detailed languages spoken at home and ability to speak English for the population 5 years and over: 2009-2013*. Retrieved from https://www.census.gov/data/tables/2013/demo/2009-2013-lang-tables.html

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *The Journal of the Acoustical Society of America*, *106*(6), 3649–3658.

Wayland, R. (1997). Non-native production of Thai: Acoustic measurements and accentedness ratings. *Applied Linguistics*, *18*(3), 345–373.

Weil, S. A. (2001). *Foreign accented speech: Adaptation and generalization*.

Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful listening: The processing of degraded speech depends critically on attention. *Journal of Neuroscience*, *32*(40), 14010–14021.

Xie, X., & Myers, E. B. (2017). Learning a talker or learning an accent: Acoustic similarity constrains generalization of foreign accent adaptation to new talkers. *Journal of Memory and Language*, *97*, 30–46.

Xie, X., Weatherholtz, K., Bainton, L., Rowe, E., Burchill, Z., Liu, L., & Jaeger, T. F. (2018). Rapid adaptation to foreign-accented speech and its transfer to an unfamiliar talker. *The Journal of the Acoustical Society of America*, *143*(4), 2013–2031.

Yang, B. (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics*, *24*(2), 245–261.