Native Korean Listeners' Social Expectations and Perception of Foreign-Accented Korean Speech

by

Yeojin Jung

A dissertation accepted and approved in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in East Asian Languages & Literatures

Dissertation Committee:

Kaori Idemaru, Chair

Zhuo Jing-Schmidt, Core Member

Rachel E. Weissler, Core Member

Jeff Holliday, Core Member

Vsevolod Kapatsinski, Institutional Representative

University of Oregon

Spring 2026

© 2026 Yeojin Jung.
All rights reserved.

DISSERTATION ABSTRACT

Yeojin Jung

Doctor of Philosophy in East Asian Languages & Literatures

Title: Native Korean Listeners' Social Expectations and Perception of Foreign-Accented Korean Speech

What factors beyond acoustic features influence native speakers' perception of foreign-accented speech? This dissertation examines how native Korean listeners' social expectations affect the perception and evaluation of American English-accented Korean in a context where East Asians constitute the racial majority and foreignness is not necessarily associated with lower socioeconomic status. Building on research on Reverse Linguistic Stereotyping (RLS) and exemplar-based accounts of speech perception, this study tests predictions involving race- and nationality-based socioeconomic expectations, ethnic in-group membership, and expectation mismatch. Using an inverted matched-guise design, identical American English-accented Korean speech was paired with one of five visual and descriptive cues: a silhouette baseline, a White American, a Korean-American, a Chinese learner, and a Korean mismatch condition. Native Korean listeners completed (1) a transcription task in noise, (2) an accentedness rating task, and (3) a language attitude survey.

The results showed limited evidence for social bias effects in perceptual measures. Intelligibility was largely unaffected by social information at the sentence level, with only minor effects at the target-word level. In contrast, accentedness judgments were influenced by expectation mismatch, with incongruent social-acoustic pairings yielding higher perceived accentedness. Language attitude results further showed that the Korean-American guise received the lowest ratings in both status and solidarity dimensions. Exploratory analyses also examined

acoustic features associated with accentedness and language attitudes. Overall, the findings demonstrate that speech perception is shaped by the interaction between acoustic properties and social expectations, and provide insight into how native Korean listeners associate foreign accents with speaker identity and how Koreanness may be construed in relation to accented speech.

ACKNOWLEDGMENTS

I would like to express my sincere appreciation to my advisor, Dr. Kaori Idemaru, for her guidance and steadfast support throughout my academic journey. I would not have been able to complete this long journey without her. Through her mentorship, I learned how to approach language science and research with rigor and gained invaluable wisdom for life. Thanks to her, I have grown into a more independent researcher and a better person.

I would also like to extend my heartfelt gratitude to my committee members. I am deeply grateful to Dr. Zhuo Jing-Schmidt for her consistent support and for giving me opportunities to teach and give guest lectures. I am also deeply thankful to Dr. Rachel Weissler, whose emotional support and encouragement helped me continue through moments when I lacked confidence. I sincerely appreciate Dr. Volya Kapatsinski for helping me think more critically about theoretical issues and for providing invaluable guidance on statistics. I am also grateful to Dr. Jeff Holliday for sharing his insights into Korean society and the Korean language, and for helping me develop my own reasoning process and critical thinking.

Beyond my committee, I would like to thank Dr. Keunyoung Lee, who has been one of my reliable supporters and friends in Eugene. I am also grateful to Dr. Lisa Redford for her care and support in many ways, and to Dr. Keli Yerian for always offering valuable insights on teaching and life. I would also like to thank Dr. Kris Kyle for inviting me into a church community, where I was able to build meaningful friendships in Eugene. Special thanks to Dr. Jeong-Im Han for her support and encouragement, and for always being willing to help when I needed to conduct experiments in Korea.

I am also deeply grateful to my beloved UO linguist friends, Carissa Diantoro, Stephanie Curiel, Tamaya Levy, Dakota Whisler, and Dr. Hakyung Sung. Although I did not have a cohort of my own, I was fortunate to have you all, who felt like one and shared my time in Eugene with warmth and support. I would also like to thank my dear friends outside the linguistics field, Dr. Yeseul Jeon, Dr. Jeehyun Jenny Lee, and Dr. Sungue Susan Yang, who went through their doctoral journeys around the same time and gave me strength in the U.S. I am also grateful to my friends, Minjee Loh, Hyojeong Seo, Minkyu Kim, and Taehee Lee, for standing by me from afar and for being beautiful companions in my life.

Finally, I would like to express my deepest gratitude to my parents, Yu-In Jung and Sunhee Park, for quietly believing in me and supporting me from afar throughout this journey and my whole life. Words cannot fully express how grateful I am for their love, patience, and unwavering support.

TABLE OF CONTENTS

DISSERTATION ABSTRACT 3

ACKNOWLEDGMENTS 5

TABLE OF CONTENTS 7

LIST OF FIGURES 14

LIST OF TABLES 16

I. INTRODUCTION 18

1.1 Listeners' social expectations in speech perception 20

1.2 Language attitudes as an evaluative form of listener bias 26

1.3 Overview of the dissertation 33

II. SOCIAL EXPECTATIONS IN INTELLIGIBILITY AND ACCENTEDNESS JUDGMENTS 36

2.1 Introduction 36

2.1.1 Summary of research questions 39

2.2 Pilot studies for stimulus creation and selection 41

2.2.1 Pilot 1 41

2.2.1.1 Stimuli 41

2.2.1.2 Participants 43

2.2.1.3 Procedure 43

2.2.1.4 Assessment of transcription 44

2.2.2 Pilot 2 45

2.2.2.1 Stimuli 46

2.2.2.2 Participants 46
2.2.2.3 Procedure 46
2.2.2.4 Results 46
2.3 Experiment 47
2.3.1 Methodology 47
2.3.1.1 Design 47
2.3.1.2 Audio stimuli 47
2.3.1.3 Visual stimuli 48
2.3.1.4 Participants 50
2.3.1.5 Procedure 51
2.3.2 Analysis 53
2.3.2.1 Intelligibility task 53
2.3.2.2 Accentedness rating task 55
2.3.2.3 Correlation 55
2.3.3 Results 56
2.3.3.1 Intelligibility task 56
2.3.3.2 Accentedness rating task 61
2.3.3.3. Correlation 65
2.3.3.3.1 Intelligibility and accentedness 65
2.3.3.3.2 Demographic and language background effects 65
2.4 Discussion 66
2.5 Conclusion 72

III. NATIVE KOREAN LISTENERS' LANGUAGE ATTITUDES TOWARD ACCENTED KOREAN: EVALUATIONS OF CLOSE OUTSIDERS AND DISTANT INSIDERS 75

3.1 Introduction 75

3.1.1 Summary of Research Questions 78

3.2 Methodology 80

3.2.1 Language attitude survey 81

3.2.2 Procedure 82

3.3 Analysis 83

3.3.1 Language attitude data 83

3.3.2 Correlation 84

3.4 Results 85

3.4.1 Language attitude 85

3.4.1.1 Positive affect 86

3.4.1.2 Negative affect 89

3.4.1.3 Status 91

3.4.1.4 Solidarity 93

3.4.1.5 Comprehensibility 95

3.4.2 Relationships among language attitude dimensions 98

3.4.3 Correlation with intelligibility and accentedness 98

3.4.4 Correlation with demographic & language background survey 99

3.5 Discussion 102

3.5.1 Positive and negative affect evaluations 102

3.5.2 Status and solidarity dimensions 104

3.5.3 Comprehensibility 108
3.5.4 Relationship between language attitudes and perceptual measures 110
3.5.5 Listeners' demographic and language background factors 112
3.6 Conclusion 115
IV. ACOUSTIC AND PHONOLOGICAL REALIZATION OF KOREAN STOPS AND VOWELS BY AMERICAN ENGLISH LEARNERS: A COMPARISON WITH NATIVE KOREAN SPEAKERS 117
4.1 Introduction 117
4.1.1 Research questions 121
4.2 Methodology 121
4.2.1 Participants 121
4.2.2 Speech materials and recording 122
4.2.3 Acoustic measurements and phonological coding 124
4.2.3.1 Vowels 124
4.2.3.2 Stops 125
4.2.3.3 Phonological realization patterns 126
4.3 Analysis 127
4.3.1 Vowels 127
4.3.2 Stops 129
4.3.3 Phonological realization patterns 130
4.4 Results 131
4.4.1 Vowels 131
4.4.1.1 Vowel quality 131

4.4.1.2 Vowel duration 134
4.4.2 Stops 137
4.4.2.1 VOT 137
4.4.2.2 F0 141
4.4.3 Phonological realization patterns 144
4.5 Discussion 146
4.5.1 Vowels 146
4.5.2 Stops 148
4.5.3 Phonological realization patterns 152
4.6 Conclusion 154
V. SPEECH FEATURES ASSOCIATED WITH ACCENTEDNESS JUDGMENTS AND LANGUAGE ATTITUDES TOWARD FOREIGN-ACCENTED KOREAN 156
5.1 Introduction 156
5.1.1 Research questions 160
5.2 Methodology 161
5.2.1 Participants 161
5.2.2 Audio stimuli 161
5.2.3 Procedure 162
5.2.4 Acoustic measurement 163
5.3 Analysis 164
5.3.1 Acoustic features associated with accentedness judgments 164
5.3.2 Acoustic features associated with language attitudes 165
5.3.3 Listeners' subjective responses 167

5.4 Results 168
5.4.1 Accentedness judgments 168
5.4.1.1 Fluency-related features 168
5.4.1.2 Stop features 171
5.4.1.3 Vowel features 173
5.4.2 Language attitudes 175
5.4.2.1 Fluency-related features 176
5.4.2.2 Stop features 179
5.4.2.3 Vowel features 182
5.4.3 Listeners' subjective responses 186
5.5 Discussion 188
5.5.1 Accentedness judgments 188
5.5.2 Language attitudes 190
5.6 Conclusion5.6 Conclusion 196
VI. CONCLUSION 198
6.1 Summary of the findings 198
6.2 Contributions 200
6.3 Limitations and future directions 202
6.4 Conclusion 204
APPENDICES 206
A. List of sentence stimuli 206
B. Examples of transcriptions judged as correct or incorrect 210
C. Language attitude questionnaire 211

D. Demographic and language background questionnaire 213

REFERENCES 216

LIST OF FIGURES

Figure Page

1. Figure 2.1. Face used in the experiment 50

2. Figure 2.2. Estimated effects of guise condition on whole sentence intelligibility relative to the Baseline (logit scale) 60

3. Figure 2.3. Estimated effects of guise condition on target word intelligibility relative to the Baseline (logit scale) 61

4. Figure 2.4. Distribution of accentedness ratings by condition 64

5. Figure 3.1. Face used in the experiment 81

6. Figure 3.2. Estimated differences in positive affect ratings from the Baseline 88

7. Figure 3.3. Estimated differences in negative affect ratings from the Baseline 91

8. Figure 3.4. Estimated differences in status ratings from the Baseline 93

9. Figure 3.5. Estimated differences in solidarity ratings from the Baseline 95

10. Figure 3.6. Estimated differences in comprehensibility ratings from the Baseline 97

11. Figure 4.1. Example waveform and spectrogram of native and learner productions of 갚았다 /kap^{h}-ass-ta/ 130

12. Figure 4.2. Normalized vowel spaces for native Korean speakers and American English learners of Korean 133

13. Figure 4.3. Normalized vowel spaces for native Korean speakers and American English learners of Korean in a single plot 134

14. Figure 4.4. Posterior estimates of vowel duration for native speakers and learners across eight Korean vowels 136

15. Figure 4.5. Posterior estimates of VOT for native and learner groups across laryngeal categories and prosodic positions 139

16. Figure 4.6. Posterior estimates of F0 for native and learner groups across laryngeal categories and prosodic positions 143

17. Figure 5.1. Example screen from the language attitude survey 163

18. Figure 5.2. GAM smooths for fluency-related predictors on accentedness judgments 170

19. Figure 5.3. GAM smooths for significant stop-related predictors on accentedness judgments 173

20. Figure 5.4. GAM smooths for significant vowel-related predictors on accentedness judgments 175

21. Figure 5.5. GAM smooths for fluency-related predictors on language attitude ratings 179

22. Figure 5.6. GAM smooths for stop-related predictors on language attitude ratings 182

23. Figure 5.7. GAM smooths for vowel-related predictors on language attitude ratings 185

LIST OF TABLES

Table Page

1. Table 2.1. Overview of experimental conditions and associated social expectations 41

2. Table 2.2. Posterior medians and 95% CrIs for the effects of condition on whole sentence (logit scale) 57

3. Table 2.3. Posterior median and 95% CrIs for the effects of condition on target word accuracy (logit scale) 57

4. Table 2.4. Posterior contrasts for whole sentence accuracy (logit scale) 59

5. Table 2.5. Posterior contrasts for target word accuracy (logit scale) 60

6. Table 2.6. Effects on perceived accentedness (logit scale) 62

7. Table 2.7. Posterior contrasts in perceived accentedness 64

8. Table 3.1. Overview of experimental conditions and associated social expectations 80

9. Table 3.2. Posterior median ratings and 95% HPD intervals for each attitudinal dimension across guise conditions 87

10. Table 3.3. Posterior contrasts for positive affect (logit scale) 88

11. Table 3.4. Posterior contrasts for negative affect (logit scale) 90

12. Table 3.5. Posterior contrasts for status dimension (logit scale) 92

13. Table 3.6. Posterior contrasts for solidarity dimension (logit scale) 94

14. Table 3.7. Posterior contrasts for comprehensibility (logit scale) 97

15. Table 3.8. Spearman correlations among language attitude dimensions 98

16. Table 3.9. Correlations between language attitude and speech perception measures 99

17. Table 3.10. Correlations between language attitude and listener background variables 101

18. Table 4.1. Posterior contrasts of normalized F1 between native and learner groups 132

19. Table 4.2. Posterior contrasts of normalized F2 between native and learner groups 133

20. Table 4.3. Posterior contrasts of vowel durations between native and learner groups..... 135

21. Table 4.4. Posterior contrasts between native merged [e] and learner [e] and [ɛ] 137

22. Table 4.5. Posterior contrasts of VOT between native and learner groups 139

23. Table 4.6. Posterior contrasts of VOT between laryngeal contrasts within native and learner groups..... 141

24. Table 4.7. Posterior contrasts of F0 between native and learner groups 142

25. Table 4.8. Posterior contrasts of F0 between laryngeal contrasts within native and learner groups..... 144

26. Table 4.9. Phonological realization patterns in learner productions..... 145

27. Table 4.10. Variation in phonological realizations in native and learner productions..... 146

28. Table 5.1. Effects of non-smooth predictors on accentedness judgments 168

29. Table 5.2. Effects of fluency-related predictors on accentedness judgments 170

30. Table 5.3. Effects of stop-related predictors on accentedness judgments 172

31. Table 5.4. Effects of vowel-related predictors on accentedness judgments 174

32. Table 5.5. Effects of non-smooth predictors on language attitude ratings 176

33. Table 5.6. Effects of fluency-related predictors on language attitude ratings 178

34. Table 5.7. Effects of stop-related predictors on language attitude ratings..... 181

35. Table 5.8 Effects of vowel-related predictors on language attitude ratings 184

36. Table 5.9. Summary of listeners' subjective responses about features perceived as foreign accents 188

I. INTRODUCTION

Communication between native and non-native speakers is often assumed to be challenging primarily because non-native speech may deviate from dominant phonetic and prosodic norms. A widely held view in second language acquisition research is that learners who begin acquiring a target language after early childhood, roughly after the ages of 6-8, tend to produce acoustic and prosodic patterns that diverge from those of native speakers (Diller, 1981; Long, 1990). Such acoustic variation can contribute to accentedness judgments, intelligibility, and comprehensibility.

However, this speaker-centered view is incomplete, because listeners' perceptions of accented speech are not shaped solely by acoustic properties of the speech signal. They are also influenced by listeners' social expectations and beliefs about who the speaker is. This tendency to locate communicative difficulty in the speaker is further reinforced by broader language ideologies, in which accented speech is framed as deficient or problematic, and communicative breakdowns are attributed to the speaker rather than the listener (Lippi-Green, 2012). At the same time, an emerging body of literature suggests that communicative difficulty cannot be fully explained by failures of intelligibility alone (Brown et al., 2020; Munro & Derwing, 1995; Van Engen & Peelle, 2014). Even when non-native speech is accurately perceived at the segmental level, listeners may still experience increased processing difficulty due to other listener-related factors, resulting in subjective judgments of reduced comprehensibility.

Munro and Derwing (1995) demonstrated this distinction by showing that Mandarin Chinese-accented English was perfectly transcribed by native English listeners, but was

nonetheless judged as difficult to understand. This distinction between objective intelligibility and perceived comprehensibility shifts attention to listener-related factors such as language background, expectations, bias, and cognitive mechanisms underlying perceived effort (Baese-Berk et al., 2020; Van Engen & Peelle, 2014). Taken together, these findings suggest that challenges in native and non-native communication arise not only from properties of the speech signal, but also from the interaction between acoustic input and the listener's perceptual and cognitive mechanisms.

To better understand the mechanisms underlying the perception of foreign-accented speech, this dissertation examines listener-related factors grounded in social information, as well as acoustic properties that affect native listeners' perception of accentedness, with a specific focus on the South Korean context, where the social configuration differs from that of many previously studied settings. In this introduction, relevant literature is reviewed to establish a framework for understanding how communicative difficulty arises in foreign-accented speech perception.

Specifically, Section 1.1 reviews research on listeners' social expectations such as bias and how mismatches between expected and actual speech may influence speech perception. Section 1.2 focuses on language attitudes toward accented speech, which represent a more evaluative form of listener bias. Finally, Section 1.3 introduces the research questions and motivates the study by explaining why the South Korean context is necessary and how it differs from previously examined settings.

1.1 Listeners' social expectations in speech perception

A substantial body of research has demonstrated that speech perception is influenced not only by acoustic features of the speech signal but also by social information associated with the speaker (Drager, 2010; Niedzielski, 1999). Listeners infer social meaning from speech and shape expectations about speakers' identities, such as gender (Strand, 1999, 2000; Strand & Johnson, 1996), age (Drager, 2011; Koops et al., 2008), regional background (Hay & Drager, 2010), ethnicity (Casasanto, 2008), race (Kutlu, 2023; Kutlu et al., 2022), and socioeconomic status (Hay et al., 2006). These expectations can shape listeners' perception in alignment with the social information attributed to the speaker. In Hay and Drager's study (2010), the experimenter pretended to accidentally find a stuffed toy kangaroo or koala (associated with Australia) or kiwi (associated with New Zealand) and placed it in the room without explicitly mentioning the regions. This subtle exposure primed a specific region for participants and influenced their performance in a subsequent perception task, where they had to decide whether vowel sounds along a continuum were Australian or New Zealand in pronunciation. Their perception of the vowels changed depending on the subtle contextual cues they had been exposed to.

Foreign-accented speech provides a particularly informative context for examining this phenomenon. Foreign-accented speech often differs from listeners' linguistic norms at both segmental and suprasegmental levels, and these differences are usually associated with social categories such as region of origin, ethnicity, or race. As a result, foreign-accented speech can activate listeners' social expectations and stereotypes, which in turn influence perceptual judgments. In the previous literature on accented speech perception, two main explanatory accounts have been proposed. The first is the bias account, which posits that negative biases or stereotypes associated with particular accents or with social information provided about the

speaker negatively affect perceptual performance, such as intelligibility, comprehensibility, and accentedness judgments. The second is the exemplar account, which suggests that perceptual performance is degraded when the speech signal (i.e., accent) does not match prototypical social information associated with the speaker, such as race or nationality.

The bias account has been most extensively illustrated in a series of studies by Rubin and colleagues (O. Kang & Rubin, 2009; Rubin, 1992; Rubin & Smith, 1990), which demonstrated that listeners' assumptions about a speaker's identity can systematically bias speech perception, even when the acoustic signal is identical. In one of Rubin's studies (1992), North American students listened to a lecture recorded with a Standard American English accent and presented with a photograph of either a White or an Asian (Chinese) female face. After listening, participants completed a cloze test based on a written transcript of the lecture to assess listening comprehension. They then responded to the Homophily instrument (McCroskey et al., 1975), which measures perceived similarity in attitude, background, values, and appearance. Finally, they rated the teaching assistant's accentedness, teaching competence, and perceived ethnicity.

Despite the fact that the audio stimuli were identical in both conditions, participants reported higher homophily (i.e., sense of similarity) ratings for the White guise, with attitude and background homophily correlating with appearance. Most importantly, they rated the speech paired with the Asian guise as more accented and performed worse on the cloze test compared to those who saw the White guise.

Rubin attributed these results to negative biases held by native English speakers, particularly stereotypes associated with the speaker's perceived ethnicity (i.e., Asian) and based

on this, he theorized that listeners' potential biases tied to a speaker's social characteristics can negatively influence their perception and comprehension of speech. This theory, referred to as *bias hypothesis*, was further elaborated by Kang and Rubin (2009) into the concept of Reverse Linguistic Stereotyping (RLS). RLS is based on the idea of linguistic stereotyping, which refers to the tendency to categorize and judge others based on their language output (Dragojevic & Goatley-Soan, 2022; Lippi-Green, 2012). Research on linguistic stereotyping provides important groundwork for understanding RLS, as many studies have demonstrated that accent is a more salient cue for ethnic categorization than physical appearance (Hansen et al., 2017; Kinzler et al., 2009; Rakić et al., 2011). Moreover, numerous studies have shown that individuals with non-native or non-standard accents are evaluated differently from those with standard accents (Fuertes et al., 2012).

RLS describes the type of stereotyping that occurs in the opposite direction from linguistic stereotyping. While linguistic stereotyping refers to the process by which language use influences social categorization and evaluation, RLS refers to the process by which listeners' social expectations about a speaker's social identity (e.g., generated from their looks) influence their perception and evaluation of the speaker's speech (Drager, 2010; Hay et al., 2006; Mulac et al., 2013). More specifically, language schemata and stereotypes (e.g., implicit and explicit beliefs about how particular social categories, such as Asian immigrants, are expected to speak English) may influence how listeners perceive and judge speech and the speaker (e.g., they must be accented) (Mulac et al., 2013). Collins and Clément (2012) describe this phenomenon as "a lens that directs and distorts cognition," emphasizing that stereotyping leads people to notice behaviors that confirm their expectations while ignoring those that contradict them.

Building on Rubin's earlier study (1992), Kang and Rubin (2009) used a cloze probability test to measure listening comprehension and asked participants to rate teaching quality and accent standardness. Their participants listened to four counterbalanced short lectures, two recorded by the same native male English speaker and two by a distinctive East-Asian speaker of intermediate intelligibility, included as a distractor. While listening to the native speaker's lectures, participants were shown either an Asian face or a White face. In this study, Kang and Rubin aimed to directly measure the RLS construct and examine individual background factors that might influence susceptibility to RLS. To capture a diverse range of experiences, they recruited participants with various background dimensions such as exposure to non-native English-speaking people or formal training in language studies. After listening to the lectures, participants assessed the speaker's traits, including perceived superiority and social attractiveness.

The results revealed that RLS-related perceptions, such as perceived superiority and social attractiveness based on the accompanying photo, significantly influenced participants' evaluations of the speaker's comprehensibility, teaching performance, and accent standardness. More specifically, when listeners believed they were hearing a native Chinese instructor, they experienced greater comprehension difficulties, perceived a heavier accent, and gave lower ratings for teaching quality, despite the identical audio recordings. Approximately 9 to 12% of the variance in non-native speaker's performance ratings was directly attributed to RLS factors. The results indicated that higher RLS in superiority and social attractiveness (perceiving the speaker as *less* superior and *less* socially attractive) associated with lower teaching performance ratings of the speaker. High RLS in social attractiveness was also associated with lower comprehension ratings. Additionally, superiority RLS was negatively correlated with accent

standardness, indicating that stronger RLS perceptions were linked to a more non-standard accent rating. These findings align with previous research, demonstrating that listeners' evaluations of speech are influenced by social expectations and biases related to the speaker's perceived identity, often resulting in negative bias against speakers.

Subsequent studies have extended this line of research by examining the role of visual cues and bias in native listeners' perception of accented speech. Using an Implicit Association Test (IAT) (Greenwald et al., 1998), Yi and colleagues (2013) demonstrated that native English listeners hold strong implicit associations between White faces and American identity, and between Asian faces and foreign identity. These associations modulated speech perception such that audiovisual information enhanced intelligibility more for native English speech than for non-native speech, relative to the audio-only condition, indicating that listener expectations and social associations can differentially facilitate or constrain speech processing. Similarly, Babel and Russell (2015) showed that visual cues interacted with speaker identity in Canadian listeners' evaluations. While White Canadian speakers benefited from the audiovisual condition, Chinese Canadian speakers were penalized, receiving lower intelligibility scores and higher accentedness ratings compared to the audio-only condition. Consistent with these perceptual results, IAT scores revealed a strong positive bias toward White Canadians and a less positive bias toward Asian Canadians. Together, these findings add to the accumulating evidence that information from racial cues affects speech processing, often disadvantaging Asian speakers.

The second explanatory account, the exemplar model of language (Bybee & Torres, 2008; Johnson, 2006; Strand, 2000) proposes that the stereotypical phonetic details associated with social categories are encoded and stored in memory and that the perceptual process is highly sensitive to this association, influencing how speech is processed and evaluated. Building

on this framework, McGowan (2015) reinterpreted the results of Rubin's study (1992), suggesting that listeners in Rubin's study may have experienced difficulty in processing the mismatch between the provided social information (i.e., Asian face) and the acoustic signal (i.e., Standard American English accent; SAE accent) because their encoded memory consists of SAE accent matched with White face. McGowan argues that this mismatch likely disrupted speech processing, leading to lower comprehension and poorer evaluations of teaching quality and accent standardness in the incongruous condition, where an Asian face was paired with a SAE accent.

In McGowan's (2015) study, the design inverted Rubin's conditions. McGowan presented Chinese-accented English paired with either an Asian (Chinese) face, a White, or a silhouette used as a baseline. In this design, the Asian guise represented the congruent condition, while the White face was the incongruent condition. The results showed that the Asian guise yielded significantly higher transcription accuracy than the White guise but did not significantly differ from the silhouette baseline, despite a slightly higher accuracy rate.

The results suggest that congruence between social cues and acoustic signals may improve intelligibility relative to the incongruence but does not necessarily *facilitate* the intelligibility of speech. This is evident from the Asian guise's performance not differing significantly from the baseline (a silhouette with Chinese-accented speech). In other words, speech perception might not be necessarily benefited by congruence, but it might be affected negatively if the expectations (based on social cues) and the actual acoustic signal are not matched due to the processing cost (Strand, 2000).

Building on this line of research, Gnevsheva (2018) directly contrasted RLS and exemplar accounts by examining accentedness perception of non-native White (German) and non-native Asian (Korean) English speakers across audio-only, video-only, and audiovisual conditions. The results showed that Korean speakers' accentedness ratings remained stable across conditions, whereas German speakers were perceived as more accented when visual cues were available than in the audio-only condition. This pattern suggests that when expectations linking White appearance to native-like English speech are violated, processing costs induced by cue mismatch outweigh any potential racial advantage for White speakers, thereby supporting mismatch-based explanations of degraded accented speech perception.

Taken together, findings from research on the exemplar accounts suggest that listeners' responses to a speaker's perceived social identity cannot be explained solely by negative bias toward race. Rather, visual social cues appear to be associated with stored representations linking social categories to expected phonetic patterns (i.e., accents). Therefore, when the perceived social information and the incoming acoustic signal are incongruent, listeners experience increased processing difficulty, which in turn leads to reduced intelligibility and less favorable evaluations of the speech such as accentedness judgment.

1.2 Language attitudes as an evaluative form of listener bias

Results from perceptual tasks such as intelligibility, comprehensibility, and accentedness judgments provide valuable evidence that listeners' perceptions vary systematically depending on given social information. However, such measures alone offer limited insight into the underlying mechanisms driving these differences. In particular, perceptual outcomes themselves

cannot reliably determine whether lower performance reflects negative social bias toward a speaker or increased processing costs caused by expectation mismatch.

To determine which mechanism plays a more central role in driving these outcomes, it would require psycholinguistic measures that provide evidence about online processing costs, such as pupillometry and event-related potential (ERP) components. If negative perceptual outcomes are driven by increased processing demands arising from expectation mismatch, this should be reflected in online measure of processing difficulty (e.g., Hansen et al., 2017). While such methods are crucial for understanding the cognitive mechanisms underlying speech perception, Chapter 3 focuses on socially grounded evaluations, examining how listeners' social evaluations of speakers and speech are patterned and how these evaluations relate to perceptual judgments.

In this line of research, studies adopting bias accounts of accented speech perception have employed various measures of listener bias, including the homophily instrument, the Implicit Association Test (IAT), and explicit stereotype questionnaires. Specifically, measures such as perceived superiority and social attractiveness used in Kang and Rubin (2009) are widely recognized dimensions of language attitudes, an established evaluative form of listener bias. Therefore, this section reviews research on language attitudes as a theoretical framework for understanding how social evaluations shape listeners' judgments of accented speech.

Language attitudes are generally understood as the social evaluations individuals hold toward specific languages, language varieties, or groups using those languages or varieties. For example, many English speakers are reported to hold a view that Italian sounds elegant and lively and French sounds romantic, while German sounds harsh and unpleasant (Giles &

Niedzielski, 1998). Those judgments may partly stem from acoustic features, for instance, sonorants in French sound elegant, while voiceless obstruents or fricative consonants in German sound harsh. However, Giles and Niedzielski (1998) argue that such judgments about a language variety are strongly tied to the social connotations and evaluations of those national and cultural identities, such as strong associations with art, fashion, or historical prestige in France and Italy. Further, many studies of language attitudes in sociolinguistics have focused on listeners' perception of a speaker's group membership and listeners' beliefs about that group, since attitudes toward languages and specific language varieties are believed to be closely linked to attitudes toward the people who use them (Preston, 2013). According to Hogg and Reid (2006), when listeners categorize speakers by their languages, they adjust their mental representation of speakers to align with the context-dependent prototype of a category. Once categorized, individuals are then perceived through a lens framed by the relevant group prototype and viewed as embodiments of the attributes of their group.

On this basis, numerous studies on language attitudes have investigated how stereotypes about "standard" and "non-standard" language varieties are organized along two evaluative dimensions: *status* and *solidarity* (Coupland & Bishop, 2007; Grondelaers et al., 2019; Kinzler & DeJesus, 2013; Rickford, 1985). Standard language usually refers to prescriptive norms that determine the correct or canonical spoken and written forms of a language with respect to grammar, vocabulary, and pronunciation, whereas non-standard varieties are those that deviates from such norms in some way (Milroy, 2006). One of the most comprehensive examples of language attitude studies is Coupland and Bishop's (2007) large-scale online survey of 5,010 U.K. respondents, which examined language attitudes toward 34 accents, including both regional varieties within the U.K. and foreign accents such as Australian, French, and South African. In

their framework, *status* was framed as *prestige* and *solidarity* as *social attractiveness* and survey questions were posed directly for each accent: "How much prestige do you think is associated with this accent?" and "How pleasant do you think this accent sounds?" Many studies investigating language attitude follow this model (See Dragojevic & Goatley-Soan, 2022 for a review)

In terms of *prestige*, the Queen's English was ranked highest, followed by the Standard accent of English. Although both are categorized as Received Pronunciation (RP), the Queen's English was more associated with royalty, poshness, and higher social class, while the Standard English accent was perceived as more neutral. Following this, respondents tended to rank their own accent third; for instance, Scottish respondents rated Scottish and Edinburgh accents as more prestigious than respondents from other regions did. Coupland and Bishop interpreted this pattern as evidence of in-group loyalty.

With regard to *social attractiveness*, respondents from Scotland, Wales, and Northern Ireland showed clear in-group loyalty. For example, Scottish respondents evaluated the Scottish accent more positively than did all other groups, and Welsh respondents rated Welsh English more favorably than any other group. However, unlike the status dimension, Queen's English received the lowest attractiveness ratings in Northern Ireland, Scotland, and Wales, although it was rated most attractive in South-East England.

As it is seen in Coupland and Bishop (2007), people tend to ascribe more *status* (*prestige*) to standard language and less status to non-standard varieties. On the other hand, they tend to attribute more *solidarity* (*social attractiveness*) to non-standard varieties especially their own accent and less to standard language. Giles and Marlow (2011) argue that this tendency is rather

consistent across different social strata and linguistic groups because language varieties within a given speech community can be ordered on a hierarchy of prestige, typically corresponding to the socioeconomic status of the social groups with which those varieties are associated. Varieties associated with socioeconomically dominant groups such as standard language tend to carry high prestige, hereafter referred to as high prestige varieties (Milroy & Milroy, 1999). In contrast, varieties associated with socioeconomically subordinate groups or minorities tend to carry low prestige including minority groups or non-standard varieties such as regional dialects, foreign accents, and varieties associated with stigmatized groups (hereafter, low prestige varieties).

Status-related dimensions normally include traits such as competence, intelligence, education, wealth and success, and they are grounded in perceptions of socioeconomic status (Fiske et al., 2007; Fuertes et al., 2012). Consequently, high prestige varieties are generally evaluated more positively on status-related dimensions compared to low prestige varieties (Coupland & Bishop, 2007; Fuertes et al., 2012). Importantly, speakers of low prestige varieties are rated less favorably on status not only by members of other groups but also by members of their own linguistic communities (Kinzler & DeJesus, 2013). Furthermore, the greater the divergence of a variety from high prestige forms, the less status it tends to be attributed, often in a gradient manner (Giles, 1972; Grondelaers et al., 2019).

By contrast, solidarity-related dimensions include traits such as warmth, friendliness, pleasantness, trustworthiness, and often in-group preference (Dragojevic et al., 2013). Accordingly, low prestige varieties tend to receive more favorable evaluations on solidarity, particularly from speakers of the same language variety, since solidarity indexes in-group loyalty. Labov (1966) referred to this phenomenon as “covert prestige,” noting that non-standard varieties can sometimes carry positive value in terms of solidarity, even if they lack overt

prestige. Such covert prestige is especially evident when varieties are seen as expressing shared history, cultural values, or collective (Powesland & Giles, 1975; Thomas & Wassink, 2010), and it can also be reinforced by institutional support, as in the case of African-American English within African-American communities (Baugh, 1983; Rahman, 2004).

As discussed above, people tend to ascribe higher prestige to standard varieties and greater solidarity to non-standard varieties, particularly when they are associated with their own group. However, among non-standard varieties, foreign accents often diverge from this pattern. Speakers of foreign accents (i.e., accented English speakers, as typically examined in the prior research) are usually members of minority groups in the society and are therefore evaluated lower on status dimensions. Moreover, because they are perceived as out-group members, they also receive less favorable ratings on solidarity. In addition, depending on the degree of accentedness, these evaluations can be gradient (Giles, 1972; Grondelaers et al., 2019).

According to Dragojevic and Goatley-Soan (2022), evaluations of foreign accents are shaped by social categorization processes. Accents associated with less stigmatized groups such as Western European varieties received comparatively higher ratings on both status and solidarity than those associated with more stigmatized groups. Importantly, in this study, they were judged solely on the basis of speech without any supplementary information about the speakers such as their appearance or nationalities. This study further showed that listeners could, to some degree, infer regional or ethnic background simply from accents, and that accents perceived as originating from socioeconomically dominant countries (e.g., Western European countries) or from groups racially similar to the listeners (e.g., White) were evaluated more favorably across both status and solidarity dimensions.

In sum, previous research demonstrates that listeners' evaluations of language varieties are systematically organized along the dimensions of status and solidarity, and that foreign-accented speech in particular is often disadvantaged on both traits. These evaluative dimensions are closely tied to social categorization, perceptions of socioeconomic status, and intergroup relations. This analysis of listener bias based on language attitude extends what Rubin and colleagues termed RLS, by introducing additional dimensions and nuance.

However, most of this work has been conducted in the U.S., the U.K., and other European contexts, leaving open the question of whether such evaluations manifest differently in other sociolinguistic settings. In English-speaking and Western European contexts, socioeconomic status and inter-group preference are often conflated, as non-accented mainstream speakers may often be in-group members and perceived as having more socioeconomic power than accented speakers, while accented speakers, often immigrants, may be seen as out-group members from countries or cultures with a lower socioeconomic status.

As a result, examining accent perception in a non-English-speaking context provides an interesting opportunity to separate factors that are often conflated in English-speaking or Western European countries. In many Western contexts, foreign-accented speakers are often positioned as both out-group members and members of groups associated with lower socioeconomic status. In South Korea, however, foreignness does not necessarily carry the same socioeconomic meaning.

South Korea has recently experienced an increasing number of foreign residents and Korean learners from diverse racial and national backgrounds (Korea Immigration Service, 2024; Park, 2014). Moreover, although White foreigners are racially out-group members in Korea, they may be associated with relatively high socioeconomic status due to their perceived

connection with Western countries such as the U.S. and the U.K. At the same time, ethnic Korean speakers from other countries, such as Korean-Japanese and Korean-Americans, are increasingly visible in Korean media as celebrities who speak Korean with a foreign accent, challenging ideologies of linguistic and ethnic homogeneity. These dynamics position South Korea as a distinct sociolinguistic context in which to examine how race, ethnicity, and socioeconomic perceptions may interact with evaluations of foreign-accented speech in a different manner than they do in English-speaking contexts like the U.S. and the U.K. Therefore, the present study seeks to contribute to the further understanding of how language attitudes are shaped by specific sociolinguistic contexts.

In addition to social and cultural factors, cognitive mechanisms such as processing fluency may also influence language attitudes (Alter & Oppenheimer, 2009; Dragojevic & Giles, 2016). Listeners often attribute their negative evaluations of foreign or other non-standard accents to perceived difficulties in comprehension or fluency (Shuck, 2006), thereby placing the burden of communication largely on the speaker (Lippi-Green, 2012). Processing fluency or comprehensibility refers to the perceived ease or effort involved in understanding information, such as speech. When speech is harder to process due to, for example, a foreign accent, listeners tend to form less favorable impressions of the speaker.

1.3 Overview of the dissertation

This dissertation investigates how native Korean listeners perceive and evaluate foreign-accented speech when it is paired with different types of social information about a speaker, including race, nationality, and ethnicity. Specifically, it examines how listeners' social

expectations shape both perceptual outcomes and evaluative judgments of accented speech, and what acoustic features are relevant to accentedness judgment and language attitudes.

Chapter 2 examines the extent to which social information about a speaker modulates perceptual outcomes such as intelligibility and accentedness judgments. Using transcription accuracy and accentedness ratings, this chapter explores whether speakers associated with socioeconomically higher-status nationalities (e.g., the U.S.) are perceived more favorably, whether speakers associated with in-group ethnic identities (e.g., Korean-American) or race (e.g., Asian) receive perceptual advantages, and whether mismatches between speech and social information (e.g., Korean nationality paired with foreign-accented speech) result in greater perceptual degradation than congruent conditions. Through these comparisons, Chapter 2 focuses on which specific aspect of speakers' social information influences perception of foreign-accented speech.

Building on the perceptual results observed in Chapter 2, Chapter 3 shifts the focus from perception to social evaluation, examining native Korean listeners' language attitudes toward speakers associated with different social information. This chapter investigates whether speakers associated with higher-status nationalities are evaluated more favorably on status dimensions, whether in-group racial or ethnic identities (e.g., Korean-American) receive preferential evaluations on solidarity dimensions relative to out-groups (e.g., White American), and how these results observed in South Korea are compared with those reported in other social contexts. Importantly, it further examines the relationship between language attitudes and perceptual outcomes from Chapter 2, assessing whether listeners' social evaluations such as bias or stereotype are systematically associated with intelligibility and accentedness judgments.

In Chapter 4, the focus turns to the acoustic features of speech, comparing the Korean productions of American learners of Korean with those of native Korean speakers. Focusing on stop consonants and vowels, this chapter examines acoustic properties such as voice onset time (VOT), fundamental frequency (F0), vowel duration, and vowel formants. Building on these acoustic analyses, Chapter 5 investigates which acoustic features are associated with accentedness judgments and language attitudes, and evaluates the extent to which the perceptual and evaluative patterns observed in earlier chapters can be linked to specific acoustic characteristics of the speech.

Chapter 6 summarizes the main findings of this dissertation, highlights the theoretical and empirical contributions to the field, and suggests directions for future research on foreign-accented speech perception

II. SOCIAL EXPECTATIONS IN INTELLIGIBILITY AND ACCENTEDNESS JUDGMENTS

2.1 Introduction

The purpose of this chapter is to examine how information about the speaker such as race, nationality and ethnicity associated with intergroup bias, socioeconomic status, and expectation mismatch interact in shaping the perception of foreign-accented speech in South Korea (hereafter, Korea), a context in which East Asians constitute the racial majority and foreignness is not necessarily associated with lower socioeconomic status. Although Korea has traditionally been considered a linguistically and ethnically homogeneous society, the number of non-citizen residents living in South Korea has increased substantially over the past decades, from approximately 110,000 in 1995 to a peak of 2.4 million in 2024 (Park, 2014; Korea Immigration Service, 2024). Given this, understanding how foreign-accented speakers are perceived in Korea has become an increasingly important issue.

In the Korean context, once Korean listeners are informed that both White Americans and Korean-Americans are from the U.S., the two groups may be perceived as sharing similar socioeconomic characteristics based on their American background. However, even when such background information leads to comparable inferences about socioeconomic status, racial differences can still activate distinct social stereotypes and shape listener expectations depending on the context where listeners are situated (Kutlu et al., 2022; Zebrowitz et al., 2007). Specifically, although White American individuals are a racial minority and typically perceived as out-group members in Korea, they may be associated with relatively higher socioeconomic status and prestige, as White appearance has long been linked to Western European countries, the

U.S., and the U.K. By contrast, Korean-Americans may be perceived as more in-group based on their East Asian appearance and Korean ethnic background, and thus elicit more favorable attitudes due to in-group favoritism. At the same time, Asian appearance may also be associated with lower socioeconomic status relative to White appearance, despite shared nationality. These competing associations raise the question of whether Korean listeners perceive accented Korean differently depending on a speaker's appearance.

By comparing White American and Korean-American guises, the present study first addresses whether race-based visual cues influence the perception of accented Korean speech. In other words, the present study investigates whether Reverse Linguistic Stereotyping (RLS) occurs and whether it is related to the race of the accented speaker. However, because Korean-Americans can plausibly be considered both racially and ethnically in-group to Korean listeners, this comparison alone may obscure the effects of race and ethnicity. To disentangle these factors, a Chinese guise was included. Chinese speakers share East Asian racial features with Koreans (racial in-group) but are ethnically and nationally out-group members, while also being associated with substantial economic power, given that China has the world's second-largest GDP after the United States (World Bank, 2024). Including this condition allows us to tease apart racial effects from ethnic in-group effect.

In addition to social categorization effects, this study also investigates expectation mismatch effects as proposed in the exemplar-based account of speech perception (McGowan, 2015). In Korea, nationality is strongly tied to language expectations, arguably more so than ethnicity (i.e., Korean-American). Although the term "Korean" can be interpreted in various ways, it is assumed that Korean listeners primarily associate it with Korean citizenship and native Korean accent. For this reason, a Korean guise is included as a mismatch condition. By

presenting an image of a Korean individual paired with the description "Korean," listeners' expectations for native accent will increase. Therefore, pairing this description with foreign-accented Korean speech creates a mismatch between expected and actual speech, which, according to the exemplar theory, should negatively influence speech perception relative to matched conditions such as the Korean-American or White guises (McGowan, 2015). To address these research questions (see Section 2.1.1), Korean listeners were presented with identical foreign-accented Korean auditory stimuli paired with one of five visual and descriptive guises: a White American, a Korean-American, a Chinese, a Korean, and a silhouette baseline (See McGowan, 2015 for the use of silhouette). In addition to the visual information, each guise was accompanied by written information that the speaker was an American learner of Korean (paired with a White image), a Korean-American (paired with an Asian image), a Chinese learner of Korean (paired with the same Asian image), a Korean (paired with the same Asian image), or a foreign learner of Korean (a silhouette image). The written descriptions were included to ensure that the intended social category was activated (Vaughn, 2019) and to prevent unintended interpretations.

For the Chinese condition, a pilot study was conducted based on the concern that Korean listeners might be familiar with the typical Chinese-accented Korean, and that using a Chinese guise in this study could produce mismatch effects. In the pilot study, where 13 Korean listeners were asked to guess the speaker's first language, responses varied widely. Only three identified the speaker as a native English speaker, and two identified the speaker as Chinese. The results of this pilot study suggested that pairing the English-accented Korean speech with a "Chinese" description would not elicit an unintended mismatch. Regarding the silhouette baseline, social

information related to race was minimized, allowing this condition to serve as a reference point for assessing listeners' perception of accented speech without specific social cues.

Participants completed two tasks: 1) a transcription task measuring intelligibility and 2) an accentedness rating task measuring perceived degree of foreign accent. Based on these considerations, different outcomes were predicted as summarized below. For ease of reference, the experimental conditions are labeled A-E, corresponding to the baseline (A), White American (B), Korean-American (C), Chinese (D), and Korean (E) conditions (see Table 2.1).

2.1.1 Summary of research questions

(1) **Research Question 1**: Does race-based speaker information influence the perception of accented Korean speech?

In other words, does being perceived as a White speaker versus a speaker of color affect accent perception even outside of predominantly white contexts? To address this question, three guises were compared: White American, Korean-American, and Chinese. If a White speaker receives a positive bias (or, conversely, if speakers of color receive a negative bias), task scores would be more favorable, indicating higher intelligibility and lower accent ratings, for the White guise than for the Korean-American or Chinese guises.

(2) **Research Question 2**: Does ethnic group membership influence the perception of accented Korean speech?

Here the comparison is also among the White American, Korean-American, and Chinese guises. As an alternative to (1), Reverse Linguistic Stereotyping may operate not primarily on race but on perceived group membership, such that ethnically in-group speakers (i.e., Korean-Americans)

receive a positive bias in the Korean context. If this is the case, task scores would be more favorable for the Korean-American guise than for the White or Chinese guises.

Furthermore, Reverse Linguistic Stereotyping may work on multiple levels. Even if race may have overall negative influences, an ethnic in-group effect may still benefit the Korean-American guise. In this scenario, the White guise would receive the highest intelligibility scores and the lowest accentedness ratings due to the race-based bias, followed by the Korean-American guise with the second-highest intelligibility scores and the second-lowest accent ratings due to the ethnicity-based bias.

(3) **Research Question 3**: Does language expectation influence the perception of accented Korean speech?

If Korean listeners' perception of accented Korean is primarily shaped by linguistic expectations (e.g., Gnevsheva, 2018; McGowan, 2015), task scores were expected to be least favorable for the Korean guise, reflected in lower intelligibility scores and higher accentedness ratings compared to the other guises. In this condition, listeners would anticipate fluent native Korean speech and encountering foreign-accented Korean does not match that expectation.

Section 2 below describes the pilot studies conducted to create and test experimental stimuli, and Section 3 presents the experiments conducted to address the research question.

Table 2.1
Overview of experimental conditions and associated social expectations

	Race (Image)	Ethnic relation	Socioeconomic status	Expectation
A. Baseline	Silhouette	Unspecified	Unspecified	Unspecified
B. White American	White	Out-group	High	Match
C. Korean-American	East Asian	In-group	High	Match
D. Chinese	East Asian	Out-group	High	Match
E. Korean	East Asian			Mismatch

Note. *Ethnic relation* refers to the perceived ethnic in-group or out-group relationship between the speaker and Korean listeners. *Expectation* indicates whether listeners' likely social expectations about the speaker's accent match the actual accent they hear.

2.2 Pilot studies for stimulus creation and selection

2.2.1 Pilot 1

Following the prior studies (Babel & Russell 2015; McGowan, 2015), noise was added to the auditory stimuli used in the transcription task to avoid the ceiling effects of the task. Pilot 1 was conducted to determine the most suitable signal-to-noise ratio (SNR) and speaker for use in the experiments.

2.2.1.1 Stimuli

The stimulus materials consisted of 60 sentences originally developed by Song et al. (2023) as part of the Korean Speech Recognition (KSR) corpus. This corpus was designed to evaluate speech recognition and lexical processing through sentences categorized by predictability levels: predictable, neutral, and anomalous, allowing it to test both native and non-native Korean speakers in terms of sentence complexity and vocabulary.

For the current pilot, 60 sentences were selected from the 120 neutral sentences. The neutral sentences rather than anomalous ones were selected because it minimizes any potential impact of nonsensical words on accentedness judgments, as hearing an anomalous word can increase the perceived foreignness of the speech (Bergeron & Trofimovich, 2017; Varonis & Gass, 1982). These sentences were also selected with a final word cloze probability for native listeners between 20-40%. This range minimized the predictability of the target word, encouraging participants to focus on listening to the speech signal rather than guessing. The sentences were controlled for the number of syllables, with sentences containing from 13 to 15 syllables and the final words containing 3-4 syllables. The sentences in the original Korean and English translation are presented in Appendix A.

Six female speakers were recorded producing the selected 60 sentences: two female native Korean speakers (aged 21 and 31) and four female American learners of Korean (aged 20, 22, 22 and 28). Of the four American learners, two had completed one year, and the other two had completed two years in the Korean program at the University of Oregon. Recordings were made individually in the sound-attenuated booth using a portable recorder (TASCAM DR-100mkII) with a 44.1kHz sampling rate and a 16-bit quantization level, using a clip microphone (Audio-Technica AT8537). Sentences were presented on a computer screen in a random order for each speaker controlled by the software Gorilla during recording. The recording lasted approximately 11 minutes for native speakers and 30 minutes for non-native speakers. The recorded utterances were later separated into individual sentence files and amplitude-normalized to 70 dB.

To determine the optimal SNR and speaker for the intelligibility task, these files were mixed with white noise in Praat (version 6.4.04; Boersma & Weenink, 2024). Two noise levels were tested: -3 dB SNR (noise at 73 dB) and -1 dB SNR (noise at 71 dB).

The total 360 sentences (60 sentences x 6 speakers) were divided into 6 lists to make the length of experiment manageable. Each list contained 10 sentences per speaker, with no sentence repeats. This resulted in two sets of 6 lists, with one set for each noise level: the -3 dB SNR set and the -1 dB SNR set.

2.2.1.2 Participants

A total of nineteen participants (10 female, 9 male), aged from 22 to 42 (M = 31.9 SD = 5.12), were recruited through word-of-mouth and the network of friends and acquaintances of the author. All participants were born and raised in Korea but residing in the U.S. at the time of the study (length of stay: 2 months to 21 years). They reported no hearing issues and were compensated at a rate equivalent to $10 per hour. Of these 19 participants, 12 took part in the -3dB SNR condition (with one removed due to incomplete data collection), and 7 participated in the -1 dB SNR condition. Because pilot testing indicated that the -3 dB SNR stimuli were too difficult for listeners, an additional -1 dB SNR condition was tested to reduce task difficulty. As such, these conditions were tested sequentially.

2.2.1.3 Procedure

Pilot 1 involved a transcription task and an accentedness judgment task, as the main experiments included these two tasks. The study was conducted online using the Gorilla Experiment Builder (Anwyl-Irvine et al., 2020), a platform facilitating various experimental

tasks. Participants received a link to the experiment directly from the researcher and were asked to complete the experiment in a quiet setting using their own headphones or earphones. Specific instructions were provided within the link as they progressed through each part.

The study included three components: a transcription task, an accentedness judgment task, and a demographic/language background questionnaire. In the transcription task, participants listened to stimuli in noise and transcribed (typed using a keyboard) what they heard in a text box. Once they finished transcribing, they clicked the 'next' button to proceed to the next trial and audio clips were automatically played only once, with a one-second delay before starting. They were instructed to transcribe as accurately as possible but could enter a question mark for words they were unsure of. Prior to the test task, participants first practiced transcribing two example stimuli.

In the accentedness judgment task, sentences without noise were presented in a randomized order for each participant. Participants rated accentedness of each stimulus on a 9-point Likert scale which appeared in the center of the screen, ranging from 0 (sounding like a native Korean speaker) to 8 (sounding heavily accented). After the accentedness rating task, participants completed a demographic and language questionnaire.

2.2.1.4 Assessment of transcription

In this report, the labels K are used to refer to the native Korean speakers, while L represents the Korean learners. Two sets of intelligibility scores were calculated: the whole sentence score and the target word score. The whole sentence score reflects the percentage of correctly transcribed words across the entire sentence, while the target word score measures the accuracy of transcribing the final word in each sentence. Participants in the -3 dB SNR condition

found the transcription task excessively difficult. The accuracy rates for entire stimuli (not just target words but all words within each sentence) were as follows: K1 – 40%, K2 – 49%, L1 – 25%, L2 – 20%, L3 – 17%, and L4 – 34%. To achieve the whole sentence accuracy of 60-70% for native speakers' speech, the noise level -1 SNR was tested subsequently. The 60-70% target was selected to place performance at an intermediate level of difficulty, thereby avoiding both ceiling and floor effects.

In the -1 dB SNR condition, sentence accuracy rates improved: K1 – 74%, K2 – 71%, L1 – 47%, L2 – 50%, L3 – 44%, and L4 – 68%. Accentedness ratings for each speaker were: K1 – 0.21, K2 – 0.22, L1 – 4.71, L2 – 7.28, L3 – 6.55, and L4 – 5.44 (0 = no accent, 8 = heavy accent). Based on these results, K1, with a slightly higher accuracy rate (74%) than K2 (71%), and L4 (68%), whose accuracy rate was similar to those of native Korean speakers, were selected for the second pilot study.

Additionally, on the 9-point Likert scale used for accentedness ratings, a score of 5 represents a mid-level accent. Therefore, L4's rating of 5.44 indicates a moderate degree of foreign accent. The speaker was considered appropriate for the study, as presenting a Korean-American face with a very heavy accent in the main study was not desirable.

2.2.2 Pilot 2

Pilot 2 was conducted to evaluate the accuracy rate of each sentence in the transcription task in order to select 30 out of 60 sentences for use in the main experiments. The accent rating task was not included.

2.2.2.1 Stimuli

Stimuli consisted of the same sentences used in Pilot 1, recorded by K1 and L4, mixed with white noise at -1 dB SNR. Each participant was randomly assigned to one list containing 60 sentences from one speaker.

2.2.2.2 Participants

Twenty participants (13 female, 7 male) were recruited from Prolific, an online research platform with a diverse participant pool. All participants identified Korean as their first and dominant language. Due to the limited Korean subject pool on Prolific, participants' countries of residence were not controlled. The participants ranged in age from 20 to 51 (M = 32.2, SD = 6.54) and were compensated at a rate equivalent to $10 per hour. None of the participants reported any hearing issues or language-related disorders.

2.2.2.3 Procedure

Participants were redirected from Prolific to the experiment run on Gorilla. Pilot 2 included only the transcription task and language background questionnaire. The transcription procedure mirrored that of Pilot 1.

2.2.2.4 Results

K1's mean sentence accuracy was 78.27% (SD = 16.61%), with a mean target word accuracy of 63% (SD = 29.70%). L4's mean sentence accuracy was 71.51% (SD = 17.99%), and target word accuracy was 63.67% (SD = 28.34%).

L4's sentence accuracy rates were then divided into three performance levels based on the distribution of accuracy scores: high, mid, and low (Whole: high > 66.68%, low <54.67%;

Target: high > 80%, low < 36.66%). Ten sentences from each performance level (total 30) were selected to include a range of performance levels. However, as the performance levels for whole sentences and target words did not perfectly align, some sentences had, for example, a mid-level accuracy for whole sentences but a low-level accuracy for target words. Including a range of performance levels served two purposes: first, to investigate potential correlations between intelligibility performance and accentedness judgments, and second, to reflect real-life conversational conditions where the intelligibility of foreign-accented speech may vary even within the same speaker.

The selected 30 stimulus sentences are presented in Appendix A.

2.3 Experiment

2.3.1 Methodology

2.3.1.1 Design

This study employed a between-subjects design to investigate the influence of visual cues and speaker description on Korean listeners' perceptions of intelligibility and accentedness in foreign-accented Korean. Participants were assigned to one of five experimental conditions, each of which paired identical audio stimuli with visual and descriptive cues designed to elicit distinct social expectations (see Table 2.1).

2.3.1.2 Audio stimuli

The stimulus materials consisted of 30 sentences originally developed by Song and colleagues (2023). Based on the results of Pilot studies 1 and 2, 30 sentences (see Appendix A)

recorded by a female native Korean speaker (K1) and a female non-native speaker (L4) speaker were selected, and white noise at -1 SNR was added to the speech for the intelligibility task.

The female native Korean speaker was an exchange student at the University of Oregon, with less than seven months of living abroad experience at the time of recording. Her audio was used as an anchor only in accentedness rating task. The female non-native speaker was an American intermediate learner of Korean who had completed second-year Korean courses at the University of Oregon and was enrolled in the third-year Korean program at the time of recording. In Pilot 1, the non-native speaker stimuli's foreign accent was assessed as moderately accented, with a rating of 5.44 on a 9-point Likert scale (0 = no accent, 8 = heavy accent).

Before the experimental stimuli were presented in the intelligibility task, listeners completed two practice items without visual cues to adjust their volume and familiarize themselves with listening to speech in noise. No practice items were included in the accentedness judgment task, as the audio was presented without noise and the Likert scale interface was straightforward.

2.3.1.3 Visual stimuli

This study employed an inverted matched guise task to investigate listeners' perceptions of accent and speaker characteristics. In this approach, participants were presented with identical audio paired with different visual and written representations to assess the influence of visual and written social cues on listeners' perceptions. Unlike the traditional matched guise task, where a single speaker produces different linguistic varieties to gauge attitudes toward language features (Lambert et al., 1960), the inverted matched guise task uses the identical audio while varying the

accompanying social information. This design enables researchers to tease apart the effect of social cues on listeners' judgments of the same linguistic input (Campbell-Kibler, 2007). Three visual conditions were included as shown in Figure 2.1. Using different descriptions of the Asian guise's racial and ethnic background resulted in five guise conditions:

1. A silhouette (baseline) described as "Foreign learner of Korean" ("한국어를 배우는 외국인")
2. A photograph of a White face described as "American learner of Korean" ("한국어를 배우는 미국인")
3. A photograph of an Asian face described as "Korean-American" (재미교포)
4. A photograph of an East Asian face described as "Chinese learner of Korean" ("한국어를 배우는 중국인")
5. A photograph of an Asian face described as "Korean" (한국인)

The descriptions were provided to ensure that the intended social cues were clearly conveyed, particularly for the East Asian-looking speakers, whose social cues are distinguished by the descriptions.

Each participant was assigned to only one of the five conditions, in which the corresponding image was displayed throughout each trial in both the intelligibility task and the accentedness judgment rating task. The photographs used in the study were obtained from speakers who participated in the recording phase of the study and who consented to the use of their images. Although these individuals participated in the recording phase, the native and non-native audio stimuli used across all guise conditions did not necessarily correspond to the individual in the

photograph. The size of all the images was normalized to a 1:1 ratio to ensure consistency across conditions. While the images were not explicitly normed for attributes such as attractiveness, racial or gender typicality, they were instructed to maintain neutral facial expressions and remove glasses to ensure visual consistency. Additionally, a minimalistic silhouette image was generated using AI to eliminate identifiable social characteristics such as race.

Figure 2.1
Face used in the experiment. From left to right: silhouette, White, and East Asian face.



2.3.1.4 Participants

A total of 149 native Korean speakers (45 male, 80 female; 24 were removed) living in South Korea were recruited through word of mouth, the network of friends and acquaintances of the researcher, and Everytime (https://everytime.kr), a social networking and academic resource site widely used among university students in South Korea. Twenty-four participants were automatically excluded due to exceeding the time limit (1 hour); most exclusions involved participants who opened the link but did not proceed with the task. All participants had no history of hearing or linguistic disorders, and were born and raised in South Korea, with little to no experience living abroad (less than 1 year) before the age of 18. Participants' age ranged from

19 to 58 years (M = 29.7, SD = 6.8). 125 participants were randomly assigned to one of five conditions, with each condition completed by 25 participants.

2.3.1.5 Procedure

The experiment was conducted online using Gorilla Experiment Builder (https://gorilla.sc), and participants received the experiment link directly from the researcher. They were instructed to complete the experiment in a quiet place using their own headphones or earphones. Specific instructions were also provided within the link. Since participants had to transcribe what they heard using a keyboard, access to the experiment on tablets or phones was blocked to ensure they used a desktop or laptop. The entire experiment typically took 15-20 minutes, and the link automatically closed if left open for over 1 hour. Although environmental factors such as internet speed and types of headphones or earphones could not be strictly controlled, participants used their own familiar keyboards and audio devices and could adjust volume as needed in comfortable environments.

Upon clicking the link, they first saw the consent form and could proceed only after providing consent. Next, Gorilla automatically displayed a pop-up box to confirm that audio was playing correctly. After that, the experiment proceeded to the intelligibility task, followed by the accentedness rating task, and finally by background survey. Instructions for the intelligibility task informed participants that they would hear 30 sentences in noise and that their job was to transcribe what they heard. They were also informed that each audio clip would play automatically with a 1-second delay and only once as they proceed to the next trial. Participants completed two practice items before starting test items, which allowed participants to adjust their volume and familiarize themselves with listening to speech in noise and with the task format.

After the practice, condition-specific prompts were provided in Korean saying they would hear Korean speech from:

- Silhouette: “a foreigner learning Korean”
- White: “an American learning Korean”
- Asian: “a Korean-American”
- Asian: “a Chinese learning Korean”
- Asian: “a Korean”

Each trial displayed the prompt in Korean “If there are words you could not hear clearly, you may enter ‘?’, but please try to transcribe what you hear as accurately as possible,” Then, 30 sentences in noise were randomly presented with condition-specific images (silhouette, White, Asian face) and a text box for transcription located below the image. Participants were instructed to click a button when they finished transcribing an item, which presented the next trial with a 1-second delay.

After the intelligibility task, participants moved onto the accentedness judgment task. In this task, the same 30 non-native speaker stimuli used in the intelligibility task and the same set of 30 sentences recorded by a native Korean speaker (K1) were used. The native speaker stimuli were added in this task to provide a reference point for participants’ accentedness judgments. Therefore, a total of 60 sentences were randomly presented with corresponding images; one of the three images paired with the speaker stimuli (see Figure 2.1) and the second Korean face paired with the native Korean stimuli. A 9-point Likert scale was displayed below the image, ranging from 0 (sounding like a native Korean speaker) to 8 (sounding heavily accented).

Participants were instructed to judge the degree of accentedness of the stimuli. To prevent participants from choosing one score as they may have noticed that they were listening to the same two speakers, the prompt was given in Korean saying "How much of a foreign accent did you perceive in the sentence?" ("들으신 문장에서 외국 억양을 얼마나 느끼셨습니까?"). Participants were instructed that, even for the same speaker, the degree of accent may vary across sentences. Then participants rated each stimulus according to how native or accented it sounded.

Finally, participants completed a demographic and language background questionnaire, which included questions on their Korean dialect, second language, experience living abroad, interactions with foreigners speaking Korean, and exposure to media featuring foreigners speaking Korean. They were also asked about their views on the importance of accent of pronunciation when learning a second language.

2.3.2 Analysis

2.3.2.1 Intelligibility task

Listeners' transcription responses in the intelligibility task were first manually normalized for spelling errors. Spelling errors involving more than two characters and that changed the meaning of the original word were coded as incorrect (see Appendix B). For example, '끓였다' /kku-lye-tta/ meaning 'boiled,' transcribed as '그렸다' /ku-lye-tta/ meaning 'drew,' was coded as incorrect. However, simple typos including a single character error or errors that were orthographically incorrect but phonologically accurate were coded as correct. Additionally, particles were generously coded to minimize penalization for minor omissions or substitutions.

For whole-sentence accuracy, transcription response was summarized as counts of correct and incorrect words per sentence. For target word accuracy, the final word in each sentence was coded as 1 if correct and 0 if incorrect, in a separate column.

To evaluate the effect of social information associated with the speaker (*Guise* condition) on intelligibility, Bayesian generalized linear mixed-effects models were fit in R (R Core Team, 2026) using the `brm` function in the *brms* package (Bürkner, 2017). For whole-sentence accuracy, a binomial likelihood was specified with the number of correctly transcribed words as the outcome and the total number of words in each sentence as the number of trials. For target word accuracy, a Bernoulli likelihood was specified. Four chains were run with 8,000 iterations each to ensure adequate exploration of the posterior distribution. Convergence was checked by inspecting potential scale reduction factors ($\hat{R}$), effective sample sizes (ESS) for both bulk and tail of the posterior distributions, and trace plots of the Markov chain Monte Carlo (MCMC) samples. All parameters showed convergence ($\hat{R} = 1.00$), with sufficient effective sample sizes (> 1000) and well-mixed chains.

Models included *Guise* condition as a fixed effect and random intercepts for *Participant* and *Item* (sentence) to account for baseline differences in listener performance and item difficulty. *Guise* was a categorical factor treatment-coded with the Baseline condition as the reference level. Given the between-subjects design, random slopes for *Guise* by-*Participant* were not included. Random slopes by-*Item* were considered but were not retained as they did not substantially affect posterior estimates.

To examine differences between *Guise* conditions, posterior contrasts were done based on the fitted Bayesian models. Inference was based on the posterior distributions of these

contrasts, summarized by posterior medians, 95% credible intervals, and posterior probabilities that contrasts exceeded zero.

2.3.2.2 Accentedness rating task

To investigate accentedness ratings, Likert scale scores ranging from 0 to 8 were used as the dependent variable. To evaluate the effect of social information associated with the speaker (*Guise* condition) on accentedness ratings, a Bayesian cumulative ordinal mixed-effects model was fit in R using the `brm` function in the *brms* package (Bürkner, 2017). The model specified a cumulative (ordered logistic) likelihood with a logit link.

Four chains were run with 12,000 iterations each. Convergence was assessed by inspecting potential scale reduction factors ($\hat{R}$), effective sample sizes for both bulk and tail of the posterior distributions, and trace plots of the MCMC samples. All parameters showed satisfactory convergence ($\hat{R} = 1.00$), with sufficient ESS (> 1,000) and well-mixed chains.

The model included *Guise* condition as a fixed effect and random intercepts for *Participant* and Item to account for baseline differences in participants' rating tendencies and item-level differences. *Guise* was treated as a categorical factor, treatment-coded with the Baseline condition as a reference level. Given the between-subjects design, random slopes for *Guise* by-*Participant* were not included. Random slopes by-*Item* were considered but were not retained as it did not substantially affect posterior estimates.

To examine differences between *Guise* conditions, posterior contrasts were done based on the fitted Bayesian model. Inference was based on the posterior distributions of these contrasts, summarized by posterior medians, 95% credible intervals, and posterior probabilities that contrasts exceeded zero.

2.3.2.3 Correlation

Although correlation analyses were not part of the main research questions, sentence difficulty levels were controlled to see potential correlations between intelligibility and accentedness judgments. Therefore, correlations between these two tasks were examined, and additional correlation analyses were conducted to explore possible relationships between perceptual results and participants' demographic and language background information.

Participants' mean scores for intelligibility and accentedness ratings were calculated. Spearman rank correlations were calculated to examine the relation between intelligibility and accentedness. In addition, relationships between perceptual outcomes and participant background variables were examined using several correlation tests. For categorical predictors such as gender and dialect type, Kruskal-Wallis tests were used. For binary predictors such as media exposure and whether participants had foreign friends speaking Korean, Pearson (point-biserial) correlations were conducted. For continuous predictors such as age and age of acquisition of English (AOA), Spearman correlations were used.

2.3.3 Results

2.3.3.1 Intelligibility task

Posterior estimates from the Bayesian models are reported in Table 2.2 (whole sentence accuracy) and Table 2.3 (target word accuracy), and posterior contrasts across guise conditions are reported in Table 2.4 (whole sentence) and Table 2.5 (target word). Figure 2.2 (whole sentence) and 2.3 (target word accuracy) provides the visual summary of the estimated condition effects on intelligibility relative to the Baseline condition.

Estimates are summarized by posterior medians and 95% credible intervals (CrIs) on the logit scale, with the Baseline condition serving as a reference level. Posterior medians for the conditions were close to zero on the logit scale, and the 95% CrIs for all conditions crossed zero. Only the intercept for whole sentence intelligibility was reliably above zero.

Posterior contrasts across guise conditions (Table 2.4 and Table 2.5) are summarized by posterior medians, 95% CrIs, and posterior probabilities that contrasts exceed zero (P(>0)). The posterior median represents the estimated difference between two conditions on the logit scale, while the 95% CrIs indicate the range within which the true value of the contrast lies with 95% posterior probability. When a credible interval includes zero, the direction and magnitude of the difference between two conditions cannot be estimated with certainty.

Table 2.2
Posterior medians and 95% CrIs for the effects of condition on whole sentence accuracy (logit scale)

Condition	Median (logit)	Lower 95% CrI	Upper 95% CrI
Intercept (Baseline)	0.69	0.27	1.11
White	0.15	-0.24	0.52
Korean-American	0.07	-0.30	0.46
Chinese	-0.11	-0.50	0.27
Korean	-0.04	-0.43	0.36

Table 2.3
Posterior medians and 95% CrIs for the effects of condition on target word accuracy (logit scale)

Condition	Median (logit)	Lower 95% CrI	Upper 95% CrI
Intercept (Baseline)	0.03	-0.64	0.70
White	0.32	-0.21	0.85
Korean-American	0.10	-0.44	0.62
Chinese	-0.24	-0.77	0.30
Korean	-0.08	-0.61	0.45

1) **RQ1: Race bias (White advantage) vs. RQ2: Ethnicity - in-group bias (Korean-American advantage)**

If race-based stereotype dominates perception, the White guise would be expected to yield higher intelligibility than both the Chinese and Korean-American guises. If ethnic in-group bias were to dominate, the Korean-American guise would be expected to yield higher intelligibility than both the Chinese and White guises.

Although White and Korean-American guises show numerically higher intelligibility scores, the contrasts were not necessarily different statistically. For whole sentence accuracy, contrasts comparing the White guise with the Chinese and Korean-American guises included zero, indicating no reliable differences. The results suggest no bias based on race and ethnicity membership for the sentence level transcription accuracy.

For target word accuracy, the White and Chinese contrast (Median = 0.56, 95% CrI [0.03, 1.10], P(>0) = 0.98) showed reliable evidence of higher intelligibility for the White guise, whereas the White and Korean-American contrast (Median = 0.22, 95% CrI [-0.31, 0.76], P(>0) = 0.79) and the Korean-American and Chinese contrast (Median = 0.33, 95% CrI [-0.19, 0.87], P(>0) = 0.89) included zero. The results suggest that race-based bias may emerge only in comparisons between ethnic out-group speakers.

2) **RQ3: Expectation mismatch**

If expectation mismatch effect dominates, the Korean guise would be expected to yield lower intelligibility than all other matched guises. For both whole sentence and target word accuracy, posterior contrasts involving the Korean guise included zero, indicating no reliable mismatch effect in intelligibility.

In summary, no evidence of race- and ethnicity-based bias was found in sentence-level transcription accuracy for accented Korean. However, at the target word level, higher intelligibility was observed for the White guise relative to the Chinese guise, suggesting a possible race-based advantage for the White guise. The mismatch between the speech and expectation of the speaker did not result in any difference in intelligibility.

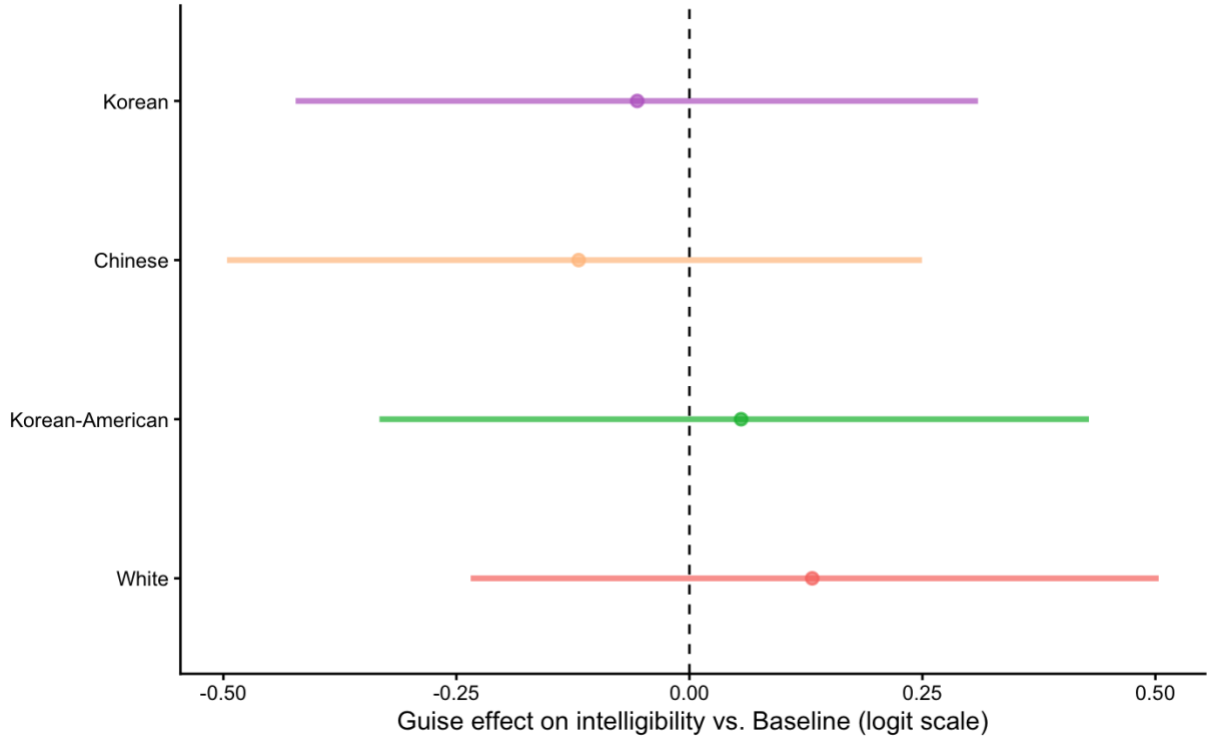
Table 2.4
Posterior contrasts for whole sentence accuracy (logit scale)

Contrast	Median (logit)	P(>0)	95%CrI
Baseline – White	-0.15	0.22	[-0.52, 0.24]
Baseline – Korean-American	-0.07	0.36	[-0.46, 0.30]
Baseline – Chinese	0.11	0.72	[-0.27, 0.50]
Baseline – Korean	0.05	0.59	[-0.36, 0.43]
White – Korean-American	0.08	0.65	[-0.31, 0.45]
White – Chinese	0.25	0.91	[-0.12, 0.64]
White – Korean	0.19	0.84	[-0.19, 0.56]
Korean-American - Chinese	0.18	0.83	[-0.19, 0.56]
Korean-American - Korean	0.11	0.72	[-0.26, 0.50]
Chinese – Korean	-0.07	0.36	[-0.46, 0.31]

Note. P(>0) indicates the posterior probability that the contrast is greater than zero.

Figure 2.2

Estimated effects of guise condition on whole sentence intelligibility relative to the Baseline (logit scale)



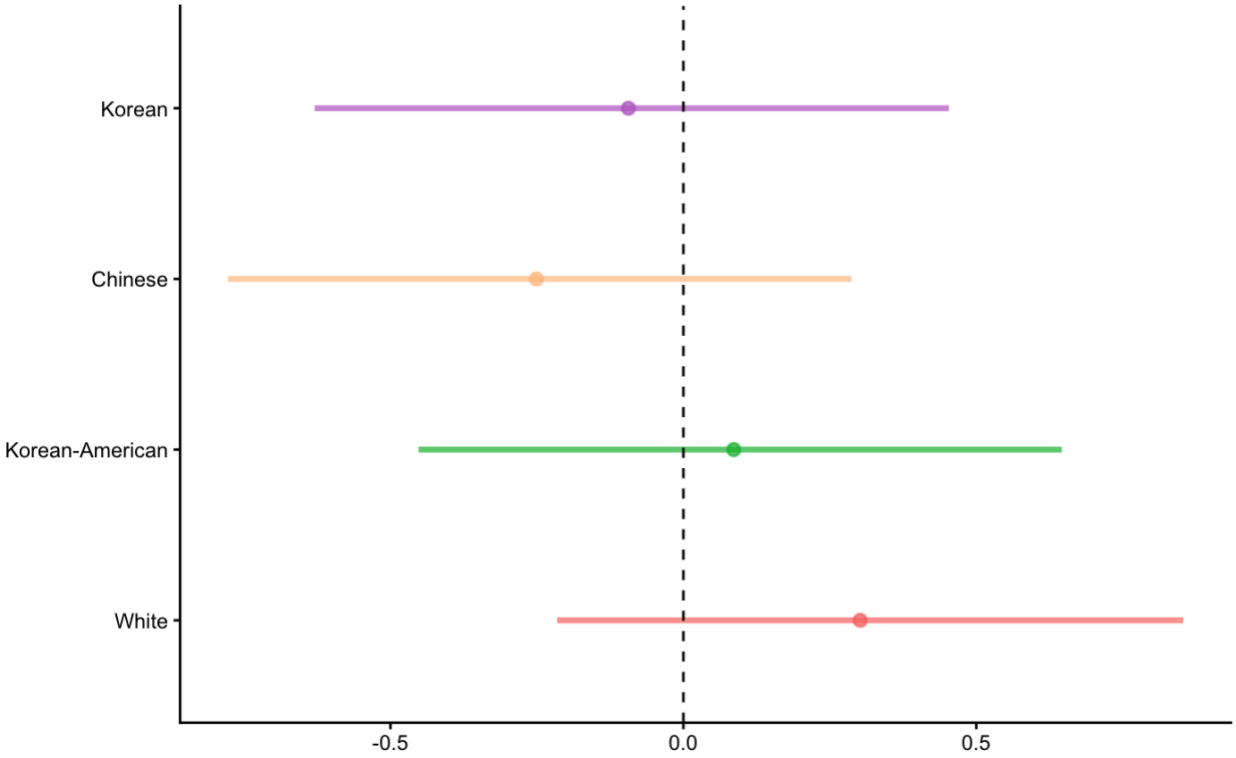
Note. Points indicate posterior medians and horizontal bars represent 95% credible intervals for each condition's effect on whole sentence intelligibility. The vertical dashed line at zero indicates no difference from the Baseline.

Table 2.5

Posterior contrasts for target word accuracy (logit scale)

Contrast	Median (logit)	P(>0)	95%CrI
Baseline – White	-0.32	0.12	[-0.85, 0.21]
Baseline – Korean-American	-0.10	0.36	[-0.62, 0.44]
Baseline – Chinese	0.23	0.81	[-0.30, 0.77]
Baseline – Korean	0.08	0.61	[-0.45, 0.61]
White – Korean-American	0.22	0.79	[-0.31, 0.76]
White – Chinese	0.56	0.98	[0.03, 1.10]
White – Korean	0.39	0.93	[-0.14, 0.93]
Korean-American - Chinese	0.33	0.89	[-0.19, 0.87]
Korean-American - Korean	0.17	0.74	[-0.35, 0.71]
Chinese – Korean	-0.16	0.28	[-0.69, 0.37]

Figure 2.3
Estimated effects of guise condition on target word intelligibility relative to the Baseline (logit scale)



2.3.3.2 Accentedness rating task

Accentedness ratings were analyzed using a cumulative mixed-effects model, which estimates category thresholds along the accentedness scale together with condition-specific effects. Posterior estimates for the effects of guise conditions on perceived accentedness relative to the Baseline condition are shown in Table 2.6, and posterior contrasts across conditions are presented in Table 2.7.

Posterior estimates are summarized by posterior medians and 95% CrIs on the logit scale. Positive values indicate higher perceived accentedness relative to the Baseline condition. Among the conditions, the Korean guise showed the largest posterior median and a 95% CrI

excluding zero, suggesting reliable evidence of higher perceived accentedness relative to the Baseline condition. The other conditions (White, Korean-American, and Chinese) showed positive posterior medians but their credible intervals crossed zero, indicating less certain differences from the Baseline condition.

Table 2.6
Effects on perceived accentedness (logit scale)

Condition	Median (logit)	Lower 95% CrI	Upper 95% CrI
White	0.89	-0.60	2.36
Korean-American	1.40	-0.12	2.88
Chinese	1.14	-0.40	2.68
Korean	2.02	0.53	3.50

1) RQ1: Race bias (White advantage) vs. RQ2: Ethnicity - In-group bias (Korean-American advantage)

If race-based bias dominates perception the White guise would be expected to receive lower accentedness ratings than both the Chinese and Korean-American guises. If ethnic in-group bias dominates, the Korean-American guise would be expected to receive lower accentedness ratings than the White and Chinese guises.

Posterior contrasts among the White, Korean-American, and Chinese guises included zero in their credible intervals (Table 2.7), which indicates no reliable differences in accentedness ratings among these conditions. Although the posterior medians suggest directional tendencies across some comparisons, the magnitudes of these effects remain uncertain. These results show no reliable evidence of race- and ethnic-based bias in perceived accentedness judgments.

2) RQ3: Expectation mismatch

If expectation mismatch dominates perception, the Korean guise would be expected to receive higher accentedness ratings than the other guises. The Korean – Baseline contrast showed the largest posterior median and a 95% credible interval excluding zero (Median = 2.02, 95% CrI [0.53, 3.50], P(>0) =.99), and it indicates reliable evidence that the Korean guise, which is a mismatched condition, was perceived as more accented than the Baseline condition.

Comparisons between the Korean guise and the other guises yielded small posterior medians and credible intervals that included zero. Although posterior probabilities for these contrasts are generally high in the positive direction, the CrIs show substantial uncertainty in the magnitude of these differences. A similar pattern is observed for contrasts including the Korean-American, Chinese, and White guises relative to the Baseline condition. All of these contrasts showed positive posterior medians and it suggests a tendency toward higher perceived accentedness when a visual cue and social information were provided than the Baseline. However, their credible intervals all crossed zero, indicating that these effects are weaker and less reliable.

Overall, no reliable evidence of race- and ethnicity-based bias was found in accentedness ratings across the White, Korean-American, and Chinese guises. However, the mismatched condition, the Korean guise, was perceived as more accented relative to the Baseline, which is consistent with an expectation mismatch effect in accentedness perception (Gnevsheva, 2018; McGowan, 2015).

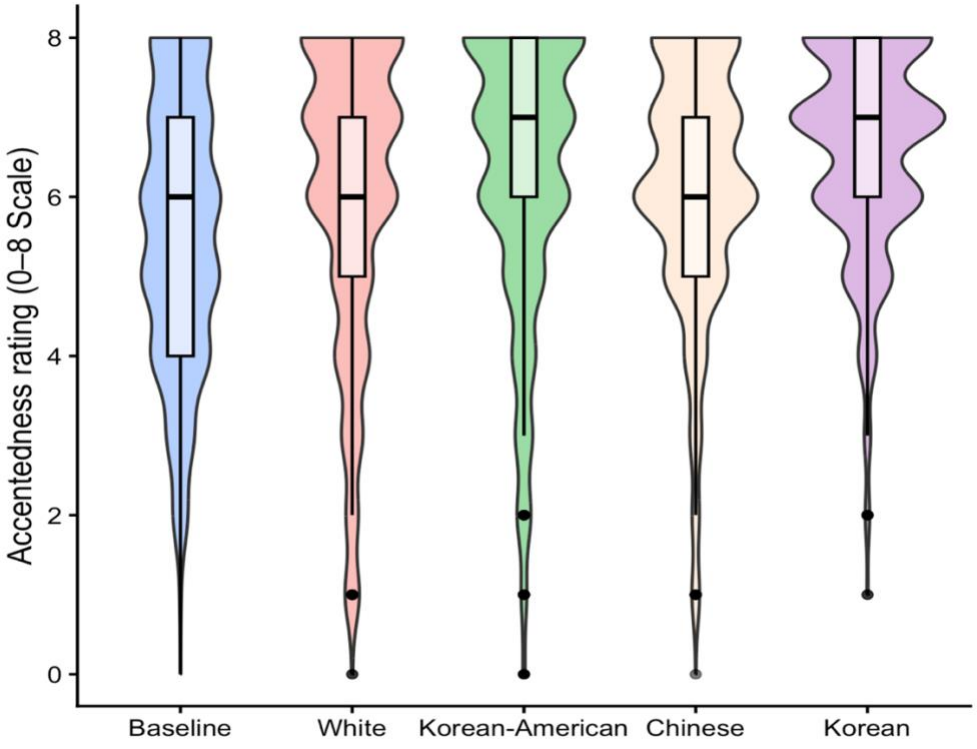
Table 2.7

Posterior contrasts in perceived accentedness

Contrast	Median (logit)	P(>0)	95%CrI
Korean – Baseline	2.02	0.99	[0.53, 3.50]
Korean – White	1.13	0.93	[-0.39, 2.65]
Korean – Chinese	0.88	0.87	[-0.68, 2.48]
Korean – Korean-American	0.63	0.80	[-0.85, 2.07]
Korean-American – Baseline	1.40	0.96	[-0.12, 2.88]
Korean-American – White	0.50	0.75	[-0.98, 2.00]
Korean-American – Chinese	0.26	0.63	[-1.29, 1.81]
Chinese – Baseline	1.14	0.92	[-0.40, 2.68]
Chinese – White	0.24	0.62	[-1.27, 1.79]
White - Baseline	0.89	0.88	[-0.60, 2.36]

Figure 2.4

Distribution of accentedness ratings by condition.



Note. Violin plots show the density of observed ratings, with boxplots indicating the descriptive median and interquartile range.

2.3.3.3. Correlation

2.3.3.3.1 Intelligibility and accentedness

Spearman correlation analyses were conducted to examine the relationship between intelligibility and accentedness judgments at the participant level. The results showed no significant correlation between whole sentence intelligibility and accentedness ratings ($\rho(123) = -0.07$, $p = .45$). Similarly, no significant correlation was observed between target word intelligibility and accentedness ratings ($\rho(123) = -0.12$, $p = .20$). These results suggest that Korean participants' ability to accurately understand foreign-accented Korean speech was not associated with their perceived accentedness ratings.

2.3.3.3.2 Demographic and language background effects

Spearman correlation analyses were conducted to examine the relationship between intelligibility and accentedness on one hand and demographic information on the other. For intelligibility, the number of participants' second languages showed a positive correlation with both whole sentence accuracy ($\rho(123) = .28$, $p = .001$) and target word accuracy ($\rho(123) = .29$, $p = .001$), and it indicates that participants who speak more second languages tended to be better at transcribing foreign-accented Korean. Participants who reported exposure to media featuring Korean-speaking foreigners tended to achieve higher transcription accuracy, but the effect was marginal for whole sentence accuracy ($\rho(123) = .17$, $p = .053$) and significant for target word accuracy ($\rho(123) = .25$, $p = .004$). People who reported that pronunciation or accent is important for learning a second language tended to show lower transcription accuracy ($\rho(123) = -0.18$, $p = .05$), although this effect was marginal. No other demographic or language background variables showed significant relationships with intelligibility.

For accentedness ratings, gender showed a significant effect ($\chi^2(1) = 6.46$, p = .011), with female participants giving higher accentedness ratings than male participants. The number of second languages showed a negative correlation with accentedness ratings ($\rho(123) = -0.28$, $p = .001$), and it suggests that participants who speak more second languages tended to perceive the speech as less accented. People who reported that pronunciation or accent is important for learning a second language tended to give higher accentedness ratings ($\rho(123) = .18$, $p < .05$). No other demographic or language background variables showed significant correlations with accentedness ratings.

2.4 Discussion

The present study examined whether visual and descriptive social information influences Korean listeners' perception of foreign-accented Korean. Building on the previous studies on Reverse Linguistic Stereotyping (RLS; Kang & Rubin, 2009; Rubin, 1992) and the exemplar model of speech perception (McGowan, 2015), this chapter tested predictions related to race- and ethnicity-based biases and expectation mismatch. Across tasks, a distinction between two tasks was observed that intelligibility was not largely affected by social information, whereas accentedness judgments were subtly affected by expectation mismatch.

At the whole sentence level, intelligibility did not systematically differ depending on the tested social factors. Korean listeners transcribed the speech with comparable accuracy across conditions. Posterior contrasts between conditions all included zero, indicating uncertainty regarding differences in transcription accuracy and no clear evidence that race, ethnicity or expectation mismatch influenced sentence-level intelligibility. Overall, these results suggest that Korean listeners' transcription accuracy for accented speech was not systematically influenced

by the speaker's perceived race or ethnicity. This finding is notable because accuracy was not at ceiling, leaving room for social information to affect performance. Nevertheless, the results did not provide consistent evidence that racial or ethnic information explained variation in listeners' transcription accuracy.

At the target word level, however, there was one reliable contrast in which intelligibility was higher for the White guise than for the Chinese guise. This result is partially consistent with race-based expectation. If the race-based bias dominated perception, the Korean-American guise would also be expected to yield lower intelligibility than the White and similar to the Chinese guise. However, the Korean-American guise did not differ reliably from either the White or the Chinese conditions, with its median estimate falling between the two. Thus, these results may suggest that race-based bias may emerge only in comparisons between ethnic out-group members.

One possible explanation for why the effect appeared only at the target word intelligibility, the target words occurred sentence-finally and were designed to be low in contextual predictability, which requires more careful attention for accurate transcription. Therefore, it is possible that subtle social factors appeared influential only under increased processing demands. In contrast, listeners rely more on contextual information across the sentence for sentence-level intelligibility, which could reduce the potential influence of social expectations. Nonetheless, because this pattern was limited to one single contrast, it should be interpreted cautiously.

Taken together, the results provide little evidence that race-based or ethnicity-based biases systematically influenced Korean listeners' perception of foreign-accented Korean. Importantly, the Korean-American guise did not show a consistent advantage in both sentence-

level and target word intelligibility. In other words, ethnic in-group advantage was not systematically observed in the intelligibility data, meaning Korean listeners do not necessarily have positive bias toward their in-group members in comparison to White American or Chinese speakers of accented Korean.

However, there is one limitation regarding the interpretation of the Chinese guise. In this Chapter, the Chinese condition was included to separate race-based effects from ethnic in-group effects of the Korean-American condition, and the norming study suggested that there was no mismatch effect, as Korean listeners did not identify the speaker as English-accented based on the audio alone. Nevertheless, this does not entirely rule out the possibility that some listeners may have drawn on their prior experience with Chinese-accented Korean when processing the Chinese guise because Chinese people make up one of the largest foreign resident groups in South Korea (Korea Immigration Service, 2024) and Chinese students are also a highly visible group in Korean universities.

A related limitation is that the Chinese condition might not be interpreted as a test of race and socioeconomic status. In the design of this chapter, the Chinese guise was included to represent an East Asian racial in-group but ethnic and national out-group condition associated with relatively high socioeconomic status. However, "Chinese" is a socially and politically complex category in the Korean context. Public attitudes toward Chinese in South Korea have become increasingly negative in recent years, and these attitudes are shaped by geopolitical tension, historical disputes, cultural conflicts, and online discourse (J. Ahn, 2023; Hwang, 2025; H. J. Jung & Jeong, 2016; E. E. Song, 2023). Therefore, the Chinese guise likely carried meanings beyond the factors explicitly manipulated in this chapter and the results should be interpreted cautiously as those factors were not isolated by the present experimental design.

On the other hand, accentedness judgments were not sensitive to race-based or ethnicity-based biases similar to intelligibility task, but were instead influenced by expectation mismatch. It is consistent with previous studies showing that cue mismatch outweigh any potential racial advantages in the accentedness rating task (Gnevsheva, 2018) and in target word intelligibility (McGowan, 2015). All guise conditions showed positive posterior medians relative to the Baseline, suggesting a general tendency for speech paired with visual cues to be perceived as more accented. However, these effects were uncertain in magnitude, as credible intervals crossed zero. However, the Korean guise, in which a Korean identity was paired with non-native speech, showed a reliable increase in perceived accentedness relative to the Baseline. This indicates that accent ratings were selectively increased when listeners' expectations about a native speaker were violated. By contrast, differences among the other guises (White, Chinese, and Korean-American) were small and uncertain.

These results suggest that the mere presence of visual information may influence subjective evaluations (i.e., accent rating), but a reliable increase in perceived accentedness emerged only when social information conflicted with listeners' expectations. Thus, mismatch between expected and actual speech appears to play a role in shaping accentedness judgments. One possible explanation is that having additional visual information while listening to speech may have increased cognitive load (Mattys & Palmer, 2015), which in turn may lead to more negative judgments in a subjective rating task. In contrast, the Baseline condition, which provided minimal visual information, may have avoided this effect. Yi et al. (2013) found that visual cues had little effect on the perception of native speech. However, this may be because processing native speech requires less cognitive effort compared to processing accented speech. In contrast, accented speech is known to impose greater processing demands than native speech

(McLaughlin et al., 2024). Thus, when social information conflicted with listeners' expectations, the additional cognitive burden associated with visual cues may have further amplified negative evaluations.

Returning to previous studies, although the present findings broadly align with those of Gnevsheva (2018) and McGowan (2015), the results do not fully align with McGowan's results. In McGowan (2015), expectation mismatch influenced target word intelligibility, whereas in the present chapter, the mismatch effect appeared only in accentedness judgments. This difference may reflect the distinct perceptual processes involved in the two tasks rather than contradicting McGowan's findings. The present findings suggest that intelligibility and accentedness judgments tap into different aspects of accented speech perception. Transcription may function as a relatively objective task that requires listeners to accurately comprehend linguistic content, whereas accentedness judgment is more about a subjective evaluation where accented speech is compared against listeners' internalized representations of native speech norms (i.e., exemplar).

The correlation analyses further supported this distinction between intelligibility and accentedness judgments. No significant correlation was observed between listeners' transcription accuracy and their accentedness ratings. In other words, participants who were better able to understand the foreign-accented Korean speech did not necessarily perceive it as less accented. This interpretation aligns with the study of Melguy and Johnson (2021), which found no significant effect of visual cues (e.g., silhouette, European face, East Asian face, and South Asian face) on transcription accuracy when listening to accented speech, and with Munro and Derwing (1995), who reported that accent ratings did not necessarily correlate with transcription accuracy. Together with the present findings, these results indicate that transcription accuracy and accent

judgment of foreign-accented speech may not always align or correlate with each other. Social expectations may therefore modulate subjective judgments without impairing intelligibility.

Additional correlation analyses also suggested that listener-related experience may influence the perception of foreign-accented speech. Participants who reported speaking more second languages tended to achieve higher transcription accuracy and to perceive the speech as less accented. This result may reflect the fact that multilingual listeners are better at understanding varied accents and more generous to accents. Exposure to media featuring Korean-speaking foreigners also showed a tendency to facilitate intelligibility. These findings suggest that listeners' experience with accented speech may play a role in comprehending and evaluation of accented speech, although direct contact with foreign friends speaking Korean did not have correlation effects.

The absence of systematic social bias effects in the present study does not imply that social biases does not affect intelligibility or speech perception in general. Rather, these findings suggest that the mechanisms underlying Reverse Linguistic Stereotyping (RLS) may not operate in the same way across tasks or cultural contexts (Graillot & Oh, 2025; Kutlu et al., 2022). In the Korean context, race-based distinctions between White and East Asian individuals may not function as strongly as they do in other settings where Whites constitute a majority group, which is where much of the RLS literature has been developed. Social categories such as race and ethnicity may carry different historical and socioeconomic meanings across sociocultural environments, potentially shaping how listeners perceive and evaluate speakers. Social bias effects observed in one sociolinguistic setting, therefore, may not directly generalize to another.

At the same time, the expectation mismatch effect observed in accentedness judgments is consistent with the exemplar model of speech perception. When listeners encounter speech

that deviates from stored expectations (i.e., “Korean” speaker producing foreign-accented Korean), the mismatch effect increases perceived accentedness. Because the credible intervals for several contrasts were relatively broad, this interpretation should remain cautious. Thus, the present findings should not be interpreted as disproving RLS. Instead, they suggest that different mechanisms may operate for different perceptual tasks, with mismatch playing a particularly salient role in subjective accent evaluation.

2.5 Conclusion

This study examined how listeners’ social perceptions of a speaker influence their perception of foreign-accented speech, focusing on intelligibility and accentedness judgments within the sociocultural context of South Korea. By testing predictions derived from Reverse Linguistic Stereotyping (RLS) and the exemplar model of speech perception, the study sought to determine whether social biases or expectation mismatch meaningfully shape speech perception.

Across experiment tasks, intelligibility proved to be largely robust to manipulation of social information. Visual and descriptive cues did not systematically alter transcription accuracy, and even when a subtle directional tendency emerged at the target word level, these effects did not generalize across conditions. This pattern suggests that actual comprehension of accented speech remains relatively robust, even when listeners are presented with conflicting social information about the speaker.

On the other hand, accentedness judgment showed selective sensitivity to expectation mismatch. When foreign-accented speech was paired with a native Korean identity, perceived accentedness reliably increased. These findings indicate that violations of social expectations

about native speech play a greater role in shaping accent evaluation than biases based on social categorization.

The distinction between intelligibility and accentedness is particularly important. Korean listeners may understand speech equally well regardless of the social information the speakers have while evaluating it differently depending on who they believe the speaker should be. In this sense, perceived accentedness is not a reliable indicator of actual intelligibility. Rather, it reflects an evaluative process that is more susceptible to expectations listeners hold. This finding reinforces previous research demonstrating that accent ratings do not necessarily correspond to intelligibility and suggests the importance of distinguishing between accurate understanding and social evaluation in studies of accented speech.

Importantly, the absence of penalties based on social bias in intelligibility in this study does not suggest that social bias is not influential in speech perception. Instead, the findings emphasize the role of sociocultural context. The social meaning of race and ethnicity varies across societies, and bias effects documented in English-speaking or Western European contexts may not generalize uniformly to South Korea. The present study therefore contributes to a broader understanding of how speech perception operates within different social ecologies.

The results further suggest that expectation-based mechanisms may be especially influential in subjective accent evaluation. When listeners encounter speech that deviates from stored sociophonetic representations, perceived accentedness increases even when intelligibility remains unimpaired. This distinction has important implications for how accented speakers are evaluated, as perceived accentedness does not necessarily reflect communicative competence or actual communication breakdown.

Some limitations should be acknowledged as well. First, the study used a between-subjects design, with 25 listeners assigned to each guise condition. Although this design made it possible to avoid repeated exposure to different guises of the same speaker, it also limited the strength of the evidence for guise effects because each listener encountered only one type of social information. As a result, observed differences across conditions may have been difficult to detect subtle effects. Future research should replicate the present study with a larger sample and adopt a within-listener manipulation of guise. A within-listener design would provide stronger evidence for the effect of social information by allowing the same listeners' responses to be compared across different guise conditions.

Additionally, future research also can explore how cognitive processing costs contribute to perception of foreign accent, since it remains unclear how much additional cognitive effort is required to process accented speech when expectations are violated. By using different methods such as EEG or pupillometry, researchers could quantify processing costs in real-time speech perception tasks, shedding light on the cognitive mechanisms underlying accented speech perception.

Beyond these cognitive and perceptual aspects, it is also crucial to consider real-world applications such as in education and workplace environments in multilingual societies. Listener-related factors can shape perceptions of accented speech, and perceived accentedness does not necessarily correspond to communicative effectiveness. Recognizing this distinction may inform second language curricula, educational policies, and hiring practices, promoting more equitable and informed evaluation of accented speakers.

III. NATIVE KOREAN LISTENERS' LANGUAGE ATTITUDES TOWARD ACCENTED KOREAN: EVALUATIONS OF CLOSE OUTSIDERS AND DISTANT INSIDERS

3.1 Introduction

The preceding chapter examined how social information associated with speakers shapes Korean listeners' perception of foreign-accented speech through intelligibility and accentedness judgment tasks. A race-based effect emerged only in target word intelligibility, and its influence was observed only between ethnic out-group speakers (i.e., the White and Chinese guises), while the effect was not found in sentence-level intelligibility or other contrasts. Similarly, the accentedness judgment task showed a mismatch effect rather than the effects of social bias or stereotype. That is, these perceptual measures provided limited insight into the underlying social beliefs that Korean listeners hold toward speakers from different social backgrounds, and how much such beliefs might relate to the results of perception tasks.

To address this gap, the present chapter employs language attitude questionnaires as a complementary measure of listener bias. Studies adopting the bias account of accented speech perception have used various instruments such as the homophily instrument (Rubin, 1992), the Implicit Association Test (Yi et al., 2013), and language attitude questionnaire (O. Kang & Rubin, 2009), to examine how listener bias or attitudes correlate with perceptual outcomes. Among these, language attitude questionnaires offer a structured means of capturing multidimensional social evaluations, and have been widely used in matched guise research (Cargile & Giles, 1997; Díaz-Campos & Killam, 2012; Loureiro-Rodriguez et al., 2013) to assess how speaker information shapes listeners' beliefs about language and its users.

Research on language attitudes has consistently demonstrated that people hold systematic beliefs about language and make judgments about others based on linguistic variation in everyday interaction (Dragojevic & Goatley-Soan, 2022; Garrett, 2010; Giles & Watson, 2013). However, these evaluations are not limited to linguistic cues themselves. Manipulated or inferred information about speakers, such as speakers' appearance can also shape listeners' beliefs and evaluations, eliciting distinct language attitudes (Campbell-Kibler, 2007; Purnell et al., 1999; Rubin, 1992), as demonstrated by studies using the matched guise technique (Lambert et al., 1960). Accordingly, this chapter adopts the framework proposed by Dragojevic and Giles (2016), in which language attitudes consist of five dimensions: positive affect, negative affect, status, solidarity, and comprehensibility.

As discussed in Chapter 1, the status and solidarity dimensions have been foundational in language attitude research. Status-related dimensions include attributes such as intelligence, education, success, and wealth, which are closely associated with perceived socioeconomic status (Fiske et al., 2007; Fuertes et al., 2012). Accordingly, high-prestige varieties or speakers from socioeconomically high status are typically rated higher on status dimensions (Coupland & Bishop, 2007; Fuertes et al., 2012). In contrast, solidarity-related dimensions includes interpersonal judgments such as warmth, friendliness, and trustworthiness (Dragojevic et al., 2013). These judgments are often rated higher for low-prestige varieties, particularly among in-group members, because solidarity indexes in-group loyalty.

In addition, the comprehensibility dimension, as included in Dragojevic and Giles's (2016) language attitude framework as well as research by Munro and colleagues (e.g., Munro & Derwing, 1995), reflects listeners' subjective assessment of how easy or difficult it is to understand the speaker. This dimension is particularly relevant in the context of foreign-accented

speech, where perceived ease of understanding can be shaped by speakers' social information rather than the speech signal alone (O. Kang & Rubin, 2009; Rubin, 1992).

Most of this work, however, has been conducted in the United States, Canada, the United Kingdom, and other Western European contexts, leaving open the question of whether such evaluative patterns appear differently in other sociolinguistic settings. In English-speaking contexts such as the U.S. and the U.K., for example, socioeconomic status and intergroup preference are often conflated, as non-accented mainstream English speakers (i.e., White race) tend to be both in-group members and socioeconomically dominant, while accented speakers are often perceived as out-group members from lower socioeconomic backgrounds. Examining accented speech perception in a non-English-speaking context, therefore, provides an opportunity to tease apart these two potentially confounding factors, as foreign-accented speakers with a different race may be associated with higher socioeconomic status than members of the host country.

As discussed in Chapter 1, South Korea provides a useful context for examining language attitudes toward foreign-accented speech because race, ethnicity, socioeconomic perceptions, and expectations about Korean language ability are not aligned in the same way as in many English-speaking contexts. In particular, racial out-group status does not necessarily imply low socioeconomic status, and ethnic in-group membership may coexist with foreign-accented Korean speech. These dynamics make Korean an informative context for examining how social information shapes listeners' evaluations of foreign-accented speech. Despite this, research on native Korean listeners' language attitudes toward foreign-accented Korean remains limited.

One study (H. Lee, 2016a) investigated general language attitudes toward foreign-accented Korean speech in terms of foreignness, language preference, dynamism, credibility, and

social status. However, this study relied solely on survey-based judgments without audio stimuli or attention to specific accent types. The findings suggested that native Koreans tended to consider foreign-accented speakers as less credible and less trustworthy and to attribute lower social status to them, even if their speech was intelligible and their pronunciation was clear. However, given that the experiment simply tapped into their beliefs rather than perceptual responses, this suggests the need for further examination of how actual speech and specific social information shape listeners' evaluations.

This chapter addresses these issues by employing the matched guise technique, in which the same foreign-accented Korean speech is paired with different images and descriptions representing distinct speaker identities. Five conditions are tested as in Chapter 2: Baseline, White American, Korean-American, Chinese, and Korean. By varying race and ethnicity while presenting the same speech, this design allows for examining how social information influences language attitudes toward accented Korean. In doing so, this study extends the status-solidarity framework beyond Western contexts and contributes new insights into how language attitudes are mediated by local sociolinguistic ideologies.

3.1.1 Summary of Research Questions

(1) **Research Question 1**: Does race-based speaker information influence language attitudes toward foreign-accented Korean speech?

If race-based social perceptions influence language attitudes, the White guise is expected to receive more favorable evaluations than the Korean-American and Chinese guises, particularly on status dimensions, as White appearance is often associated with higher prestige and socioeconomic status.

Alternatively, because the status dimension reflects perceived socioeconomic status, and the White, Korean-American, and Chinese guises may all be associated with relatively high socioeconomic status based on their nationalities (i.e., the U.S. and China), it is also possible that no clear differences would emerge among these guises (see Table 3.1)

(2) **Research Question 2**: Does ethnic group membership influence language attitudes toward foreign-accented Korean speech?

If ethnic group membership affects language attitudes, the Korean-American guise is expected to receive more favorable ratings than the White and Chinese guises, particularly on solidarity dimensions, as solidarity dimensions often reflect in-group loyalty.

(3) **Research Question 3**: Does language expectation influence language attitudes toward foreign-accented Korean speech?

If expectation mismatch influences language attitudes, the Korean guise is expected to receive less favorable evaluations than the other guises due to increased processing cost. However, because the speaker is ethnically Korean, similar to the Korean-American guise, the Korean guise may still receive relatively higher ratings on solidarity dimensions.

(4) **Research Question 4**: Does perceived comprehensibility covary with status and solidarity evaluations of foreign-accented Korean?

(5) **Research Question 5**: Are listeners' evaluations of accented Korean correlated with intelligibility, accentedness, and individual differences in demographic and language

background (e.g., age, gender, education, L2 experience, and exposure to Korean-speaking foreigners)?

Table 3.1
Overview of experimental conditions and associated social expectations

	Race (Image)	Ethnic relation	Socioeconomic status	Expectation
A. Baseline	Silhouette	Unspecified	Unspecified	Unspecified
B. White American	White	Out-group	High	Match
C. Korean-American	East Asian	In-group	High	Match
D. Chinese	East Asian	Out-group	High	Match
E. Korean	East Asian	In-group		Mismatch

Note. This table is reproduced again for ease of reference. *Ethnic relation* refers to the perceived ethnic in-group or out-group relationship between the speaker and Korean listeners. *Expectation* indicates whether listeners' likely social expectations about the speaker's accent match the actual accent they hear.

3.2 Methodology

The language attitude survey followed intelligibility and accentedness rating tasks reported in Chapter 2. In other words, the same 125 participants completed the intelligibility task, followed by the accentedness rating task, and then followed by the language attitude survey within a single assigned condition. The participants therefore responded to the same audio, written, and visual information as described in Chapter 2. The five conditions and the images used in the experiment are provided again here for the ease of reference.

Five conditions were used, using the images shown in Figure 3.1:

1. A silhouette labeled “Foreign learner of Korean” (baseline)
2. A photograph of a White female labeled “American learner of Korean”
3. A photograph of an Asian female labeled “Korean-American”
4. A photograph of an Asian female labeled “Chinese learner of Korean”
5. A photograph of an Asian female labeled “Korean”

Figure 3.1
Faces used in the experiment. From left to right: silhouette, White, and Asian face.



3.2.1 Language attitude survey

Language attitude is an abstract construct that cannot be directly observed but can be inferred from evaluative responses such as questionnaire items or rating scales (Ryan & Giles, 1982). In this study, listeners’ language attitudes toward the foreign-accented speaker were measured using three sets of items: 1) affective responses toward the speaker, 2) status and solidarity dimensions, and 3) perceived comprehensibility of the speaker’s speech. All responses were provided on a 7-point Likert scale (1 = not at all, 7 = extremely).

First, participants rated how they felt about the foreign-accented speaker on a total of 6 affects, including three positive (i.e., interested, proud, glad) and three negative (i.e., frustrated,

annoyed, irritated) affects. These items were adapted from Dragojevic and Giles (2016), which in turn drew on the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988). See Appendix C for the survey items. Then, participants rated the speaker on 10 personality traits: five status-related traits (intelligent, educated, smart, competent, proficient) and five solidarity-related traits (friendly, trustworthy, social, nice, warm), using a 7-point scale (1 = not at all, 7 = extremely). These items were adapted from previous research on language attitudes (Dragojevic & Giles, 2016; Heaton & Nygaard, 2011).

The survey items were translated into Korean and several adjustments were made to ensure semantic appropriateness. For example, "happy" appears in the survey (see #3 in Appendix C). The direct translation 행복하다 ('happy') typically denotes a long-term state of life satisfaction in Korean, so it was replaced with 기쁘다 ('glad'), which more accurately reflects an immediate emotional reaction denoted in the item. Likewise, 'enthusiastic' was rendered as 대견하다 ('proud of someone'), as this better reflects the listener's evaluative stance rather than the speaker's character. Finally, since both 'annoyed' and 'irritated' are usually translated as a single Korean word, 짜증나다, an additional item, 귀찮다 ('bothered/annoyed') was included to capture nuance (see Appendix C).

Next, participants evaluated the comprehensibility of the speaker on the same 7-point scale, indicating the extent to which it was easy to understand, clearly articulated, natural and smooth to hear, and effortful to understand.

3.2.2 Procedure

As described in Chapter 2, each participant was assigned to one of the five experiment conditions, with 25 participants in each condition (Baseline, White American, Korean-American,

Chinese, and Korean), and completed the experiments online through the Gorilla Experiment Builder platform (https://gorilla.sc). The experiment consisted of four sequential components: an intelligibility task, an accentedness rating task, a language attitude survey, and a questionnaire on participants' demographic and language backgrounds. The entire study took approximately 15 to 20 minutes in total to complete, and sessions were automatically terminated if left inactive for more than one hour. In this chapter, I report the analysis and results of the attitude survey.

3.3 Analysis

3.3.1 Language attitude data

Prior to analyzing the language attitude ratings, internal consistency reliability of the positive affect, negative affect, status, solidarity, and comprehensibility scales was assessed using Cronbach's alpha (Cronbach, 1951), calculated using the `alpha` function in the *psych* package (Revelle, 2026) in R (R Core Team, 2026). This analysis evaluated whether items within each attitudinal dimension behaved consistently as a group, allowing the item responses to be summed into one composite rating. The item "effortful to understand" in comprehensibility was reverse-coded so that higher scores reflected easier understanding. Following conventional guidelines, values of alpha between .70 and .95 were considered acceptable (Tavakol & Dennick, 2011).

Responses within each attitudinal dimension were summed to create composite scores for positive affect, negative affect, status, solidarity, and comprehensibility. These composite scores (*Rating*) were the dependent variable.

To examine the effect of social information associated with the speaker (*Guise* condition) across attitudinal dimensions, Bayesian multilevel regression models were fit in R (R Core

Team, 2026) using the `brm` function in the *brms* package (Bürkner, 2017). The model included *Guise* condition, attitudinal dimension (*Attitude*), and their interaction as fixed effects. Random intercepts for *Participant* were included, and random slopes for *Attitude* by-*Participant* were specified.

Although responses were collected using Likert scales, the dependent variable *Rating* consisted of composite scores and was therefore treated as continuous. Models were fit using a Student-t likelihood, which allows heavier-tailed error distributions than a Gaussian model and provides more stable estimation when participants differ in their use of the rating scale such as some consistently giving more extreme ratings. Four Markov chains with 8,000 iterations per chain were used to sample from the posterior distribution. Model convergence was checked by inspecting potential scale reduction factors ($\hat{R}$), effective sample sizes (ESS), and trace plots of the Markov chain Monte Carlo (MCMC) samples. All parameters met standard convergence criteria ($\hat{R} = 1.00$), with sufficient effective sample sizes (> 1000) and well-mixed chains.

3.3.2 Correlation

To further investigate whether comprehensibility covaried with the status and solidarity dimensions, whether the perceptual measures from Chapter 2 (i.e., intelligibility and accentedness) were associated with social evaluations, and whether social evaluations are related to participants' demographic and language background, several correlation analyses were conducted. Because responses were collected across multiple trials and items, all analyses were aggregated at the participant level. For each participant, intelligibility, accentedness ratings, and attitude scores were averaged within condition prior to analysis.

For the relationship between intelligibility and language attitudes, whole sentence intelligibility and target word intelligibility were analyzed with composite ratings of positive affect, negative affect, status, solidarity, and comprehensibility. Because language attitude scores consisted of summed Likert scale responses and intelligibility measures were proportions of correctly transcribed words, Spearman's rank correlation coefficients (ρ) were calculated.

To examine whether accentedness judgments were related to social evaluations, accent ratings were analyzed with the same composite attitude measures. Accent ratings were averaged for each participant within condition and correlated with the attitude scores using Spearman's rank correlations.

Relationships between language attitudes and listeners' language and demographic background were also investigated. For continuous or ordinal variables (e.g., age, age of acquisition, frequency of media exposure), Spearman correlations were conducted. For binary variables (e.g., living Abroad experience, having Korean-speaking foreign friends), Pearson (point-biserial) correlations were used. For categorical variables (e.g., gender, dialect type), Kruskal-Wallis tests were conducted.

3.4 Results

3.4.1 Language attitude

Prior to reporting results for each attitudinal dimension, internal consistency reliability using Cronbach's alpha was assessed. The three positive affect items (α = .44) and three negative affect items (α = .57) showed low internal consistency. Accordingly, both individual item scores and composite scores are reported below, but the composite scores should be interpreted with caution due to their low internal consistency reliability. By contrast, the other scales showed

acceptable internal consistency reliability: status (α = .91), solidarity (α = .84), and comprehensibility (α = .79).

Posterior estimates from the Bayesian models across the five guise conditions are summarized in Table 3.2 and visualized in Figure 3.2. The table provides an overview of overall patterns across conditions, and pairwise contrasts of the posterior estimates are reported separately for each attitudinal dimension below.

3.4.1.1 Positive affect

1) RQ1: Race bias (White advantage) vs. RQ2: Ethnicity – In-group bias (Korean-American advantage)

If race-based bias influences language attitudes, the White guise would be expected to receive higher positive affect ratings than the Chinese and Korean-American guises. On the other hand, if ethnic in-group bias dominates, the Korean-American guise would be expected to receive higher ratings than the other guises.

Although the White guise showed numerically higher ratings than the Baseline, and the Korean-American and Chinese guises showed lower ratings (see Figure 3.2), the credible intervals for these contrasts included zero (see Table 3.3), indicating no reliable differences across conditions. Overall, these results provide no evidence of race- or ethnicity-based bias in positive affect ratings.

Table 3.2
Posterior median ratings and 95% HPD intervals for each attitudinal dimension across guise conditions

	Baseline	White	KAmerican	Chinese	Korean
Positive	5.58	5.62	5.36	5.43	4.94
	[5.13, 6.00]	[5.18, 6.07]	[4.92, 5.84]	[4.95, 5.89]	[4.46, 5.38]
Negative	2.23	2.43	3.12	3.20	3.21
	[1.64, 2.79]	[1.79, 3.00]	[2.53, 3.74]	[2.59, 3.81]	[2.64, 3.80]
Status	4.45	4.25	3.58	3.64	3.86
	[3.90, 5.07]	[3.67, 4.85]	[3.00, 4.18]	[3.04, 4.22]	[3.27, 4.45]
Solidarity	4.48	4.77	3.94	4.31	4.31
	[4.01, 4.97]	[4.28, 5.25]	[3.46, 4.42]	[3.85, 4.80]	[3.83, 4.79]
Comprehensibility	3.33	3.21	2.61	2.63	3.19
	[2.93, 3.75]	[2.79, 3.63]	[2.19, 3.03]	[2.21, 3.04]	[2.77, 3.60]

Note. KAmerican = Korean-American; Higher values indicate more positive evaluations. For negative affect, higher values indicate more negative evaluations.

2) RQ3: Expectation mismatch

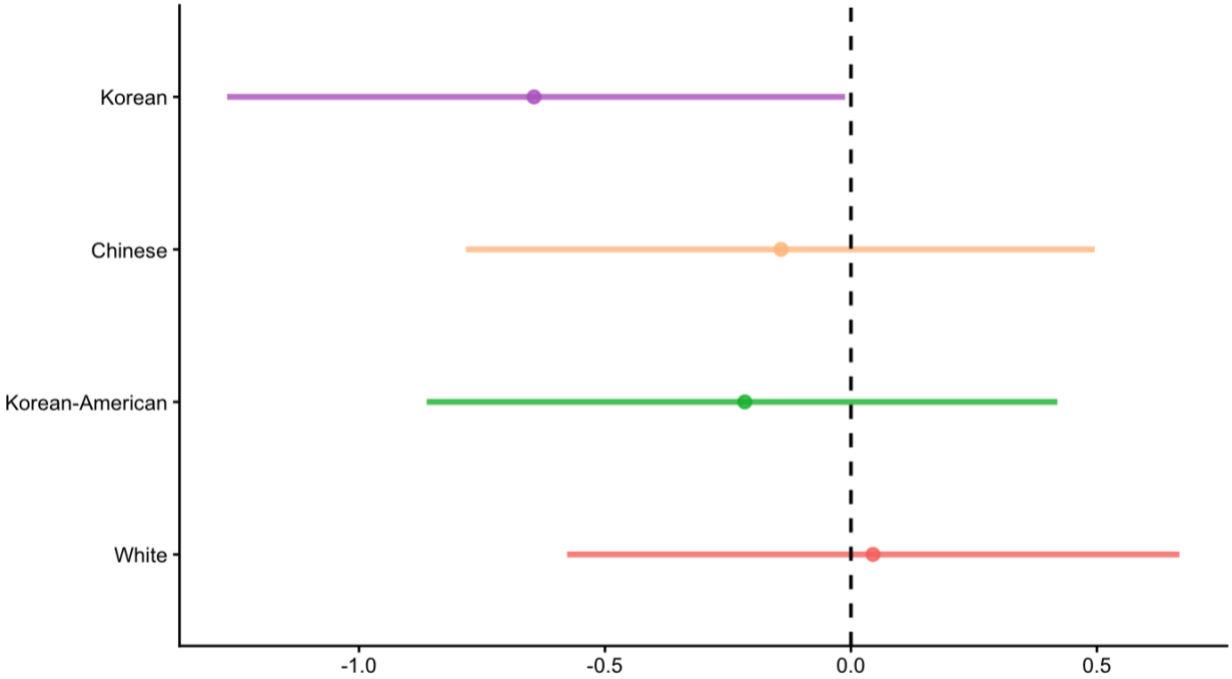
If expectation mismatch influences language attitudes, the Korean guise would be expected to receive lower positive affect ratings than the other guises.

The Baseline (median = 0.64, 95% CI [0.01, 1.27], P(>0) = .98) and the White guise (median = 0.68, 95% CrI [0.05, 1.33], P(>0) = .98) showed higher positive affect ratings than the Korean guise. However, comparisons between the Korean guise and the Korean-American and Chinese guises were not reliable as the credible intervals included zero. The mismatch effect was found in positive affect, as the Korean guise received lower ratings than the Baseline and White guises, although comparisons with the Korean-American and Chinese guises were not reliable. However, the positive affect ratings showed lower internal reliability, and the results should be interpreted cautiously.

Table 3.3
Posterior contrasts for positive affect (logit scale)

Contrast	Median (logit)	P(>0)	95%CrI
Baseline – White	-0.05	0.44	[-0.67, 0.58]
Baseline – Korean-American	0.21	0.74	[-0.43, 0.86]
Baseline – Chinese	0.14	0.67	[-0.49, 0.79]
Baseline – Korean	0.64	0.98	[0.01, 1.27]
White – Korean-American	0.26	0.79	[-0.39, 0.89]
White – Chinese	0.19	0.72	[-0.46, 0.84]
White – Korean	0.68	0.98	[0.05, 1.33]
Korean-American - Chinese	0.07	0.58	[-0.59, 0.73]
Korean-American - Korean	0.43	0.90	[-0.22, 1.08]
Chinese – Korean	0.49	0.93	[-0.16, 1.15]

Figure 3.2
Estimated differences in positive affect ratings from the Baseline



Note. Points indicate posterior medians and horizontal bars represent 95% credible intervals.

3.4.1.2 Negative affect

Negative affect ratings showed a slightly different pattern from positive affect. Unlike the other four dimensions, higher ratings indicate less favorable evaluations in negative affect. Although the contrasts are reported in one direction for consistency with other subsections (see Table 3.4), they are described in the reverse direction in this section for ease of interpretation.

1) RQ1: Race bias (White advantage) vs. RQ2: Ethnicity – In-group bias (Korean-American advantage)

If race-based bias affects language attitudes, the White guise would be expected to receive lower negative affect ratings than the other guises.

Compared to the Baseline, the Korean-American (median = 0.89, 95% CrI [0.07, 1.71], P(>0) = .98) and Chinese (median = 0.97, 95% CrI [0.13, 1.83], P(>0) = .99) guises received higher negative affect ratings.

Additionally, compared to the White guise, the Chinese and Korean-American guises tended to receive higher negative affect ratings. Although the credible intervals for these contrasts slightly overlapped zero (Chinese: median = 0.77, 95% CrI [-0.09, 1.61], P(>0) = .96; Korean-American: median = 0.69, 95% CrI [-0.09, 1.61], P(>0) = .94), the posterior probabilities indicate strong directional evidence that the Chinese and Korean-American guises were evaluated more negatively than the White guise.

2) RQ3: Expectation mismatch

If expectation mismatch influences language attitudes, the Korean guise would be expected to receive higher negative affect ratings than the other guises.

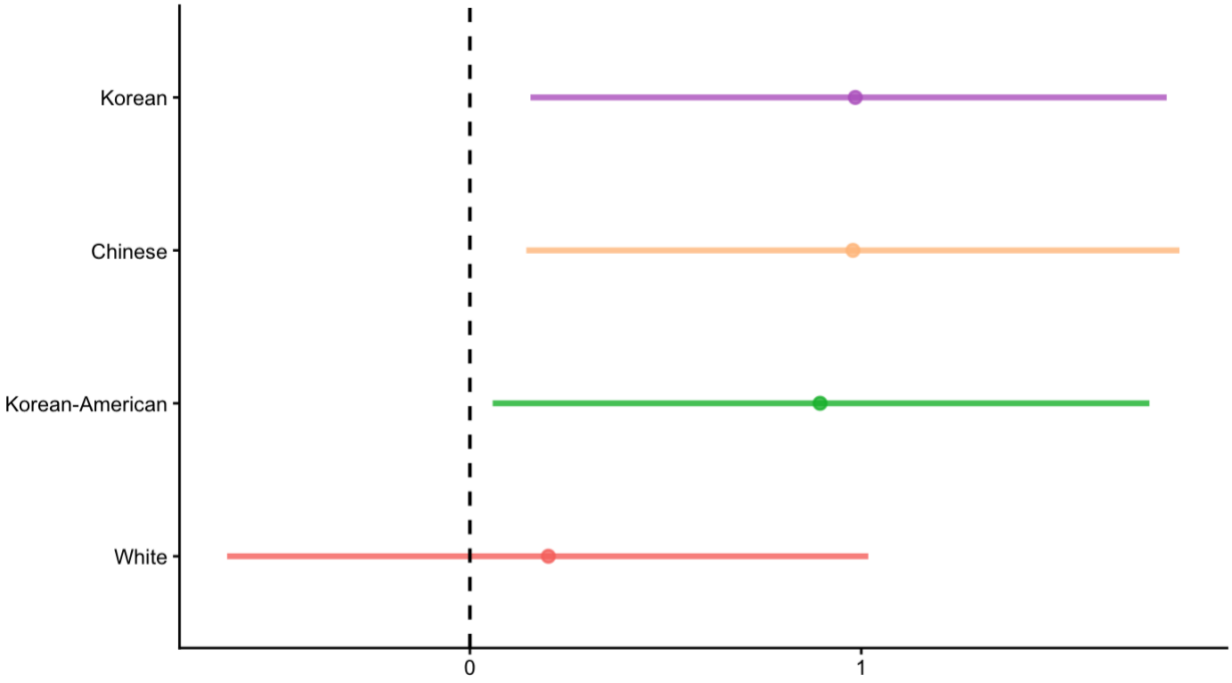
The Korean guise showed higher negative affect ratings than the Baseline (median = 0.98, 95% CrI [0.16, 1.81], P(>0) = .99) and the White guise (median = 0.78, 95% CrI [-0.04, 1.61], P(>0) = .97), although the CrI for the contrast with the White guise slightly included zero.

The expectation mismatch effect was found in negative affect, as the Korean guise received higher negative affect ratings than the Baseline and White guises, although comparisons with the Korean-American and Chinese guises were not reliable. However, the negative affect ratings showed low internal reliability, and these results should be interpreted with caution.

Table 3.4
Posterior contrasts for negative affect (logit scale)

Contrast	Median (logit)	P(>0)	P(<0)	95%CrI
Baseline – White	-0.20	0.31	0.68	[-1.02, 0.62]
Baseline – Korean-American	-0.89	0.02	0.98	[-1.71, -0.07]
Baseline – Chinese	-0.97	0.01	0.99	[-1.83, -0.13]
Baseline – Korean	-0.98	0.01	0.99	[-1.81, -0.16]
White – Korean-American	-0.69	0.06	0.94	[-1.55, 0.17]
White – Chinese	-0.77	0.04	0.96	[-1.61, 0.09]
White – Korean	-0.78	0.03	0.97	[-1.61, 0.04]
Korean-American - Chinese	0.08	0.42	0.57	[-0.78, 0.93]
Korean-American - Korean	-0.09	0.41	0.59	[-0.94, 0.77]
Chinese – Korean	-0.01	0.49	0.51	[-0.87, 0.84]

Figure 3.3
Estimated differences in negative affect ratings from the Baseline



3.4.1.3 Status

1) RQ1: Race bias (White advantage) vs. RQ2: Ethnicity – In-group bias (Korean-American advantage)

If race-based bias influences language attitudes in the status dimension, the White guise would be expected to receive higher status ratings than the Chinese and Korean-American guises. Conversely, if ethnic in-group bias dominates, the Korean-American guise would be expected to receive higher ratings than the other guises.

The Baseline received higher status ratings than the Korean-American guise (median = 0.87, 95% CrI [0.05, 1.70], P(>0) = .98). The Baseline also tended to receive higher status ratings than the Chinese guise (median = 0.81, 95% CrI [-0.04, 1.65], P(>0) = .97), although the credible interval included zero. No reliable differences were observed among the other contrasts.

These results are consistent with the alternative possibility outlined in Section 3.1.1, in that no clear differences were observed among the White, Korean-American, and Chinese guises. However, the finding that the Korean-American guise received lower status ratings than the Baseline may suggest that Korean-American speakers are evaluated less favorably in terms of status than speakers presented without explicit social information.

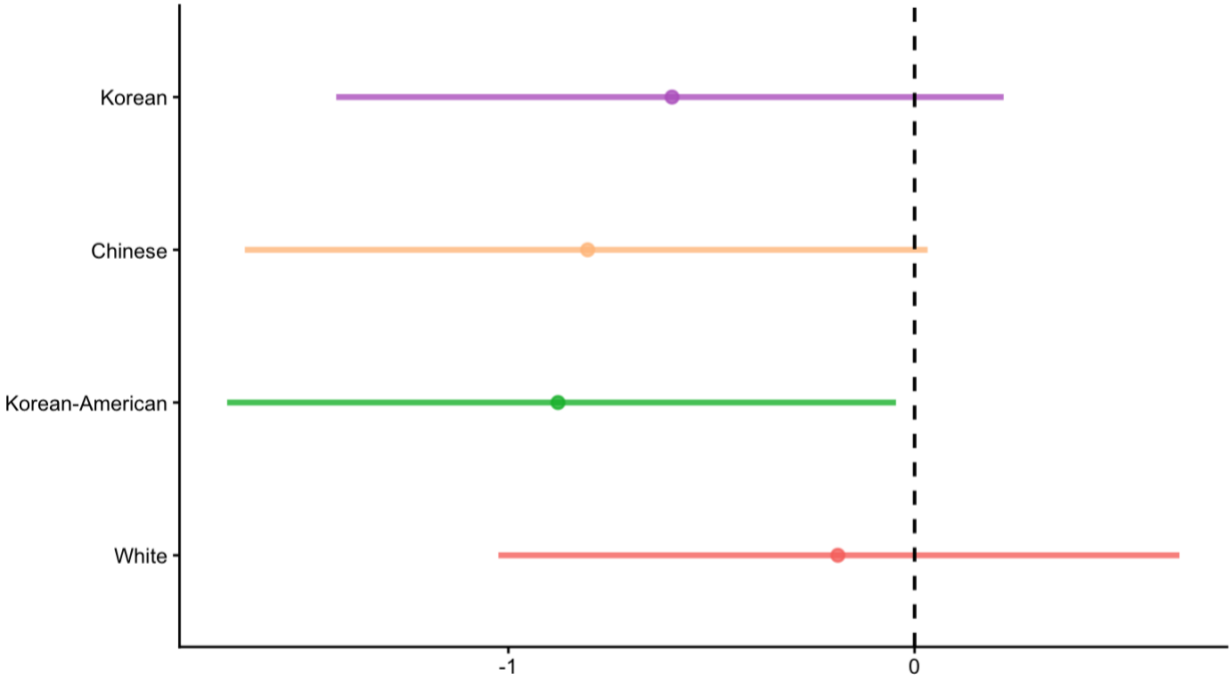
2) RQ3: Expectation mismatch

If expectation mismatch influences status ratings, the Korean guise would be expected to receive lower ratings than other guises. However, the Korean guise did not receive reliably lower ratings than the other guises, indicating that there is no evidence for the mismatch effect in status ratings.

Table 3.5
Posterior contrasts for status dimension (logit scale)

Contrast	Median (logit)	P(>0)	95%CrI
Baseline – White	0.20	0.68	[-0.63, 1.03]
Baseline – Korean-American	0.87	0.98	[0.05, 1.70]
Baseline – Chinese	0.81	0.97	[-0.04, 1.65]
Baseline – Korean	0.60	0.92	[-0.22, 1.43]
White – Korean-American	0.67	0.94	[-0.16, 1.51]
White – Chinese	0.61	0.93	[-0.23, 1.45]
White – Korean	0.40	0.83	[-0.43, 1.24]
Korean-American - Chinese	0.07	0.56	[-0.75, 0.90]
Korean-American - Korean	-0.28	0.26	[-1.11, 0.56]
Chinese – Korean	-0.21	0.31	[-1.04, 0.62]

Figure 3.4
Estimated differences in status ratings from the Baseline



3.4.1.4 Solidarity

3) RQ1: Race bias (White advantage) vs. RQ2: Ethnicity – In-group bias (Korean-American advantage)

If race-based bias influences language attitudes in the solidarity dimension, the White guise would be expected to receive higher solidarity ratings than the Chinese and Korean-American guises. Conversely, if ethnic in-group bias dominates, the Korean-American guise would be expected to receive higher ratings than the other guises.

The White guise received higher solidarity ratings than the Korean-American guise (median = 0.82, 95% CrI [0.12, 1.51], P(>0) = .99). All other contrasts were not reliable, as the credible intervals included zero. These results provide no evidence for ethnic in-group preference in the solidarity dimension. Instead, the observed difference between the White and Korean-

American guises suggests a more consistent pattern for race-based bias between the two American guises.

4) RQ3: Expectation mismatch

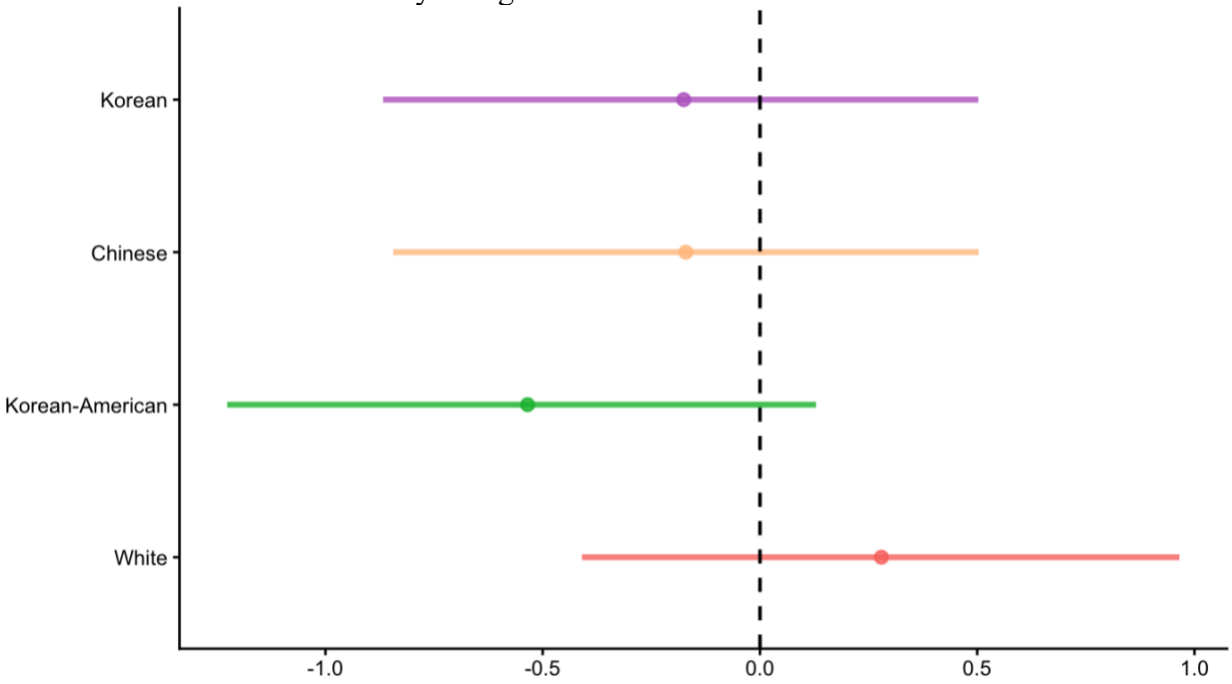
If expectation mismatch influences solidarity ratings, the Korean guise would be expected to receive lower solidarity ratings than the other guises. Alternatively, if ethnic in-group bias remains influential, the Korean guise may still receive relatively higher ratings due to shared ethnicity.

However, the Korean guise did not receive reliably lower or higher ratings than the other guises, as the credible intervals for these contrasts included zero. These results do not provide any evidence for either an expectation mismatch effect or ethnic in-group bias in the solidarity dimension, as outlined in Section 3.1.1.

Table 3.6
Posterior contrasts for solidarity dimension (logit scale)

Contrast	Median (logit)	P(>0)	95%CrI
Baseline – White	-0.28	0.21	[-0.99, 0.40]
Baseline – Korean-American	0.54	0.94	[-0.14, 1.22]
Baseline – Chinese	0.17	0.69	[-0.51, 0.84]
Baseline – Korean	0.17	0.69	[-0.51, 0.85]
White – Korean-American	0.82	0.99	[0.12, 1.51]
White – Chinese	0.45	0.90	[-0.24, 1.12]
White – Korean	0.45	0.90	[-0.23, 1.14]
Korean-American - Chinese	0.37	0.86	[-0.31, 1.06]
Korean-American - Korean	-0.37	0.14	[-1.05, 0.32]
Chinese – Korean	0.01	0.50	[-0.69, 0.69]

Figure 3.5
Estimated differences in solidarity ratings from the Baseline



3.4.1.5 Comprehensibility

5) RQ1: Race bias (White advantage) vs. RQ2: Ethnicity – In-group bias (Korean-American advantage)

If race-based bias influences perceived comprehensibility, the White guise would be expected to receive higher comprehensibility ratings than the Chinese and Korean-American guises. On the other hand, if ethnic in-group bias affects comprehensibility, the Korean-American guise would be expected to receive higher ratings than the other guises.

The Baseline was rated as more comprehensible than both the Korean-American (median = 0.73, 95% CrI [0.12, 1.31], P(>0) = .99) and the Chinese guises (median = 0.70, 95% CrI [0.12, 1.29], P(>0) = .99). Additionally, a similar directional pattern was observed for the White guise. The White guise tended to receive higher comprehensibility ratings than both the

Korean-American (median = 0.60, 95% CrI [-0.001, 1.21], P(>0) = .97) and Chinese guises (median = 0.58, 95% CrI [-0.02, 1.17], P(>0) = .97), although the credible intervals slightly included zero.

Overall, these results show that the Baseline was rated as more comprehensible than the Korean-American and Chinese guises, while there is no evidence for ethnic in-group preference and only a strong directional tendency favoring the White guise than the Korean-American and Chinese guises.

6) RQ3: Expectation mismatch

If expectation mismatch influences comprehensibility ratings, the Korean guise would be expected to receive lower comprehensibility ratings than the other guises.

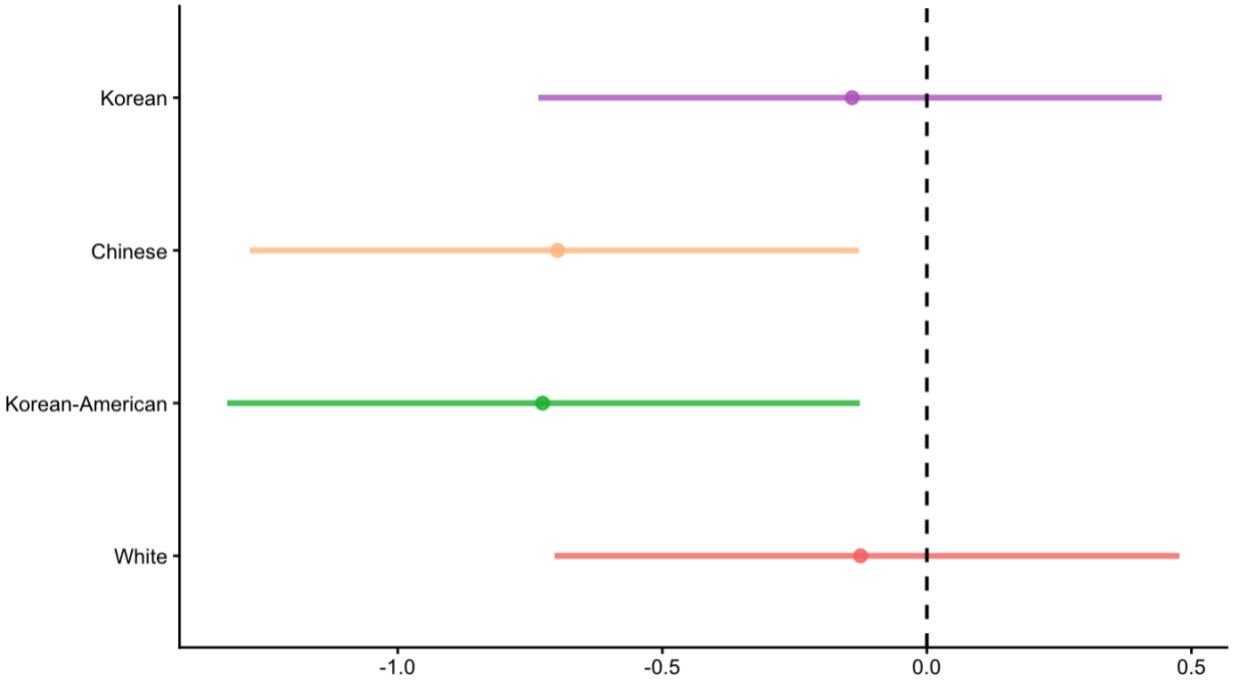
However, the Korean guise was rated as more comprehensible than both the Korean-American (median = 0.58, 95% CrI [-0.01, 1.18], P(>0) = .97) and Chinese guises (median = 0.56, 95% CrI [-0.04, 1.15], P(>0) = .97), although the credible intervals for these estimates slightly included zero.

These results show no evidence for the mismatch effect in comprehensibility ratings and are not consistent with mismatch prediction.

Table 3.7
Posterior contrasts for comprehensibility (logit scale)

Contrast	Median (logit)	P(>0)	95%CrI
Baseline – White	0.12	0.66	[-0.47, 0.70]
Baseline – Korean-American	0.73	0.99	[0.12, 1.31]
Baseline – Chinese	0.70	0.99	[0.12, 1.29]
Baseline – Korean	0.14	0.69	[-0.45, 0.73]
White – Korean-American	0.60	0.97	[-0.001, 1.21]
White – Chinese	0.58	0.97	[-0.02, 1.17]
White – Korean	0.02	0.53	[-0.57, 0.60]
Korean-American - Chinese	0.02	0.47	[-0.57, 0.62]
Korean-American - Korean	-0.58	0.03	[-1.18, 0.01]
Chinese – Korean	-0.56	0.03	[-1.15, 0.04]

Figure 3.6
Estimated differences in comprehensibility ratings from the Baseline



In sum, the language attitude results showed different patterns across dimensions.

Expectation mismatch was evident only in the affective dimensions, as the Korean guise received

less favorable evaluations in positive and negative affect. By contrast, status, solidarity, and comprehensibility did not show clear evidence of mismatch effects. Evidence for race- or in-group bias also showed mixed results in that negative affect showed limited directional evidence of a race-based effect, while status, solidarity, and comprehensibility showed only limited or contrast-specific patterns.

3.4.2 Relationships among language attitude dimensions

All three dimensions, status, solidarity, and comprehensibility, were positively correlated (Table 3.8). Status and solidarity showed a relatively strong association ($\rho(123) = .58$, $p < .001$). Comprehensibility was moderately correlated with status ($\rho(123) = .46$, $p < .001$) and slightly more weakly correlated with solidarity ($\rho(123) = .32$, $p < .001$).

Table 3.8
Spearman correlations among language attitude dimensions

Measure	Status	Solidarity	Comprehensibility
Status	-		
Solidarity	.57***	-	
Comprehensibility	.46***	.32***	-

Note. Values represent Spearman's ρ. ***$p < .001$.

3.4.3 Correlation with intelligibility and accentedness

For whole sentence intelligibility, no reliable correlations were observed with positive affect, status, solidarity, and comprehensibility ratings. However, negative affect showed a significant negative correlation with whole sentence intelligibility ($\rho(123) = -.27$, $p = .003$), and it indicates that less intelligible speech (i.e., speech that included more transcription errors) tended to receive more negative affect evaluations (Redford et al., 2018).

A similar pattern appeared for target word intelligibility. No correlations were found for positive affect, status, solidarity and comprehensibility, but negative affect was significantly associated with target word intelligibility ($\rho(123) = -.20$, $p = .023$), meaning lower intelligibility corresponding to stronger negative affect judgments.

On the other hand, accentedness was not significantly correlated with any of the language attitude dimensions.

Table 3.9
Correlations between language attitude and speech perception measures

Attitude	Whole sentence Intelligibility	Target word intelligibility	Accentedness judgment
Positive affect	.03	.06	-.03
Negative affect	-.27**	-.20*	.10
Status	.05	.04	-.09
Solidarity	.05	.06	-.03
Comprehensibility	-.03	.01	-.04

Note. Values are Spearman's ρ, * $p < .05$, ** $p < .01$ ***$p < .001$.

3.4.4 Correlation with demographic & language background survey

The overall correlations are presented in Table 3.10. For positive affect, exposure to media featuring Korean-speaking foreigners (Media) and frequency of consuming such media (MediaF) were positively correlated with positive affect ratings (Media: $\rho(123) = .31$, $p < .001$; MediaF: $\rho(123) = .35$, $p < .001$). Listeners who reported exposure to media featuring Korean-speaking foreigners gave higher positive affect ratings than those who reported no such exposure, and more frequent consumption of such media was also associated with higher positive affect ratings. Other background variables showed negligible correlations.

For negative affect, perceived importance of pronunciation in second language learning (Importance) was positively correlated with negative ratings ($\rho(123) = .29$, $p < .01$). Listeners who think pronunciation is important for acquiring a second language rated the speaker more negatively. Frequency of conversation with Korean-speaking foreign friends (ConvF) also showed a positive correlation ($\rho(123) = .18$, $p < .05$). It means that listeners who have more conversation with Korean-speaking foreign friends in Korean tended to rate the speaker more negatively.

For the status dimension, experience living abroad (LivingA) was positively correlated with status ratings ($\rho(123) = .19$, $p < .05$), and frequency of consuming media featuring Korean-speaking foreigners (MediaF) showed a small positive correlation ($\rho(123) = .09$, $p < .01$).

For solidarity, exposure to media featuring Korean-speaking foreigners (Media; $\rho(123) = .34$, $p < .05$) and frequency of consuming media (MediaF; $\rho(123) = .20$, $p < .01$) were positively associated with solidarity ratings, meaning greater media exposure was associated with higher solidarity ratings.

Finally, for comprehensibility, perceived importance of pronunciation was negatively associated with comprehensibility ratings ($\rho(123) = -.20$, $p < .001$). It indicates that listeners who think pronunciation is important for learning a second language tended to give lower comprehensibility ratings. Across dimensions, effect sizes were generally small to moderate, and many demographic variables (e.g., Age, Education) showed no correlations with language attitude measures.

Table 3.10
Correlations between language attitude and listener background variables

Attitude	Age	Educ	Dialect	AOA	L2N	LivingA	Friend	ConvF	Media	MediaF	Importance
Positive	.03	.07	-.09	-.06	-.02	.07	-.00	.03	**.31*** **	**.35*** **	.15
Negative	.03	.10	-.16	-.09	-.06	-.11	.06	**.18***	.01	.18	**.29****
Status	.02	-.02	.08	-.01	.05	**.19***	.11	-.02	.28	**.09****	-.15
Solidarity	-.06	-.13	-.05	-.11	.00	-.00	.05	.03	**.34***	**.20*** **	.03
Comprehensibility	-.15	-.05	.10	-.09	-.07	.07	.03	-.05	.05	-.05	**-.28****

Note. $p < .05$*, $p < .01$**, $p < .001$***.
Abbreviations. Educ = highest education level; Dialect = whether listeners use a regional dialect in Korean; AOA = age of acquisition of L2; L2N = number of L2s; LivingA = experience living abroad; Friend = having Korean-speaking foreign friends; ConvF: frequency of conversation with Korean-speaking foreigners; Media = exposure to media featuring Korean-speaking foreigners; MediaF = frequency of consuming media that feature Korean-speaking foreigners; Importance = perceived importance of pronunciation and accent in L2 learning

Overall, media exposure to Korean-speaking foreigners was more associated with favorable evaluations such as positive affect, status, and solidarity, whereas beliefs about the importance of pronunciation were associated with less favorable evaluations such as negative affect and low comprehensibility.

3.5 Discussion

The purpose of this chapter was to examine how native Korean listeners' language attitudes toward foreign-accented Korean differ when identical speech is paired with different social information, and how these evaluations relate to the results of speech perception tasks reported in Chapter 2. Specifically, this chapter investigated whether 1) race-based expectations are reflected in language attitude, especially on status evaluations, 2) ethnic in-group preference appears in language attitudes, particularly in solidarity evaluations, 3) expectation mismatch effects influence language attitudes, and 4) these evaluations are associated with perceived comprehensibility and 5) listener-related factors.

Overall, the results show that social information influenced each language attitude dimension in specific ways. Expectation mismatch effects emerged only in affective evaluations, and negative affect showed limited directional evidence consistent with a race-based effect. In contrast, status, solidarity, and comprehensibility showed no consistent evidence for race-based or in-group biases, despite some directional tendencies in specific contrasts.

3.5.1 Positive and negative affect evaluations

In positive affect ratings, the Korean (mismatch) guise was rated less favorably than both the Baseline and the White conditions. In negative affect ratings, the Korean guise received

higher negative affect ratings than the Baseline and directionally than the White guise. These patterns indicate that the mismatch effects influence positive and negative affective responses overall and align with the findings of the accentedness judgment task reported in Chapter 2, in which the Korean guise received higher accentedness ratings relative to the Baseline. This result suggests that expectation mismatch may extend beyond accentedness judgment and influence affective evaluations, consistent with the exemplar account of speech perception (Gnevsheva, 2018; McGowan, 2015). When incoming signal (i.e., accented speech) does not match the expectation (i.e., Korean speaker should not have a foreign accent), it leads to negative judgment of the speaker.

However, correlation analyses indicated that positive affect was not associated with accentedness ratings. This distinction suggests that affective evaluations are not directly driven by accentedness judgment, instead, expectation mismatch may have influenced different types of perception task and evaluation in a different degree in that the Korean guise was less favorably evaluated than the White in affective ratings, which was not observed in the accentedness judgment task reported in Chapter 2.

In addition, in negative affect ratings, the Korean-American and Chinese guises were evaluated more negatively than the Baseline, and posterior probabilities showed strong directional evidence that the Chinese guise was also rated more negatively than the White guise. It hints at the effect of race, such that Asian-looking speakers were more negatively evaluated.

Overall, the mismatch effects are most clearly observed in affective evaluations, while the negative effect of race cannot be ignored.

3.5.2 Status and solidarity dimensions

Unlike the affective dimensions, status and solidarity did not show any clear evidence of expectation mismatch. This may suggest that these dimensions more directly reflect listeners' social inferences about the speaker.

The status dimension was one of the primary foci of this chapter, as it allowed for testing whether race-based bias or socioeconomic expectations associated with speakers' nationality would be reflected in language attitude evaluations. If White (White) appearance were associated with higher socioeconomic status or prestige, the White guise would be expected to receive reliably higher status ratings compared to the Korean-American and Chinese guises. The results provided only limited support for this prediction. Although posterior probabilities suggested a directional tendency for higher status ratings for the White guise relative to the Korean-American and Chinese guises, the 95% CrI included zero, indicating no reliable evidence for a robust race-based socioeconomic perception.

In addition, no reliable differences were observed among the White, Korean-American and Chinese guises. This pattern is consistent with the alternative possibility outlined in Section 3.1.1. Namely, all three guises may be associated with relatively high socioeconomic status based on their national backgrounds (i.e., the U.S. and China), and therefore, no clear differences would appear in status ratings among them.

At the same time, the Korean-American guise was rated lower than the Baseline condition, and the Chinese guise showed a similar but no reliable tendency. This pattern suggests that Korean listeners' perception of certain social information, particularly in the Korean-American condition, may lower perceived socioeconomic status relative to a speaker presented without any specific social information (i.e., the Baseline). It is notable that the Korean-

American guise received the lowest status ratings despite its association with a socioeconomically high status nationality, the United States, suggesting that shared ethnicity combined with foreign-accented speech may have created a socially ambiguous position that negatively affects status evaluation.

In the solidarity dimension, ethnic in-group preference would predict that the Korean-American guise and potentially the Korean (mismatch) guise should receive higher ratings than the Baseline, White, and Chinese conditions. However, neither the Korean-American nor the Korean guise showed any advantage based on shared ethnicity in solidarity. Instead, the Korean-American was rated lower than the White condition. This finding is not consistent with previous findings where clear in-group favoritism was observed (Abrams & Hogg, 1987; Dragojevic et al., 2013; Maass et al., 1996). Although racial in-group preference was not the main focus of this chapter, it remains noteworthy that the Chinese guise, which is racially, geographically, and culturally closer to Korea than the White American guise, did not receive higher solidarity ratings either. In fact, the White guise, which is the most socially distant from Korea among the conditions, received the highest solidarity ratings, whereas the Korean-American guise, arguably the closest to Korea, received lowest ratings.

The fact that the White guise received the highest solidarity ratings, despite being racially and culturally most distant from Korean listeners, therefore requires closer examination. One possible interpretation is the presence of White favoritism or Western prestige bias in the Korean context. Although the race-based socioeconomic status hypothesis was not robustly supported in the status dimension, solidarity ratings may be influenced by symbolic associations that extend beyond socioeconomic evaluation. Prior research suggests that whiteness has carried considerable cultural value in South Korea, as it has often been associated with Western

modernity and positive representations of the U.S. (J.-H. Ahn, 2015; N. Y. Kim, 2008). Research on Korean media and celebrity culture has documented a preference for whiteness, particularly in the prominent visibility and popularity of mixed-race celebrities born to one Korean and one White parent. Public figures such as Daniel Henney (Korean-American mother and Irish-American father), Julien Kang (Korean father and French mother), and mixed-race appearing in Korean media, such as the siblings Naeun and Geonhoo (Korean father and Swiss mother) and William and Bentley (Korean mother and Australian father), have been interpreted as evidence of a social preference for whiteness over other racialized identities (J.-H. Ahn, 2015; Lo & Kim, 2011; Park, 2011). These cultural patterns suggest that whiteness may carry positive social meanings that extend beyond perceptions of socioeconomic status.

On the other hand, the Korean-American guise shares racial and ethnic features with Korean listeners, but was paired with foreign-accented Korean. This combination may have created a socially ambiguous position in which the speaker is ethnically proximate but linguistically marked as non-native. According to Social Identity Theory (Tajfel & Turner, 1986), individuals are motivated to maintain a positive self-concept and reduce self-uncertainty through their group memberships (Hogg, 2000). In doing so, they tend to build favorable social identities that often leads to strong in-group preferences (Hewstone et al., 2002). However, when group boundaries become ambiguous, uncertainty may increase, potentially weakening the strength of perceived group membership.

Accent may further affect this boundary perception. Prior research has shown that accent can function as a cue for group membership, sometimes overriding racial appearance (Bestelmeyer et al., 2015a; Kinzler et al., 2009). It means native accents tend to signal in-group membership, whereas foreign accents index out-group membership. In the present chapter, the

Korean-American guise was consistently paired with foreign-accented Korean. It is therefore possible that accent strength functioned as a cue for boundary categorization, positioning the Korean-American outside the linguistic in-group despite shared ethnicity.

Related findings have been reported by Hansen et al. (2018), who examined how the order of racial information and accent influences evaluations. When a foreign-looking appearance (e.g., Turkish) was paired with native-accented speech (e.g., German), participants showed positive expectancy violations and evaluated the candidate as more competent and more hirable. In contrast, when native German-accented speech was presented first and then paired with a Turkish appearance, evaluations became more negative. Although the design of the present study differs from that experiment, a similar expectancy mechanism may be relevant. In the current study, the written description presented first may have shaped listeners' expectations. The description "American learner" paired with a foreign accent may have been perceived as relatively expected, whereas the description "Korean-American," which might imply higher proficiency in Korean, may have produced a stronger negative expectancy violation when paired with foreign-accented speech than the White guise.

Notably, in addition, this phenomenon appears to be specific to the Korean-American guise rather than the Korean (mismatch) guise. This suggests that this evaluative pattern cannot be explained solely by mismatch between identity and speech. Instead, the combination of ethnic similarity and linguistic deviation may have played a particularly salient role for the Korean-American condition.

The observed pattern – higher solidarity ratings for the White and lower ratings for the Korean-American - resonates with two neologisms used in contemporary South Korea, 대한외국인 (Daehan oegugin, "*honorary Korean*") and 검머외 (Geom-meo-oe, short for 검은머리

외국인 [geomeun meori oegugin], "*black-haired foreigners*"). The term honorary Korean refers to non-Koreans who are highly proficient in Korean and are perceived as having become culturally "Koreanized." In contrast, the term black-haired foreigners refers to ethnic Koreans, often Korean-Americans, who are perceived as "foreign" due to their overseas upbringing or citizenship.

These two expressions reflect South Koreans' contrasting perceptions of two groups. Korean-speaking foreigners remain ethnic out-group members, but they may be socially and culturally included through linguistic competence and association with "Koreanness." In contrast, overseas Koreans share ethnic background but may be treated as out-group members when they are perceived as culturally or linguistically distant. This contrast helps explain why the Korean-American guise was evaluated less favorably despite its ethnic similarity to the in-group.

3.5.3 Comprehensibility

Another dimension examined in this chapter was listeners' perceived comprehensibility of the speaker. Unlike objective intelligibility measures, comprehensibility reflects listeners' subjective judgment of how easy the speech is to understand, and therefore may be influenced by both speech and social expectations.

In comprehensibility ratings, the Baseline condition was rated as more comprehensible than both the Korean-American and Chinese guises, and a similar directional tendency was observed for the White and Korean guises relative to these two conditions. These patterns suggest that listeners' perceived ease of understanding was not influenced solely by the speech signal, but also by social information about the speaker. However, it is difficult to attribute the

results solely to race-based bias. If race-based bias strongly shaped listeners' evaluations, it is expected that the White guise shows a clear advantage over the other conditions. As the Korean (mismatch) guise also showed a similar directional tendency with the White guise relative to the Korean-American and Chinese guises, it suggests that race-based bias alone cannot fully explain this pattern. Importantly, this pattern is not consistent with the expectation mismatch account either. If expectation mismatch increased processing costs, the Korean guise would be expected to receive lower comprehensibility ratings than the other guises. However, this was not observed as the Korean guise showed comparable ratings to the Baseline and White conditions. This suggests that expectation mismatch may not necessarily result in lower perceived comprehensibility, but may instead primarily influence affective evaluations.

While one explanation suggests that social information may influence language attitude and perceived comprehensibility, the correlation analyses further indicate that comprehensibility itself is also closely related to listeners' social evaluations of the speaker. In this chapter, comprehensibility was positively correlated with both status and solidarity, although the association was stronger with status than solidarity. This result suggests that listeners' higher comprehensibility was more associated with higher perceived intelligence, education, or competence, while its relationship with solidarity-related traits was relatively weaker.

These results also help explain why the mismatch effect did not appear in comprehensibility. Given that comprehensibility is closely associated with status and solidarity, the absence of mismatch effects in these dimensions accounts for the absence of the mismatch effect in comprehensibility.

Although the direction and causality of these association remain unclear, this finding is partly consistent with previous studies that argued that listeners tend to attribute communication

difficulties primarily to the speaker (Lippi-Green, 2012). Under this assumption, speakers are expected to convey their message in a way that is easy to understand. As a result, when listeners encounter difficulty understanding speech, they may interpret this difficulty as reflecting limitations in the speaker's communicative competence and evaluate the speaker less favorably, particularly on status and solidarity dimensions (Dragojevic & Giles, 2016; Oppenheimer, 2006). In the study of Dragojevic et al. (2017), heavily accented speakers are generally evaluated more negatively than mildly accented speakers because their speech disrupts listeners' processing fluency (i.e., comprehensibility). As processing fluency decreases, negative affective responses increase, which in turn lowers evaluations of status and solidarity (Dragojevic & Giles, 2016). Although the present study did not directly manipulate accent strength, the observed correlations between comprehensibility and social evaluation dimensions suggest that listeners' perceived ease of understanding speech may influence their evaluations of the speaker in a similar way. These findings, therefore, indicate that accentedness, language attitudes, and comprehensibility may be interrelated to some extent.

3.5.4 Relationship between language attitudes and perceptual measures

This chapter examined the correlations between language attitude evaluations and listeners' performance on the intelligibility and accentedness judgment tasks reported in Chapter 2 in order to further assess the bias account.

The overall results showed limited associations between perceptual performance and language attitudes. Whole sentence and target word intelligibility were not reliably correlated with most evaluative dimensions, including positive affect, status, solidarity, and

comprehensibility. This pattern suggests that listeners' social evaluations of the speaker were largely independent of their actual ability to decode and understand the speech.

However, negative affect showed a significant negative correlation with intelligibility, which indicates that speech perceived as less intelligible tended to receive more negative evaluations. In other words, when listeners experienced difficulty understanding the speech, they were somewhat more likely to evaluate the speaker negatively such as frustration or annoyance. Although the direction of association remains unclear, it is partly consistent with previous research suggesting that listeners may attribute communication difficulty to the speaker (Lippi-Green, 2012).

Similarly, accentedness judgments showed limited associations with language attitude evaluations. This pattern is consistent with the findings reported in Chapter 2, where accent ratings were not related to social bias but were more sensitive to expectation mismatch.

Taken together, these results indicate that perceptual measures and social evaluations capture different aspects of listeners' responses to foreign-accented speech in the Korean context. Intelligibility reflects listeners' objective ability to decode and understand the speech and showed only limited associations with language attitude evaluations. By contrast, accentedness judgments represent a more subjective perceptual evaluation that compares the speech with listeners' expectations about native speech norms. Therefore, as shown in Chapter 2, accentedness ratings were not systematically associated with listeners' social evaluations but sensitive to expectation mismatch.

3.5.5 Listeners' demographic and language background factors

In addition, this chapter explored whether listeners' demographic and language backgrounds were associated with their evaluations of the speaker. The results suggested that individual differences among listeners may play a role in shaping language attitudes toward foreign-accented Korean. With regard to living abroad experience, although all participants in this study had no experience living abroad for more than one year before the age of 18, some reported having lived abroad after adulthood. Those participants tended to rate the speaker higher on status dimension. This result shows that exposure to different cultural and linguistic environments may lead listeners to evaluate foreign-accented speakers more positively in status dimension.

Similarly, participants who reported exposure to media featuring Korean-speaking foreigners, such as the JTBC talk show *Non-Summit*, showed higher ratings in positive affect and solidarity. Furthermore, as the frequency of such media exposure increased, listeners tended to evaluate the speaker more positively across positive affect, status and solidarity dimensions. These results align with intergroup contact hypothesis (Allport, 1954; Pettigrew, 1998), which posits that contact between members of different social groups can reduce prejudice and improve intergroup attitudes, particularly when such contact occurs under conditions of equal status, cooperation, and institutional support. Although the contact of participants in this study was largely indirect through media, frequent exposure to media featuring foreigners speaking Korean may have functioned as a form of extended contact (Wright et al., 1997). According to this hypothesis, observing positive interactions between in-group and out-group members can enhance intergroup attitudes by providing positive exemplars of out-group members and promoting their inclusion within one's self-concept. In this context, media portrayals of friendly

and cooperative interactions between Koreans and foreigners may serve a similar role, shaping viewers' perceptions and enhancing positive affect toward foreign-accented speakers. These results provide promising outlook that more experience abroad and more exposure to accented speakers could reduce negative perception of accents and accented speakers.

However, not all listener-related factors showed consistent patterns with the contact hypothesis. In particular, the frequency of conversations with Korean-speaking foreign friends was positively associated with negative affect ratings. In other words, listeners who reported that they have frequent interactions with foreign friends learning Korean tended to evaluate the speakers slightly more negatively. One possible explanation is that direct interaction with foreign-accented speakers may have communicative challenges. In contrast to media contexts, where accented speakers are often relatively fluent and their speech is sometimes accompanied by subtitles, real-life conversations may involve greater communicative effort or misunderstandings. Although the current study did not collect qualitative data on this question, listeners who frequently communicate with foreign speakers may have experienced moments of misunderstanding or communicative effort during real conversations. Such experiences could lead to feelings of frustration or impatience, which may in turn contribute to more negative affective evaluations of accented speech.

This interpretation suggests the possibility that different forms of contact may influence attitudes in different ways. While exposure through media may provide positive or entertaining representations of accented speakers, direct interaction may also involve communicative difficulties that shape listeners' perceptions of foreign-accented speech.

Another pattern observed was that listeners' beliefs about the importance of pronunciation and accent in second language learning were associated with negative affect and

comprehensibility ratings. It indicates that listeners who put emphasis on pronunciation and accent may evaluate foreign-accented speech more strictly, and interpret deviations of pronunciation and accent as indicators of lower communicative competence or effort. In this sense, listeners' metalinguistic beliefs about language learning may shape how they evaluate accented speech.

Finally, many other demographic and language background factors did not show significant correlation with language attitude evaluations. It demonstrates that although certain types of experience and language beliefs may influence Korean listeners' attitudes, many individual background factors do not systematically shape evaluations of foreign-accented Korean.

Future research may further examine how different comparison frames influence listeners' evaluations of foreign-accented Korean. In this study, a between-subject design was used, in which the same foreign-accented speech was paired with five different guise conditions. Because language attitudes are not fixed cognitive structures but fluid and context-dependent, this design may have reduced sensitivity to subtle differences across conditions. At the same time, listeners' evaluations of a particular speaker may depend on the relative comparison with other accents and the listener's frame of reference (Dragojevic & Giles, 2014). For example, the Korean-American guise in this study did not show an in-group advantage in solidarity. However, different patterns could emerge if Korean-American guise was paired with less accented Korean speech, or evaluated alongside multiple accented speakers within a within-subject design. Exploring such comparisons may provide further insight into how Korean listeners perceive in-group membership and how they position different types of foreign-accented Korean within their internal hierarchy of accent evaluations.

Beyond its theoretical implications, these findings have practical relevance for understanding intercultural communication in South Korea's rapidly changing social environment. As the number of foreigners learning and using Korean continues to increase, listeners' attitudes toward foreign-accented Korean may influence not only interpersonal communication but also broader issues of social integration and inclusion of foreign residents. The results suggest that increasing positive and meaningful direct and indirect exposure to foreign-accented speakers through media, education and everyday interaction, could help reduce affective bias and foster more inclusive and positive perceptions of linguistic and racial diversity.

3.6 Conclusion

This study contributes to the relatively small body of language attitude research by applying the status-solidarity framework to speech perception data in the South Korean context. While most previous studies have focused on English-speaking and Western European societies or have relied on survey-based responses in Korea, the present study incorporated real audio stimuli and manipulated social cues to examine how Korean listeners evaluate foreign-accented Korean. In doing so, it highlights how listeners' evaluations of accented speech are shaped not only by the speech signal itself but also by sociocultural expectations and perceptions, and listeners' individual experiences.

Contrary to expectations derived from some previous studies (Bestelmeyer et al., 2015b; Kinzler et al., 2009; Shamoa-Nir et al., 2021; Tajfel & Turner, 1986; Wade, 2025), ethnic or racial in-group preference was not observed in the present data. Although the Korean-American and Chinese guises shared ethnic or racial similarities with Korean listeners, they were not evaluated more favorably than the Baseline or White guises across several attitudinal

dimensions. This result suggests that language attitudes toward foreign-accented Korean in South Korea may be influenced by a complex interaction of sociocultural expectations and social categorization rather than by a simple in-group preference.

These findings underscore how communication and social integration can be disrupted when biases based on race or ethnicity influence how speakers are perceived. However, the responsibility for successful communication should not fall solely on non-native speakers, as communication is inherently bidirectional. Raising awareness of such biases among native listeners is therefore essential for fostering more inclusive and mutually accommodating communicative environments.

Finally, language attitudes are not static. They evolve alongside the sociopolitical and educational changes. As South Korea continues to implement policies and educational initiatives to support multicultural children and families, such efforts are likely to influence Koreans' language ideologies and social perceptions. Future research could extend this work by examining how language attitudes shift over time in response to policy and educational changes, by including a wider range of accents and sociocultural cues, and by investigating how individual ideologies and media representations altogether shape attitudes toward non-native speakers of Korean.

IV. ACOUSTIC AND PHONOLOGICAL REALIZATION OF KOREAN STOPS AND VOWELS BY AMERICAN ENGLISH LEARNERS: A COMPARISON WITH NATIVE KOREAN SPEAKERS

4.1 Introduction

The preceding chapters examined how native Korean listeners' social expectations influence speech perception and language attitudes toward foreign-accented Korean. However, such perceptions and evaluations are still made based on speech that has specific phonetic and phonological features. For this reason, the present chapter provides a supplementary analysis of the learner speech by comparing American English learners' Korean productions with those of native Korean speakers.

In particular, this chapter examines two aspects of learner production. First, it investigates the acoustic realization of Korean vowels and stops, including vowel formants (F1 and F2), vowel duration, voice onset time (VOT), and fundamental frequency (F0). Second, it explores recurrent phonological realization patterns such as failures to apply resyllabification and phonological assimilation. Vowels and stop consonants were chosen because vowel quality has often been considered an important cue to perceived accent, and Korean three-way stop contrasts are known to differ from the two-way stop system of English. By examining both acoustic and phonological aspects, this chapter provides a more detailed account of the acoustic and phonological characteristics of American English-accented Korean and of how these productions differ from native Korean speech. Further, it also provides background for the following chapter, which examines how acoustic features of American English accented speech relate to accentedness judgments and language attitude evaluations.

First, Korean and American English differ in both vowel inventory and vowel space. Yang's (1996) acoustic comparison of Korean and American English vowels showed that vowels in the two languages differ in their phonetic realization. This difference does not mean that Korean vowels are completely different from English vowel categories. Rather, American English learners may initially perceive and produce some Korean vowels in relation to phonetically similar English vowels, while other Korean vowels may be less clearly assimilated to existing English categories. From the perspective of L2 speech learning, the relevant issue is whether learners whose vowel categories are shaped by English can produce stable Korean vowel targets. Even when a Korean vowel is broadly similar to an English vowel, the two vowels may still differ in vowel qualities and category boundaries, so direct transfer from English may be incomplete.

At the same time, the Korean vowel system itself is undergoing change. In contemporary Seoul Korean, the mid-front vowels [ɛ] and [e] have been reported to be merged (Kang, 2014; Lee & Cho, 2021), and ongoing changes in the back vowel space have also been documented where [o] is raising and [u] and [ɨ] are fronting (J.-I. Han & Kang, 2013; Y. Kang, 2014; Y. Lee, 2025). These changes are relevant for learners because they may encounter different vowel qualities or category patterns in native Korean speech depending on speakers' generation or regional background. Such variation may make it harder for learners to identify a single stable target for some Korean vowels. Thus, although learners whose first language has a larger and more complex vowel inventory may show some advantages in second language vowel perception (Iverson & Evans, 2007), this does not necessarily mean that American English learners can easily acquire Korean vowels. The learning challenge therefore includes not only the number of

vowel categories in the two languages, but also the phonetic qualities of those categories, and variability in contemporary Korean productions.

When examining learner vowel production, both vowel quality and vowel duration could be considered. Vowel quality, which is normally indexed by F1 and F2, captures the spectral features of vowel categories and vowel duration reflects their temporal properties. However, the role of vowel duration as a phonological contrast has been limited in contemporary Seoul Korean and the historical vowel length contrast has been reported to be nearing completion of merger (Y. Kang et al., 2015; Magen & Blumstein, 1993). In contrast, although vowel duration does not usually function as an independent phonological contrast in English, it can function as a secondary cue in certain English vowel contrasts (Hillenbrand et al., 1994; Hillenbrand et al., 2000). Therefore, vowel duration is not treated as a primary contrastive property of Korean vowels in native production, but it may still capture differences between native and learner productions.

With respect to stops, Korean and American English differ in the phonetic cues used to distinguish stop categories. English is generally described as having a two-way stop contrast. In word-initial position, this contrast is typically realized through a difference in voice onset time (VOT), with voiced stops showing short-lag VOT and voiceless stops showing long-lag VOT. By contrast, Korean has a three-way laryngeal contrast among lenis, aspirated, and fortis stops. Earlier studies described Korean stops as differing primarily in VOT, with fortis stops having the shortest VOT, lenis stops intermediate VOT, and aspirated stops the longest VOT (M. S. Han & Weitzman, 1970; Lisker & Abramson, 1964). More recent studies, however, have reported a sound change in Seoul Korean in which the VOT difference between lenis and aspirated stops has been reduced, while F0 in the following vowel has become more important in maintaining

the contrast. In particular, lenis stops remain distinct from aspirated stops through lowering F0 values, whereas aspirated and fortis stops are associated with relatively higher F0 and still differentiated by VOT (K.-H. Kang & Guion, 2006; H. Lee, 2016b; Silva, 2006).

These contrasts may be especially challenging for American English learners of Korean because English does not employ the same combination of VOT and F0 to distinguish stop categories. Kim and Lotto (2002) found that American English learners relied more on VOT than on F0 in distinguishing Korean stops, but not in a native-like manner. Several studies have also shown substantial interspeaker variation in producing three-way Korean stops, and suggested that even learners who share the same L1 do not necessarily converge on a single strategy for constructing Korean stop categories (Chang, 2009; J.-I. Han & Kim, 2014; Holliday, 2015). The present chapter therefore reexamines this issue by comparing American English learners' productions with those of younger native Seoul Korean speakers. In doing so, it provides a basis for the following chapter, which examines how acoustic properties of learner speech is related to accentedness judgments.

Finally, in addition to segmental differences, learner speech may also differ from native speech in applying phonological patterns of the target language. In Korean, the phonetic instantiation (i.e., pronunciation) of a segment is often influenced by its phonological environment rather than by its orthographic form alone. Phonological rules such as resyllabification across syllable boundaries, nasal or liquid assimilation, and [h]-induced aspiration frequently affect the surface form in Korean. Accordingly, producing native-like Korean requires not only accurate articulation of segments, but also appropriate application of those phonological patterns. These patterns may be difficult for American English learners of

Korean because Hangul orthography can cause more of segment-by-segment pronunciation when the target form differs from the written representation.

In sum, this chapter examines how American English learners' Korean differs from native Korean at both the acoustic and phonological levels, and which aspects of learner production may later serve as potential sources of accentedness judgments.

4.1.1 Research questions

(1) **Research Question 1**: How do American English learners of Korean differ from native Korean speakers in the acoustic realization of Korean vowels, in terms of vowel quality (F1 and F2) and vowel duration?

(2) **Research Question 2**: How do American English learners of Korean differ from native Korean speakers in the acoustic realization of Korean stops in terms of VOT and F0 across prosodic positions and laryngeal categories?

(3) **Research Question 3**: What recurrent phonological realization patterns appear in the productions of American English learners of Korean, and what do these patterns suggest about the sources of foreign accent in connected speech?

4.2 Methodology

4.2.1 Participants

Eleven female American English learners of Korean, aged 20-28 (M = 22.6, SD = 2.41) and eleven female native Korean speakers, aged from 19 to 31 (M = 22.2, SD = 3.37) participated in a production task in which they read Korean sentences. All participants were assigned female at birth, and one American participant identified as agender. This study included

only speakers assigned female at birth to minimize variability related to sex-based acoustic differences and to maintain consistency with earlier chapters. The learners and two native speakers were students at the University of Oregon at the time of recording (one native speaker was an exchange student and the other a graduate student), and other native Korean speakers were students at a university in Seoul, Korea. All participants reported no language or hearing problems. One learner reported stuttering, but stuttering was not observed in the recordings and the data were retained. Data from one native speaker were partly lost and, therefore, ten native Korean speakers were included in the analysis.

Among the American English learners of Korean, all identified that English is their first and primary language, and three of them additionally reported that they are early bilinguals at home in Spanish, Japanese, and Chinese respectively. Their ages of acquisition of Korean ranged from 6 to 24 years (M = 16.27, SD = 4.79), and their length of learning Korean was from 2 to 19 years (M = 6.36, = 4.88). Ten participants were enrolled in a Korean language program (1 completed first-year Korean, 3 completed second-year Korean, and 6 completed third-year Korean). Their length of stay in Korea ranged from 0 to 36 months (M = 5.18, SD = 11.32).

Among the 10 native Korean speakers, eight people were living in Seoul or the surrounding metropolitan area and two were living in the United States at the time of recording. All reported using standard Seoul Korean. They have little to no experience living abroad, with the exception of one graduate student had been living in the U.S. for three years.

4.2.2 Speech materials and recording

The Korean sentences were drawn from the Korean Speech Recognition (KSR) corpus developed by Song et al. (2023), which contains sentences designed to assess speech recognition

and lexical processing. The corpus includes sentences grouped by predictability levels (predictable, neutral, and anomalous), which allows it to be used with both native and non-native Korean speakers across varying levels of sentence complexity and vocabulary. For the production task, 60 sentences were selected from the set of 120 neutral sentences. Neutral sentences were chosen instead of anomalous ones to avoid potential effects of nonsensical lexical items on accentedness judgments (see Chapter 2) (Bergeron & Trofimovich, 2017; Varonis & Gass, 1982). In addition, the selected sentences had final-word cloze probabilities between 20% and 40% for native speakers and were controlled for length: each sentence had 13-15 syllables, and the final words had 3-4 syllables. The original Korean sentences and their English translations are provided in Appendix A.

The order of the 60 sentences was randomized for each participant and presented on a computer screen using Gorilla Experiment Builder (https://gorilla.sc). The sentences were recorded individually in the sound-attenuated booth using a portable recorder (TASCAM DR-100mkII) with a 44.1kHz sampling rate and a 16-bit quantization level, using a clip microphone (Audio-Technica AT8537). Participants were instructed to read each sentence aloud three times at a natural pace. If they made a mistake, they were allowed to repeat the sentence.

Of the three repetitions of each sentence, the second repetition was later separated into individual files using Praat (version 6.4.47; Boersma & Weenink, 2025). For the learners' productions, if the second repetition contained a mispronunciation or severe disfluency, the third repetition was instead selected and split into a separate file. In total, 1,260 sentence tokens were made (21 speakers x 60 sentences).

4.2.3 Acoustic measurements and phonological coding

All individual sentence files were first segmented using the Montreal Forced Aligner (MFA; McAuliffe et al., 2017). The resulting segmentation boundaries were manually checked in the waveform and confirmed both aurally and using wideband spectrograms. Some MFA-generated labels were manually converted to commonly used IPA-based notation for Korean stops such as [t*] for fortis ㄸ and [t^h] for aspirated ㅌ in Praat.

In addition, sentence-final -ta tokens were excluded from the acoustic analysis because sentence-final position may induce F0 lowering. This exclusion removed 1,260 tokens in total (21 speakers x 60 items)

4.2.3.1 Vowels

Vowel duration was measured as the temporal interval between the onset and offset boundaries of each vowel segment in the TextGrid. Because vowel duration can be influenced by speech rate (Solé & Ohala, 2010), speech rate was calculated for each sentence and included as a covariate in the statistical analysis. Speech rate was calculated as the number of syllables in a sentence divided by the total duration of the sentence, measured from the onset of the first segment to the offset of the final segment. The frequencies of the first and second formants (F1 and F2) were also measured for eight Korean monophthongs, [a], [ʌ], [o], [u], [ɨ], [i], [ɛ], and [e], at the vowel mid-point using Python scripts with the *parselmouth* package (Jadoul et al., 2018). To reduce variation in formant values attributable to physiological differences across speakers, such as differences in vocal tract size, the vowel formants were normalized using the Lobanov method (Lobanov, 1971; Thomas & Kendall, 2007), following the formula:

$$F_{n[V]}^{N} = (F_{n[V]} - MEAN_n)/S_n$$

where $F_{n[V]}^{N}$ represents the normalized value of formant n for vowel V, $MEAN_n$ is the mean value of formant n for the given speaker, and S_n denotes the standard deviation for the speaker's formant n for that speaker (Thomas & Kendall, 2007).

4.2.3.2 Stops

VOT was measured from the onset of the stop release burst to the onset of periodic voicing. Because VOT is also known to be influenced by speech rate (Beckman et al., 2011; Kessinger & Blumstein, 1997; Oh, 2008) and the place of articulation of stops (Nearey & Rochet, 1994; Volaitis & Miller, 1992), speech rate was calculated for each sentence produced by each speaker and the place of articulation of each stop was also coded. In addition, each stop token was coded for position within the sentence to distinguish phrase-initial position (i.e., the first segment of the sentence), word-initial position, and word-medial position. Aspirated and fortis stops did not occur in phrase-initial position in the current dataset. For intervocalic stops, fully voiced stop tokens were excluded because VOT could not be reliably measured for these cases. This exclusion affected 2226 tokens in 1260 sentences (21 speakers x 60 sentences), including 1170 tokens from 10 native Korean speakers and 1056 tokens from 11 American English learners.

F0 was measured at 15 ms after the onset of the following vowel. Previous studies of Korean stop production have measured F0 at different temporal points within the following vowel, including the mean over the first 15 ms of the vowel (Holliday, 2015), at 20 ms after the voicing onset (Kong et al., 2011), and at the vowel midpoint (K.-H. Kang & Guion, 2006). Despite these methodological differences, these studies report robust F0 differences across Korean phonation types. A 15 ms measurement point was chosen instead of 20 ms because some

following vowels were shorter than 20ms, which was unable to measure F0 for certain tokens. Measuring F0 at 15 ms after vowel onset therefore ensured a consistent measurement point across tokens.

4.2.3.3 Phonological realization patterns

In addition to acoustic measurements, phonological realization patterns in the learners' productions were examined. Tokens were inspected for systematic deviations from the target pronunciations based on both auditory judgment and visual inspection of the waveform and spectrogram in Praat. The term "target pronunciation" is used here not to refer to prescriptive norms, but phonological patterns that are conventionally described in Korean phonology and commonly reflected in native Korean speech. In other words, the analysis was based on descriptive patterns of native pronunciation, although many of these patterns are also codified in the Standard Pronunciation Rules issued by the National Institute of Korean Language (National Institute of Korean Language, 2026).

The analysis focused on recurring patterns that appeared consistently across multiple tokens and speakers rather than individuals' isolated errors. The observed patterns were categorized into the following types: 1) failure to apply resyllabification across syllable boundaries (e.g., 갚았다 /kap^{h}-ass-ta/, [ka$\overrightarrow{p}$.a.t*a] instead of [ka.p^{h}a.t*a]; see Figure 4.1), 2) failure to apply nasal or liquid assimilation (e.g., 업무 /up-mu/, [ʌ$\overrightarrow{p}$.mu] instead of [ʌm:u]), 3) failure to realize [h]-induced aspiration after a final lenis consonant (e.g., 급한 /keup-han/, [kɨp.han] instead of [kɨ.p^{h}an]), 4) failure to delete or weaken [h] in contexts where [h] is often deleted or weakened, such as intervocalic or non-initial onset position, or hyperarticulation of [h]

(e.g., 결혼 /kyeol-hon/ [kjʌl.hon] instead of [kjʌ.ron]), and 5) deletion of coda [l] before onset [l] (e.g., 몰래 /mol-le/ [mo.re] instead of [mol:e]) (see also Isbell et al., 2019).

4.3 Analysis

4.3.1 Vowels

Vowels produced by native Korean speakers and learners were analyzed in terms of vowel quality and vowel duration. Vowel quality, which is indexed by the first and second formants (F1 and F2), was normalized using the Lobanov z-score method prior to statistical analysis and vowel duration was analyzed in milliseconds (ms). To compare group differences in vowel quality and duration, Bayesian linear mixed-effects models were fit in R (R Core Team, 2026) using the `brm` function in the *brms* package (Bürkner, 2017). Because F1, F2, and duration were continuous measures, Gaussian likelihoods were specified for all models. Convergence was assessed using the potential scale reduction factor ($\hat{R}$), effective sample sizes (ESS) for both bulk and tail of the posterior distributions, and visual inspection of trace plots of the Markov chain Monte Carlo (MCMC) samples. All parameters met the convergence criteria ($\hat{R} = 1.00$), with adequate effective sample sizes (> 1000) and well-mixed chains.

The models for F1 and F2 included *Group* (native vs. learner), *Vowel*, and their interaction as fixed effects, and random intercepts for *Speaker* and *Sentence* to account for speaker and item variability, as well as a by-*Speaker* random slope for *Vowel*. Four chains were run with 4,000 iterations respectively for both models. To examine differences between groups, posterior contrasts were computed from the fitted Bayesian models. Inference was based on the posterior distributions of these contrasts, summarized by posterior medians, 95% credible

intervals, and posterior probabilities that the learner-native difference was greater than zero or less than zero.

Vowel duration was analyzed in a separate Bayesian linear mixed-effects model with *Vowel*, *Group*, and their interaction as fixed effects, and *Speech Rate* as a covariate. Random effects included intercepts for *Speaker* and *Sentence*, as well as a by-*Speaker* random slope for *Speech Rate*. The duration model was run with four chains and 4,000 iterations each. Posterior contrasts were computed as well to compare group differences, and inference was based on posterior distributions summarized by posterior medians, 95% credible intervals, and posterior probabilities.

In addition, because Korean ㅖ [e] and ㅐ [ɛ] have been reported to be merged (Eychenne & Jang, 2018; Y. Lee & Cho, 2021), a supplementary analysis was conducted for these vowels. In this analysis, native [e] and [ɛ] tokens were combined into a single category due to their merger, and learner [e] and [ɛ] were treated as separate categories. Separate Bayesian linear mixed-effects models were then fit for F1 and F2 with *Vowel* (native [e], learner [e], and learner [ɛ]) as a fixed effect and random intercepts for *Speaker* and *Sentence*. Vowel duration for this subset was analyzed in a separate model with *Vowel* as a fixed effect, *Speech Rate* as a covariate, random intercepts for *Speaker* and *Sentence*, and a by-*Speaker* random slope for *Speech Rate*. The merged native [e] served as the reference level, allowing direct comparison with learner [e] and learner [ɛ] respectively, as well as comparison between learner [e] and learner [ɛ]. Four chains were run with 12,000 iterations for the F1 and F2 models and 4,000 iterations for the duration model. Convergence was evaluated using the same criteria described above.

4.3.2 Stops

For stop sounds, separate Bayesian mixed-effects models were conducted for VOT and F0. Because VOT is known to vary depending on speech rate, *Speech Rate* was included as a covariate in the VOT models. For phrase-initial stops, the VOT model included *Group* (native vs. learner) as a fixed effect, *Speech Rate* as a covariate, and *Place of Articulation* (bilabial, denti-alveolar, velar) as an additional fixed effect. Random intercepts for *Speaker* and *Vowel* were included, with by-*Speaker* random slopes for *Speech Rate*. For word-initial and medial stops, the VOT models included *Group*, *Laryngeal Category* (lenis, aspirated, fortis), and their interaction as fixed effects, together with *Speech Rate* and *Place of Articulation*. These models also included random intercepts for *Speaker* and *Vowel*, with by-*Speaker* random slopes for *Laryngeal Category* and *Speech Rate*.

For F0, separate Bayesian mixed-effects models were also conducted. For phrase-initial stops, the model included *Group* and *Place of Articulation* as fixed effects, with random intercepts for *Speaker* and *Vowel*. For word-initial and medial stops, the F0 models included *Group*, *Laryngeal Category*, their interaction, and *Place of Articulation* as fixed effects. Random intercepts for *Speaker* and *Vowel* were included, with by-*Speaker* random slopes for *Laryngeal Category*.

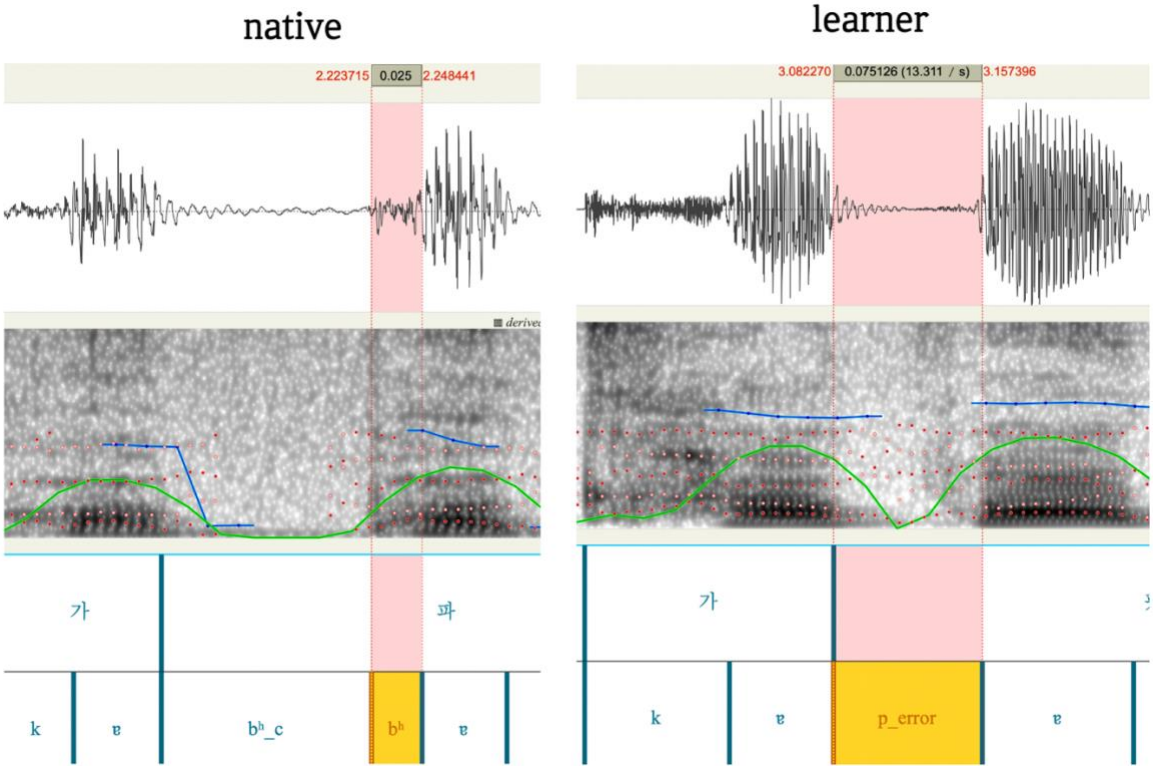
All models were estimated using the `brm` function in the *brms* package (Bürkner, 2017), and each model was run with four chains of 8,000 iterations. Group differences were examined using posterior contrasts. For phrase-initial stops, contrasts compared native and learner speakers in lenis stops as no other laryngeal categories were included in this position. For word-initial and medial stops, native and learner contrasts were conducted separately for each laryngeal category.

Inference was based on the posterior distributions of these contrasts, summarized by posterior medians, 95% credible intervals, and posterior probabilities.

4.3.3 Phonological realization patterns

In addition to acoustic measurements, phonological realization patterns in the learners' productions were examined. Each token was first assessed auditorily, and tokens that appeared to deviate from the expected target realization were then checked through visual inspection of the waveform and spectrogram in Praat. The final coding was based on the combination of auditory judgment and acoustic-visual confirmation.

Figure 4.1
Example waveform and spectrogram of native and learner productions of 갚았다 /kap^{h}-ass-ta/



Note. Native production shows resyllabification, [ka.p^{h}a.t*a], whereas the learner production shows a non-resyllabified form [kap̄.a.t*a].

4.4 Results

4.4.1 Vowels

In this section, group differences in vowel production between native Korean speakers and American English learners of Korean are reported in terms of spectral quality based on Lobanov-normalized F1 and F2 values, and temporal quality based on vowel duration.

4.4.1.1 Vowel quality

The group contrasts in F1 and F2 are presented in Table 4.1 and Table 4.2 respectively, and the vowel spaces of native speakers and learners are presented visually in separate panels in Figure 4.2 and in a single panel in Figure 4.3. With respect to the interpretation of normalized vowel formants, higher F1 values correspond to lower tongue height, whereas lower F1 values correspond to higher tongue height. Higher F2 values indicate a more fronted tongue position and lower F2 values indicate a more backed tongue position.

In normalized F1, there was limited evidence of robust group differences for most vowels, as many of the 95% credible intervals included zero. However, native speakers showed lower values than learners for [o] and higher values than learners for [ʌ], with 95% credible intervals excluding zero. Notably, the largest F1 difference was observed for [o], where native speakers showed substantially lower F1 values than learners (median = -0.44, 95% CrI [-0.58, -0.30], P(< 0) = 1.00), indicating a higher tongue position for [o] in the native productions. Native speakers also showed higher F1 values than learners for [ʌ] (median = 0.28, 95% CrI [0.05, 0.51], P(> 0) = 0.99), indicating a lower tongue position in native productions. As shown in Figure 4.3, native speakers' [o] is located in a relatively high position in the back vowel space, closer to that of [u], and is more clearly distinguished from [ʌ] than in learner productions.

Table 4.1
Posterior contrasts of normalized F1 between native and learner groups

Vowel	Median	P(native > learner)	P(native < learner)	95%CrI
a	-0.06	0.05	0.94	[-0.13, 0.01]
e	-0.15	0.05	0.95	[-0.33, 0.03]
ɛ	-0.07	0.17	0.82	[-0.22, 0.08]
i	0.06	0.86	0.13	[-0.05, 0.18]
ɨ	0.05	0.78	0.21	[-0.08, 0.18]
o	-0.44	0.00	1.00	[-0.58, -0.30]
u	0.02	0.63	0.37	[-0.11, 0.15]
ʌ	0.28	0.99	0.01	[0.05, 0.51]

Note. Contrast values (Median) are computed as native minus learner in normalized F1. Positive values indicate higher normalized F1 values for native speakers than for learners, and vice versa. P(native > learner) and P(native < learner) represent the posterior probabilities that native speakers have higher or lower values than learners respectively.

In normalized F2, native speakers showed a tendency toward higher values than learners for [a], [ɨ], [u], and [ʌ], whereas learners tended to show higher values for [ɛ], [e], [i] and [o]. However, most of these differences were not robust, as the credible intervals included zero. The vowels [ɨ] and [u] are still noteworthy, given previous work (Y. Kang, 2014) reporting that these vowels are undergoing fronting in contemporary Seoul Korean. In the present data, native speakers showed somewhat higher F2 values than learners for these vowels, suggesting relatively more fronted realizations, although the evidence for group differences was limited.

Table 4.2
Posterior contrasts of normalized F2 between native and learner groups

Vowel	Median	P(native > learner)	P(native < learner)	95% CrI
a	0.05	0.82	0.17	[-0.05, 0.16]
e	-0.01	0.46	0.54	[-0.29, 0.27]
ɛ	-0.12	0.13	0.86	[-0.34, 0.10]
i	-0.29	0.05	0.94	[-0.65, 0.07]
ɨ	0.07	0.81	0.19	[-0.10, 0.25]
o	-0.04	0.28	0.72	[-0.19, 0.10]
u	0.14	0.91	0.09	[-0.07, 0.35]
ʌ	0.15	0.97	0.03	[-0.01, 0.32]

Note. Contrast values (Median) are computed as native minus learner in normalized F2.

Figure 4.2
Normalized vowel spaces for native Korean speakers and American English learners of Korean

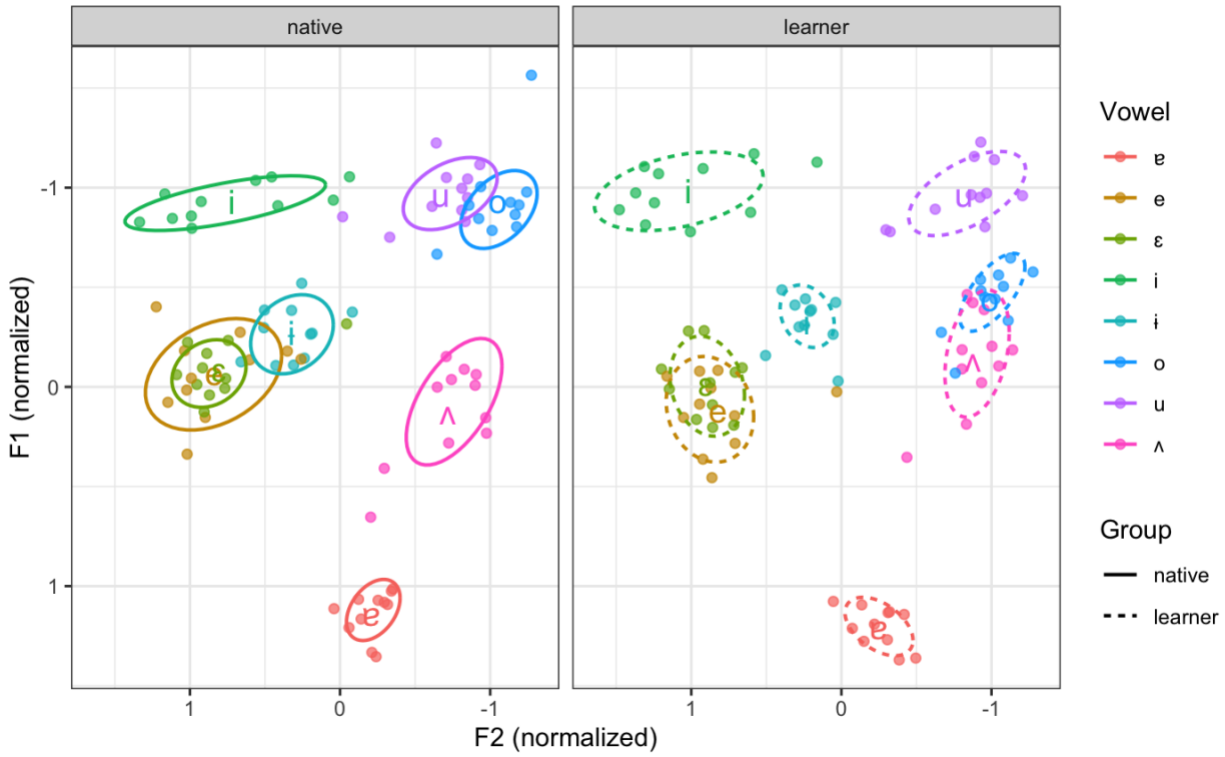
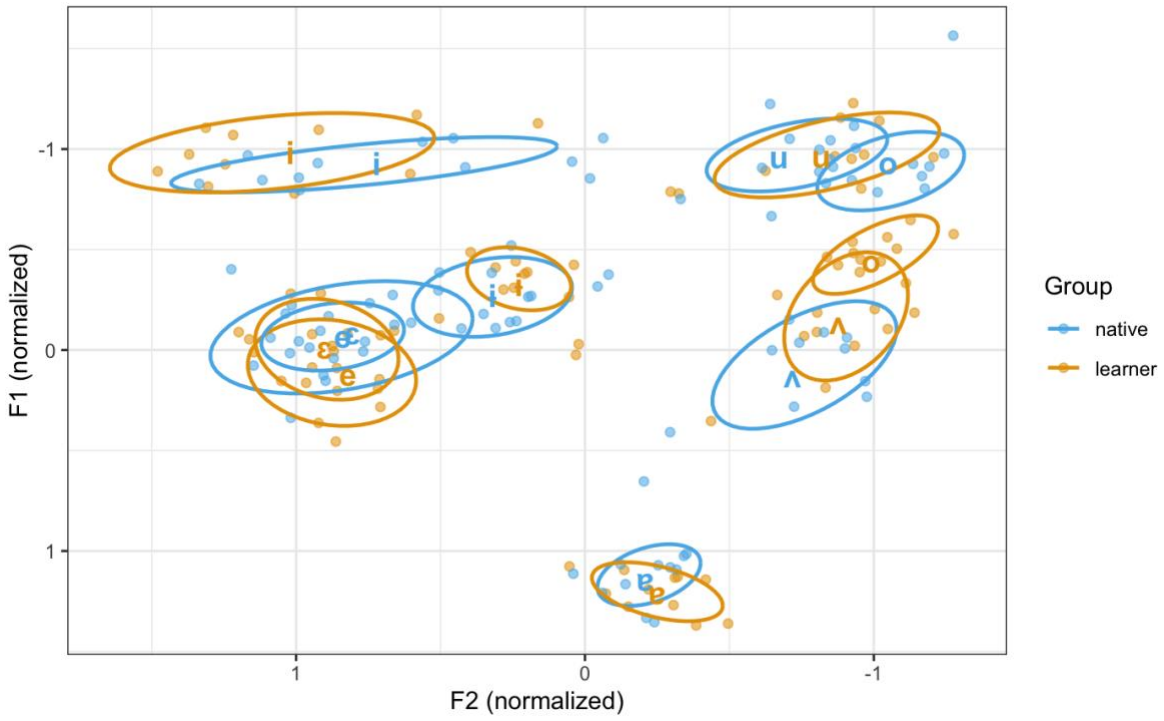


Figure 4.3
Normalized vowel spaces for native Korean speakers and American English learners of Korean in a single plot



4.4.1.2 Vowel duration

Vowel duration contrasts are presented in Table 4.3 and illustrated visually in Figure 4.4. Native speakers produced longer vowel durations than learners for [a], [ɛ], and [ʌ], whereas learners produced [e] substantially longer than native speakers. However, the remaining vowels, [i], [ɨ], [o], and [u] showed little evidence of group differences, as their credible intervals included zero. Notably, the largest duration difference was observed for [e], where learners produced considerably longer vowels than native speakers (median = -56.79, 95% CrI [-66.20, -46.77], P(native < learner) = 1.00).

In addition, although the statistical contrasts reported in Table 4.4 focus on between-group differences across vowels, visual inspection of Figure 4.4 further shows that [e] was longer than [ɛ] within native productions.

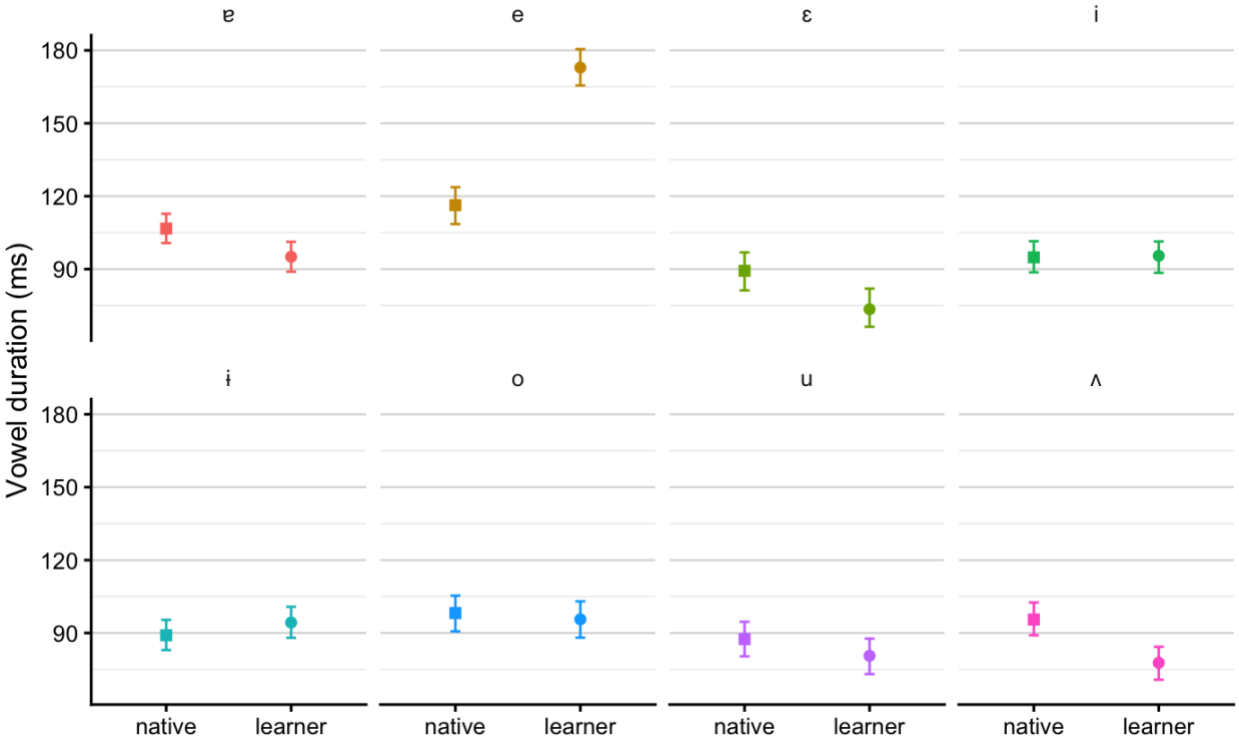
Table 4.3
Posterior contrasts of vowel durations between native and learner groups

Vowel	Median (ms)	P(native > learner)	P(native < learner)	95% CrI
a	11.63	0.99	< .001	[4.64, 18.92]
e	-56.79	0.00	1.00	[-66.20, -46.77]
ɛ	15.79	0.99	< .001	[5.81, 25.70]
i	-0.68	0.42	0.57	[-8.35, 7.08]
ɨ	-5.23	0.08	0.92	[-12.52, 2.24]
o	2.62	0.70	0.29	[-6.68, 11.76]
u	6.95	0.93	0.06	[-1.97, 16.16]
ʌ	17.85	1.00	0.00	[9.61, 26.03]

Note. Contrast values (Median) are computed as native minus learner in vowel duration (ms). Positive values indicate longer duration for native speakers than for learners, and vice versa. P(native > learner) and P(native < learner) represent the posterior probabilities that native speakers have longer or shorter values than learners respectively.

A supplementary analysis was also conducted for the comparison of learner [e] and [ɛ] relative to the merged native [e] and Table 4.4 presents posterior contrasts for F1, F2, and duration between the merged native [e] and learner [e] and [ɛ]. In F1, there was clear evidence of a difference between the merged native [e] and learner [e] (median = -0.19, 95% CrI [-0.34, -0.05], P(< 0) = .99), whereas the contrast between the merged native [e] and learner [ɛ] was not reliably different, as the credible interval included zero. In addition, learner [e] and [ɛ] were clearly differentiated from each other in F1 in that [e] is produced with higher F1 values (median = 0.16, 95% CrI [0.08, 0.23], P(>0) = 1.00). In F2, none of the contrasts showed clear evidence of differences, as all 95% credible intervals included zero.

Figure 4.4
Posterior estimates of vowel duration for native speakers and learners across eight Korean vowels



An interesting pattern emerged in duration data. All three contrasts showed clear differences. The duration of the merged native [e] was reliably shorter than that of the learner [e] (median = -52.3, 95% CrI [-69.2, -34.5], P(<0) = 1.00) and longer than that of the learner [ɛ] (median = 52.2, 95% CrI [33.7, 69.8], P(>0) = 1.00), and learner [e] was reliably longer than learner [ɛ] (median = 104.6, 95% CrI [95.8, 113.0], P(>0) = 1.00). In sum, these results suggest that the learner [e]-[ɛ] distinction was reflected most robustly in duration, with some evidence in F1 and no evidence in F2.

Table 4.4
Posterior contrasts between native merged [e] and learner [e] and [ɛ]

Dimensions	Contrast	Median	P(> 0)	P(< 0)	95% CrI
F1	K [e] - L [e]	-0.19	0.004	0.99	[-0.34, -0.05]
	K [e] - L [ɛ]	-0.03	0.31	0.68	[-0.17, 0.13]
	L [e] - L [ɛ]	0.16	1.00	0.00	[0.08, 0.23]
F2	K [e] - L [e]	-0.02	0.40	0.59	[-0.26, 0.20]
	K [e] - L [ɛ]	-0.10	0.18	0.81	[-0.33, 0.13]
	L [e] - L [ɛ]	-0.07	0.10	0.89	[-0.19, 0.04]
Duration	K [e] - L [e]	-52.3	0.00	1.00	[-69.2, -34.5]
	K [e] - L [ɛ]	52.2	1.00	0.00	[33.7, 69.8]
	L [e] - L [ɛ]	104.6	1.00	0.00	[95.8, 113.0]

Note. K = native Korean, N = Learner.

In summary, the vowel results showed that differences between native and learner productions were not consistent across vowels or acoustic features. However, the back vowel space was particularly noteworthy as native speakers produced [o] with lower F1 and [ɨ] and [u] with higher F2 than learners. While native [o] and [u] are partially merged due to sound change, learner [o] and [u] showed a clear distinction but [o] and [ʌ] did not.

4.4.2 Stops

4.4.2.1 VOT

The group contrasts in VOT are presented for phrase-initial, word-initial, and word-medial stops in Table 4.5 and visually in Figure 4.5. Posterior contrasts between laryngeal contrasts are further illustrated in Table 4.6.

For phrase-initial lenis stops, there was little evidence of a group difference in VOT. The posterior median contrast was small, and the 95% credible interval included zero. As shown in Figure 4.5, native and learner productions are similar in phrase-initial lenis VOT.

In word-initial position, group differences in VOT emerged across laryngeal categories. Learners produced longer VOTs than native speakers for lenis (median = -10.06, 95% CrI [-21.35, -0.25], P(native < learner) = .98) and aspirated (median = -28.16, 95% CrI [-46.56, -10.61], P(native < learner) = .99) stops even after adjusting for speech rate, with the larger difference for aspirated stops. However, word-initial fortis stops showed weaker evidence of a group difference, as the credible interval included zero.

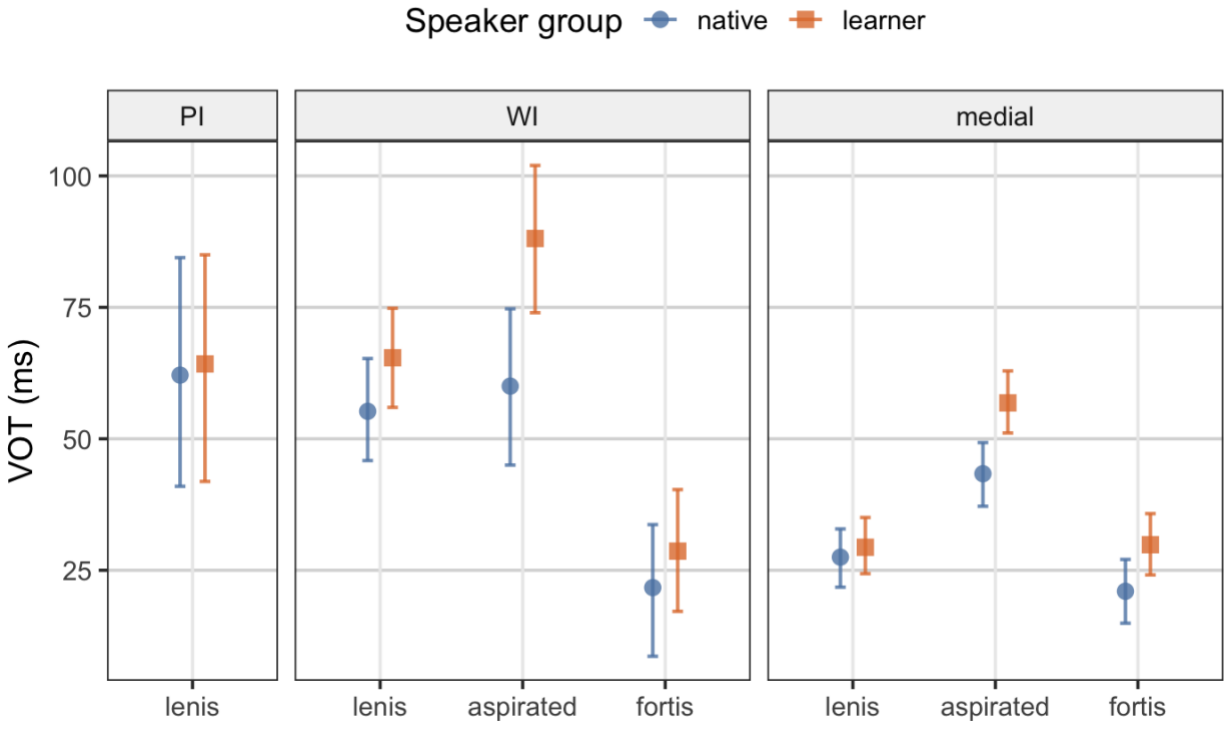
In word-medial position, there was little evidence of a group difference for medial lenis stops, as the 95% credible interval included zero. However, learners produced longer VOTs than native speakers for both medial aspirated (median = -13.46, 95% CrI [-21.52, -6.11], P(native < learner) = .99) and medial fortis (median = -8.84, 95% CrI [-16.67, -1.64], P(native < learner) = .99) stops, with credible intervals excluding zero.

Table 4.5
Posterior contrasts of VOT between native and learner groups

Stops	VOT (ms)	P(native < learner)	95% CrI
PI_lenis	-2.08	0.62	[-15.42, 11.66]
WI_lenis	-10.06	0.98	[-21.35, -0.25]
WI_aspirated	-28.16	0.99	[-46.56, -10.61]
WI_fortis	-6.94	0.82	[-22.47, 7.44]
Medial_lenis	-1.87	0.71	[-9.25, 4.65]
Medial_aspirated	-13.46	0.99	[-21.52, -6.11]
Medial_fortis	-8.84	0.99	[-16.67, -1.64]

Note. Contrast values are posterior median differences (native – learner) in milliseconds. Negative values indicate longer VOTs for learners. PI = phrase-initial, WI = word-initial, medial = word-medial.

Figure 4.5
Posterior estimates of VOT for native and learner groups across laryngeal categories and prosodic positions



The within-group contrasts between laryngeal categories in word-initial and word-medial position are presented in Table 4.6. In word-initial position, native speakers showed little evidence of a robust VOT difference between lenis and aspirated stops (difference of -4.86 ms), as the 95% credible interval included zero. However, both lenis and aspirated stops had substantially longer VOTs than fortis stops, with credible intervals excluding zero for both contrasts. By contrast, the learner group showed a different pattern. Learners' aspirated stops showed substantially longer VOTs than lenis stops (difference of -22.67 ms), and both lenis and aspirated stops had longer VOTs than fortis stops. Therefore, it is noteworthy that while native speakers indeed show the merger of lenis and aspirated in terms VOT in this position, learners' patten indicate the category distinction remaining on the basis of VOT. However, this pattern should be interpreted cautiously because group-level patterns may mask substantial individual variation in how learners realize Korean stop contrasts.

In word-medial position, native speakers showed clear VOT distinctions among all three laryngeal categories. Aspirated stops had longer VOTs than lenis and fortis stops, and lenis stops had longer VOTs than fortis stops, and aspirated stops had the longest VOT overall. In the learner group, aspirated stops showed clearly longer VOTs than both lenis and fortis stops. However, unlike the native group, learners showed little evidence of a robust difference between lenis and fortis stops in medial position, as the credible interval for this contrast included zero.

Table 4.6
Posterior contrasts of VOT between laryngeal contrasts within native and learner groups

Group	Contrast	VOT (ms)	P(>0)	P(<0)	95% CrI
Native (WI)	lenis - aspirated	-4.86	0.23	0.76	[-17.60, 8.47]
	lenis - fortis	33.51	1.00	0.00	[23.90, 43.51]
	aspirated - fortis	38.34	1.00	0.00	[24.90, 52.05]
Learner (WI)	lenis - aspirated	-22.67	0.00	0.99	[-35.90, -10.47]
	lenis - fortis	36.79	1.00	0.00	[27.00, 45.77]
	aspirated - fortis	59.42	1.00	0.00	[46.50, 72.35]
Native (medial)	lenis - aspirated	-15.81	0.00	1.00	[-20.15, -11.71]
	lenis - fortis	6.49	0.99	0.00	[2.50, 10.23]
	aspirated - fortis	22.33	1.00	0.00	[17.43, 27.06]
Learner (medial)	lenis - aspirated	-27.44	0.00	1.00	[-31.59, -23.39]
	lenis - fortis	-0.43	0.40	0.59	[-3.98, 3.37]
	aspirated - fortis	26.99	1.00	0.00	[22.25, 31.84]

Note. WI = Word-initial, medial = word-medial.

4.4.2.2 F0

The group contrasts in F0 are presented for phrase-initial, word-initial, and word-medial stops in Table 4.7 and visually in Figure 4.6. Posterior contrasts between laryngeal contrasts are further illustrated in Table 4.8.

For phrase-initial lenis stops, there was little evidence of a group difference in F0. The posterior median contrast was small and the 95% credible interval included zero (median = -8.64, 95% CrI [-25.60, 10.40], P(<0) = .83), which indicates substantial overlap between native and learner productions.

In word-initial position, the pattern differed across laryngeal categories. For lenis and fortis stops, there was little evidence of a robust group difference. However, native speakers

produced higher F0 values than learners for aspirated stops, with a posterior median difference of 31.26 Hz and a 95% credible interval excluding zero (median = 31.26, 95% CrI [8.10, 54.70], P(>0) = .99).

In word-medial position, there was little evidence of robust group differences in F0 for any laryngeal category.

Table 4.7
Posterior contrasts of F0 between native and learner groups

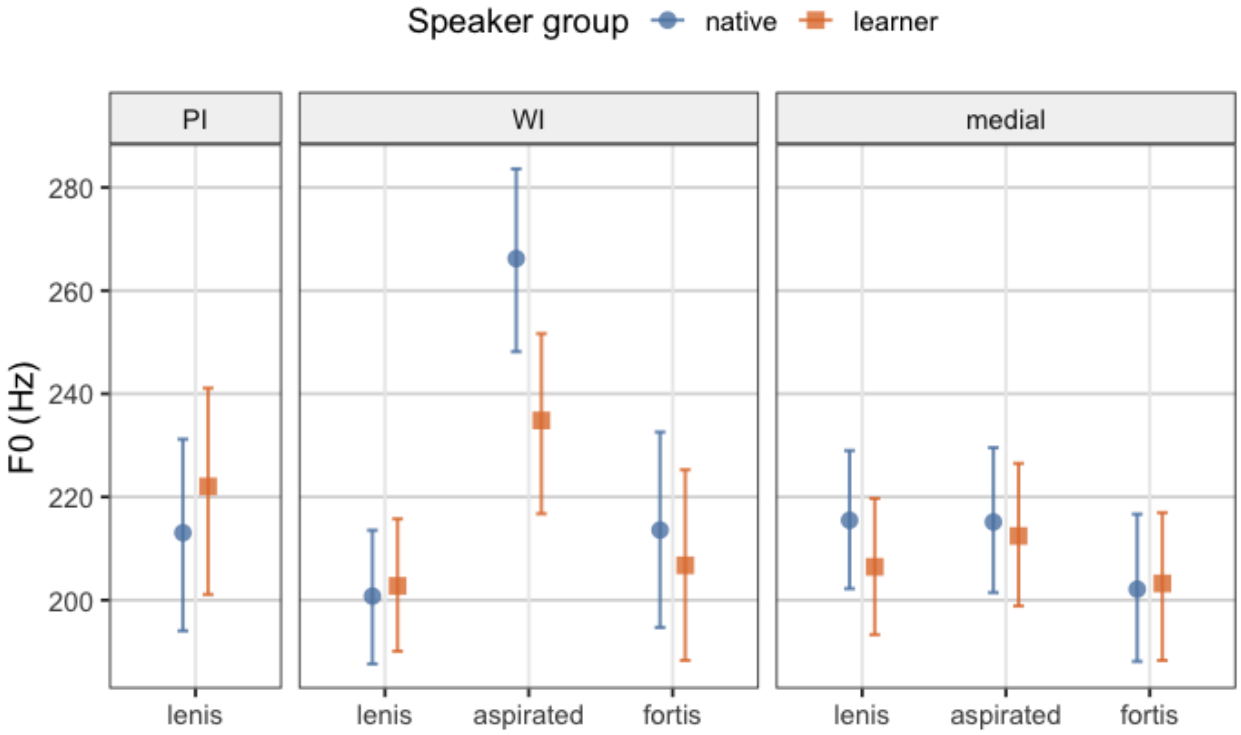
Stops	F0 (Hz)	P(>0)	P(<0)	95% CrI
PI_lenis	-8.64	0.17	0.83	[-25.60, 10.40]
WI_lenis	-1.98	0.40	0.49	[-18.40, 15.10]
WI_aspirated	31.26	0.99	0.00	[8.10, 54.70]
WI_fortis	6.85	0.71	0.28	[-18.70, 32.50]
Medial_lenis	9.16	0.85	0.14	[-8.93, 26.40]
Medial_aspirated	2.68	0.61	0.38	[-16.71, 20.60]
Medial_fortis	-1.18	0.45	0.54	[-19.59, 18.3]

Note. Contrast values are posterior median differences (native – learner) in Hertz. Positive values indicate higher F0s for native speakers, and negative values indicate higher F0s for learners. PI = phrase-initial, WI = word-initial, medial = word-medial.

The within-group contrasts between laryngeal categories in word-initial and word-medial positions are presented in Table 4.8. In word-initial position, both native speakers and learners showed clear F0 differences between aspirated and the other two laryngeal stops. For native speakers, lenis stops were produced with substantially lower F0 than aspirated stops (median = -65.43, 95% CrI [-77.80, -54.06], P(<0) = 1.00) and aspirated stops were produced with substantially higher F0 than fortis stops (median = 52.44, 95% CrI [32.80, 70.41], P(>0) = 1.00), with credible intervals excluding zero and posterior probabilities being 1 in both contrasts. However, there was little evidence of a robust difference between lenis and fortis stops, as the

credible interval for this contrast included zero. Learners showed a similar pattern to that of native speakers, but weaker as the F0 for aspirates stops were lower than that of native speakers. Here again, the pattern of learner group is not consistent with the recent language change, in which F0 is a more robust cue to distinguish lenis and aspirated stops.

Figure 4.6
Posterior estimates of F0 for native and learner groups across laryngeal categories and prosodic positions



In word-medial position, the pattern was slightly different between the two groups. For native speakers, there was little evidence of F0 difference between lenis and aspirated stops, but both lenis and aspirated stops showed higher F0 values than fortis stops, with credible intervals excluding zero for these two contrasts. It indicates that fortis stops in word-medial positions were associated with relatively lower F0 values in native speech. In contrast, in the learner group, lenis

stops showed lower F0 than aspirated stops (median = -5.94, 95% CrI [-12.11, -0.07], P(<0) = .98) and aspirated stops showed higher F0 than fortis stops (median = 9.16, 95% CrI [1.31, 17.40], P(>0) = .98), whereas the lenis and fortis contrast showed little evidence of a robust difference.

Table 4.8
Posterior contrasts of F0 between laryngeal contrasts within native and learner groups

Group	Contrast	F0 (Hz)	P(>0)	P(<0)	95% CrI
Native (WI)	lenis - aspirated	-65.43	0.00	1.00	[-77.80, -54.06]
	lenis - fortis	-12.90	0.05	0.94	[-29.50, 3.32]
	aspirated - fortis	52.44	1.00	0.00	[32.80, 70.41]
Learner (WI)	lenis - aspirated	-32.02	0.00	1.00	[-43.70, -20.70]
	lenis - fortis	-3.97	0.29	0.70	[-19.40, 11.97]
	aspirated - fortis	28.02	0.99	0.001	[10.20, 46.45]
Native (medial)	lenis - aspirated	0.42	0.55	0.44	[-5.66, 6.78]
	lenis - fortis	13.35	0.99	0.00	[6.53, 20.16]
	aspirated - fortis	12.92	0.99	0.00	[4.80, 21.22]
Learner (medial)	lenis - aspirated	-5.94	0.02	0.98	[-12.11, -0.07]
	lenis - fortis	3.17	0.83	0.16	[-3.23, 9.83]
	aspirated - fortis	9.16	0.98	0.01	[1.31, 17.40]

Note. WI = Word-initial, medial = word-medial.

4.4.3 Phonological realization patterns

The results are reported in terms of recurrent patterns observed across multiple tokens and speakers, rather than single errors produced by individual learners. Five types of learners' phonological realization patterns are summarized in Table 4.9. As shown in the table, failure to delete [h] in non-initial onset position or hyperarticulation of [h] was the most frequent pattern as its pattern occurred 167 times across all 11 learners. This was followed by failure to apply

resyllabification across syllable boundaries, which occurred in 69 tokens across all 11 learners. Nasal or liquid assimilation failure was the third most frequent pattern with 31 tokens across 8 learners and deletion of coda [l] before onset [l] and [h]-induced aspiration were less frequent, which occurred in 17 and 14 tokens across 7 learners respectively.

Table 4.9
Phonological realization patterns in learner productions

Pattern (failure)	Tokens	Speakers	Examples
Resyllabification	69	11	[kap̄.a.t*a] instead of [ka.p^{h}a.t*a]
Nasal or liquid assimilation	31	8	[ʌp̄.mu] instead of [ʌm:u]
[h]-induced aspiration	14	7	[kɨp̄.han] instead of [kɨ.p^{h}an]
[h] deletion in non-initial onset or hyperarticulation of [h]	167	11	[kjʌl.hon] instead of [kjʌ.ron]
Coda [l] before onset [l]	17	7	[mo.re] instead of [mol:e]

Note. Speaker counts are based on 11 learners.

Although this was not the main focus of this chapter, two items showed variable realizations in both native and learner productions. As shown in Table 4.10, 부엌에서 /pueok-eseo/ and 닭을 /talg-eul/ showed two realizations. In both items, one realization corresponded to the prescriptive standard pronunciation specified by the National Institute of Korean Language (National Institute of Korean Language, 2026) and the other reflecting a non-prescriptive variant. Native and learner speakers showed broadly similar distributions across the two variants, although the speaker distributions were not identical across the two items. However, there was a tendency for speakers who produced the prescriptive realization in [puʌk^{h}esʌ] to produce the prescriptive realization in [talkɨl] as well.

Table 4.10
Variation in phonological realizations in native and learner productions

Unit	Realization	Speakers
부엌에서 /pueok-eseo/	[pu.ʌ.k^{h}e.sʌ]*	K03, K05, K06, K07, K10, L01, L02, L03, L04, L05, L07
	[pu.ʌ.ke.sʌ]	K01, K02, K08, K09, K11, L06, L08, L09, L10, L11
닭을 /talg-eul/	[tal.kɨl]*	K03, K07, K09, K10, L01, L02, L03, L05, L07, L09
	[ta.kɨl]	K01, K02, K05, K06, K08, K11, L04, L06, L08, L10, L11

Note. *Note*. K = native Korean, L = Learner. Data from K04 were unavailable due to data loss. An asterisk (*) indicates the prescribed standard variant.

4.5 Discussion

4.5.1 Vowels

First, the vowel results suggest that native and learner productions were broadly similar overall. Under the present model, robust group differences in vowel quality were limited to normalized F1 for [o] and [ʌ]. Specifically, native speakers produced [o] with lower F1 values than learners and [ʌ] with higher F1 values than learners, indicating a clearer separation between these vowels in native productions. By contrast, no robust group differences were observed in normalized F2. These findings suggest that the most reliable native-learner difference in the vowel space was the [o]-[ʌ] contrast, rather than a broader pattern of group differences across vowels.

This pattern is particularly relevant to the interpretation of [o]. Native speakers produced [o] with lower F1 values than learners, which means native [o] is higher than learner [o] in the back vowel space. Although the data in this chapter are not intended to demonstrate sound

change, this pattern is broadly compatible with previous studies which have reported that contemporary Seoul Korean [o] is raising (J.-I. Han & Kang, 2013; Y. Lee, 2025). For the purposes of the current research questions, the more important point is that learners did not fully approximate the native [o]-[ʌ] contrast in terms of vowel height and [o]-raising in contemporary Seoul Korean.

With respect to [u], native speakers showed higher F2 values than learners, suggesting a tendency of a more fronted realization. However, this difference was not statistically reliable, indicating that native and learner productions were broadly similar in the present data. Given previous reports of [u]-fronting in Seoul Korean (Y. Kang, 2014), one possible interpretation is that learners may have approximated this pattern. At the same time, because [u]-fronting is also observed in American English (Clopper et al., 2019; Koops, 2010), the similarity between native and learner productions of [u] may instead reflect L1 influence rather than approximation to the contemporary Seoul Korean change. However, the present data do not clearly show which explanation is more likely.

In addition, Y. Kang (2014) reported that [ɨ] may be fronting as part of a chain shift-like change with [u] and [o]. In the present data, an additional pattern was that both native and learner productions placed [ɨ] lower in the vowel space than expected for a high vowel, with realizations closer to the front-mid region, closer to [e] and [ɛ] in the present data. Because this chapter is not designed to examine sound change directly, these results cannot be taken as evidence of sound change. One possible explanation is that F1 values for high vowels, including [ɨ], may vary by syllable position, with lower F1 values in word-initial syllables than in word-medial or word-final syllables (Y. Kang, 2014). If so, the distribution of [ɨ] tokens across syllable positions may have influenced the mean F1 values for native productions. Further investigation is therefore

needed to determine whether the present pattern reflects an ongoing chain shift, or positional effects in the dataset.

With regard to the [e]-[ɛ] contrast, given previous reports that [e] and [ɛ] are merged in contemporary Seoul Korean (Y. Kang, 2014; Y. Lee & Cho, 2021), a vowel quality difference in the native productions was not expected and the two categories were observed to overlap (Left panel in Figure 4.2). A supplementary analysis was therefore conducted to examine whether learners' [e] and [ɛ] appear similarly to the merged native category. In F1, merged native [e] differed from learner [e] but not clearly from learner [ɛ], and in F2 none of the contrasts showed clear evidence of difference. By contrast, duration provided the clearest separation among the categories: merged native [e] was shorter than learner [e], longer than learner [ɛ], and learner [e] was substantially longer than learner [ɛ]. In other words, the learner [e]-[ɛ] distinction was the most robust in duration. This suggests that learners may be maintaining a distinction between [e] and [ɛ] with temporal properties than spectral properties when vowel quality is not strongly differentiated. However, this duration distinction should be interpreted cautiously. In the present dataset, [e] may have occurred frequently as the locative particle -에, often in word-final position, and this positional or morphological effects may have contributed to lengthening in both native and learner productions. Further investigation is therefore needed to determine whether this pattern reflects learners' greater reliance on temporal cues, or is attributable to particular morphological components or syllable positions.

4.5.2 Stops

The stop results suggest that American English learners broadly patterned differently to produce the Korean three-way laryngeal contrast from native Korean speakers and that different

pattern showed which acoustic cue American English use more primarily when producing the Korean stops compared to native Korean speakers.

The first evidence can be found in the word-initial VOT pattern. Learners' VOTs for lenis and aspirated stops were both longer than those of native speakers, but the difference was larger for aspirated stops. Within groups, the native speakers did not show a robust VOT difference between lenis and aspirated stops, while fortis stops remained clearly separated from both through shorter VOT. This pattern is consistent with previous research of contemporary Seoul Korean, in which the VOT contrast between lenis and aspirated stops has been neutralized and F0 has become more important in maintaining the distinction between them (K.-H. Kang & Guion, 2006; H. Lee, 2016b; H. Lee et al., 2020; Silva, 2006). In contrast, the learners maintained a clearer VOT distinction across the three-way stop contrast, with aspirated stops showing the longest VOT and fortis stops the shortest. This suggests that the learners had not fully approximated the cue neutralization between lenis and aspirated stops in contemporary Seoul Korean, but instead rely more on VOT to distinguish the three-way stops. Such pattern is compatible with previous findings that English speaking learners of Korean tend to weight VOT more strongly than F0 in perception and production of the three-way stop (H. Kim, 2023; M.-R. C. Kim & Lotto, 2002). However, this interpretation should again be treated with caution, as the group-level pattern may not fully capture the extent of individual variability among learners (J.-I. Han & Kim, 2014; Holliday, 2015).

The word-medial VOT pattern showed a slightly different pattern. Native speakers showed shorter VOT for lenis stops than for aspirated stops, while lenis remained longer than fortis. This pattern is consistent with the positional features of Korean lenis stops, which are susceptible to intervocalic weakening and often become fully voiced in word-medial position

(Choi, 2020). In the present dataset, fully voiced medial lenis stops were excluded because VOT was not measurable. Therefore, the results reflect the positional effect on lenis VOT in word-medial position. The learners also showed shortening of medial lenis VOT, but learners' lenis and fortis were not robustly distinguished. This suggests that learners may have patterned similarly in that lenis VOT was shortened in intervocalic position, but not in a native-like way.

Regarding F0, learners patterned similarly to native speakers in that aspirated stops had higher F0 than lenis and fortis stops in word-initial position. In that sense, the overall direction of the contrast was native-like. However, the difference was smaller in learner speech because learners' aspirated stops had lower F0 than native aspirated stops. This suggests that native speakers relied more strongly on F0 to distinguish aspirated from lenis stops, which are not clearly separated by VOT. Learners, by contrast, appeared to maintain that distinction more through VOT than solely through F0. This result is consistent with earlier work suggesting that learners whose L1 relies on VOT to distinguish a two-way stop contrast tend to rely more heavily on VOT across three-way stop contrast (Holliday, 2015; H. Kim, 2023; M.-R. C. Kim & Lotto, 2002).

The group difference in stop contrast is therefore better understood as a difference in cue weighting than as a categorical difference in whether F0 was used at all. Previous work on the Seoul Korean stop contrast has shown that younger speakers rely more on F0, especially for the lenis-aspirated contrast, as the VOT contrast between those two has weakened over time (H. Lee, 2016b; H. Lee et al., 2020; Silva, 2006). The results in this chapter suggest that learners did use F0, but relied on this cue less than native speakers did, while VOT remained a more primary cue for them. In other words, rather than fully acquiring the reweighted cue structure of Seoul

Korean, the learners appear to have implemented the three-way contrast through a cue-weighting pattern of their own, which differs from the native Seoul Korean pattern (Chang, 2009).

The native word-initial F0 pattern is also noteworthy on its own. In many studies of Korean stops, fortis and aspirated stops have been grouped together as high F0 categories relative to lenis stops, or fortis stops show intermediate-high F0, falling between aspirated and lenis stops (Cho et al., 2002; K.-H. Kang & Guion, 2006; H. Kim & Jongman, 2022; M. Kim, 2004). In the present data, however, aspirated stops showed clearly higher F0 than both lenis and fortis stops, while lenis and fortis did not differ reliably from each other. One possible interpretation is that, as lenis VOT has lengthened and become less distinct from aspirated VOT, F0 has become increasingly necessary for distinguishing lenis from aspirated, while fortis remains robustly identifiable through short VOT. Under such circumstances, fortis may no longer need to maintain as high F0 as aspirated in order to remain contrastive. Kim (2023) recently reported that native Korean listeners still use F0 as a primary cue to distinguishing lenis stops from both aspirated and fortis stops in perception. Thus, the present results may reflect an ongoing shift in the cue structure of Seoul Korean production. However, the present data alone are not sufficient to conclude that fortis F0 is undergoing sound change, and further work is needed to determine whether this pattern reflects a change in Seoul Korean or a specific pattern observed in the present dataset.

The medial F0 pattern is more difficult to interpret. Native speakers showed similar F0 values for lenis and aspirated stops, while F0 for fortis was lower than both. In contrast, learners showed a pattern closer to aspirated being higher than lenis and fortis, which resembles their word-initial organization more than the native pattern in word-medial position. It is difficult to determine why fortis stops were especially low in medial position in the native data, even after

excluding all sentence-final -ta forms from the dataset, since these are typically subject to tensification and pitch lowering due to sentence-final intonation. One possibility is that positional and prosodic conditioning in non-initial contexts affects the realization of F0 differently from the word-initial position that has been the main focus of most previous studies. Because the literature on the Seoul Korean stops has concentrated primarily on word-initial stops, the present medial pattern may reflect contextual effects that remain less well understood. For the purposes of the current chapter, the more important point is that the learners did not reproduce the native medial F0 pattern either, which further supports the conclusion that they had not yet acquired the positional distribution of cues in a native-like way.

4.5.3 Phonological realization patterns

The phonological realization patterns suggest that foreign-accented Korean in the learner data may be influenced not only by segmental deviations, but also by difficulty in producing connected speech phonology. In this chapter, five recurrent patterns were observed in the learner data. One possible explanation for four of the five patterns (e.g., failure of resyllabification, assimilation, [h]-induced aspiration, and [h] deletion in non-initial position, or hyperarticulation of [h]) is the influence of Hangul orthography, since the target spoken forms of these cases differ from segment-by-segment reading of the written form. This interpretation is consistent with research showing that orthographic forms can influence L2 pronunciation during reading, but also learners' phonological representations more generally (Bassetti, 2008, 2024; Hayes-Harb & Barrios, 2021). It has also been suggested that, when orthographic input is available, instructed learners may both benefit from and experience interference from orthography in L2 phonology acquisition (Hayes-Harb & Barrios, 2021).

At the same time, some patterns may reflect incomplete acquisition of specific phonological rules rather than orthography alone. In particular, [h]-induced aspiration may require explicit instruction that a coda stop followed by [h] is realized as an aspirated stop. This type of phonological rule may be less likely to be acquired automatically by learners, especially in reading tasks. Recent work on Hangul reading instruction found that explicit instruction was more effective than implicit learning for some phonological rules such as obstruent nasal assimilation and aspiration, while resyllabification and [h]-deletion were relatively easier to learn (Sim & Kim, 2025).

The deletion of coda [l] before onset [l] may be due to limited knowledge of the allophonic behavior of Korean [l] in coda position rather than orthographic influence alone. Overall, the recurrent patterns in Table 4.9 suggest that learners' deviant phonological patterns may arise from orthographic influence, incomplete acquisition of phonological rules, and limited knowledge of Koran allophonic structure.

Although this was not the main focus of this chapter, the variable realizations of 부엌에서 /pueok-eseo/ and 닭을 /talg-eul/ also deserve brief discussion. Both native and learner speakers produced a prescriptive form and a non-prescriptive variant and the latter was not limited to learner speech. The speaker distributions were also somewhat mixed across the two items in that speakers who produced the prescriptive form in /puʌk^{h}esʌ/ did not necessarily also produce the prescriptive form in /talkɨl/. This suggests that the choice between the two realizations was not consistent even within speakers. For native speakers, the non-prescriptive forms may have resulted from a process in which coda neutralization was applied before the following morpheme (i.e., locative particle *-esʌ* and objective marker *-ɨl*) was attached. However, even when learners produced the same surface forms, it is difficult to determine whether learners applied the same

phonological process. These items therefore raise an interesting question for the following chapter, whether the two variants would be evaluated differently in accentedness judgments.

4.6 Conclusion

This chapter examined how American English learners of Korean differed from native Seoul Korean speakers in the acoustic and phonological realization of Korean speech. Overall, the results showed that learner productions differed from native productions at both the segmental and connected-speech levels.

First, the vowel results indicated that native and learner productions were largely similar overall. The clearest group differences in vowel quality were limited to normalized F1 for [o] and [ʌ], suggesting that the most noticeable difference between native and learner vowel productions is the height distinction between these two vowels rather than across the vowel system as a whole. In addition, the supplementary analysis of [e] and [ɛ] indicated that the learner distinction between these vowels was reflected more clearly in temporal than in spectral features.

Second, the stop results showed that learners produced the Korean three-way laryngeal contrast with a different cue structure from native speakers. Specifically, learners relied more on VOT than native speakers did, especially in word-initial position, where they maintained a clearer VOT distinction between lenis and aspirated stops than native speakers. Although learners also used F0 in a similar pattern to that of native speakers, they did not use the cue to the same extent or in the same positional pattern as native speakers. These patterns suggest that the learners had acquired the three-way stop contrast using a cue-weighting strategy different from that of native speakers of contemporary Seoul Korean.

Finally, the phonological realization patterns demonstrated that learner accent in connected speech may not be explained solely in terms of segmental deviations. Recurrent failures of resyllabification, assimilation, [h]-related alternations, and producing coda [l] patterns suggest that learners also experience difficulty in producing Korean phonological patterns in a native-like way. These patterns can be attributable to orthographic influence, incomplete acquisition of specific phonological rules, and limited knowledge of Korean allophones.

The findings of this chapter show that American English-accented Korean differs from native Seoul Korean not only in the acoustic realization of vowels and stops, but also in the phonological realization. At the same time, the present chapter focused on group-level differences and did not examine individual learner variation in detail. It is possible that learners' productions vary depending on factors such as Korean proficiency, age of acquisition, or amount of Korean experience. Therefore, further research is needed to examine how such learner-specific factors influence acoustic and phonological pattern.

Another limitation is that the analysis was based on read sentence productions. Read speech made it possible to control segmental and prosodic contexts, but it may not fully reflect how learners produce Korean in spontaneous interaction. Future studies could compare read and spontaneous speech to determine whether the acoustic and phonological patterns observed here also emerge in more naturalistic speech.

Overall, this chapter provides an important foundation for the following chapter, which examines how these acoustic and phonological features of learner speech may contribute to accentedness judgments and language attitude evaluations.

V. SPEECH FEATURES ASSOCIATED WITH ACCENTEDNESS JUDGMENTS AND LANGUAGE ATTITUDES TOWARD FOREIGN-ACCENTED KOREAN

5.1 Introduction

Chapters 2 and 3 of this dissertation explored how social expectations influence intelligibility, accentedness judgments, and language attitudes toward foreign-accented Korean speech. The findings showed that mismatches between acoustic social information affected accentedness judgments, and that different types of social information influenced language attitudes. At the same time, speech perception is not based solely on social information, but is also grounded in the acoustic properties of the speech signal. Therefore, this chapter provides an exploratory investigation of the speech features that are associated with accentedness judgments and language attitude evaluations of foreign-accented Korean speech.

Previous research on the speech features underlying perceived foreign accent has shown mixed findings, both in terms of which features are associated with accentedness and the relative importance of those features (Crowther et al., 2018; Saito et al., 2016; Sereno et al., 2016; Trofimovich & Isaacs, 2012; Winters & O'Brien, 2013). Across studies, accentedness has been associated with a wide range of speech properties, including segmental features such as vowel and consonant production (Gunter et al., 2020; McCullough, 2011; Munro, 1993; Sereno et al., 2016; Wayland, 1997), syllable structure-related patterns such as epenthesis and deletion (Anderson-Hsieh et al., 1992), prosodic features such as rhythm, stress, intonation, and pitch (Idemaru et al., 2019; Mareüil & Vieru-Dimulescu, 2006), and fluency-related features such as speech rate, pause frequency and duration (Polyanskaya et al., 2017; Trofimovich & Baker,

2006). However, studies have yielded mixed findings on which features are most strongly associated with perceived foreign accent, and the relative contribution of these features appears to vary across target languages, learner populations, and research methods. As Idemaru et al. (2019) noted, various findings on the acoustic correlates of accentedness likely reflect both methodological differences across studies and language-specific features of the first and second languages.

At the segmental level, earlier work has shown that deviations in vowel and consonant production can influence listeners' judgments of foreign accent. For example, Munro (1993) demonstrated that acoustic features of English vowels produced by Arabic speakers were associated with accentedness ratings. In addition, McCullough (2011) examined multiple foreign-accented varieties of English and showed that vowel quality and VOT of stop consonants significantly contributed to listeners' ratings of foreign accent, especially Korean- and Mandarin-accented English. More recently, Jung and Holliday (2026) extended this line of research to Korean by examining foreign accent perception in isolated Korean words. They found that fricative- and fortis-initial words, and F0 and F1 values that deviated from L1 Korean norms predicted stronger foreign accent ratings. These findings suggest that certain segmental cues may be relevant across different L1 backgrounds.

At the same time, other studies have shown that suprasegmental features may also play a major role in accent perception, and in some cases may be even more influential than segmental features. Idemaru and colleagues (2019) found that prosodic features related to pitch accent and intonation in Japanese were strong predictors of accent ratings. They further suggested that the acoustic correlates of perceived accent may vary depending on the relationship between learners' first language and the target language. In their study of L2 Japanese, tonal patterns were more

predictive of accent ratings for L1 English learners than for L1 Mandarin learners, possibly because L1 Mandarin learners, who speak a tonal language, were better able to approximate Japanese pitch accents. This suggests that the influence of acoustic cues may differ across learner groups depending on cross-linguistic similarities and differences.

This perspective of cross-linguistic differences is important in the present study. As discussed in Chapter 4, Korean and English differ in several phonetic and phonological respects, and some of these differences were reflected in the learner productions. In particular, Korean stop contrasts rely on both VOT and F0, whereas English relies primarily on VOT in a two-way stop contrast. Chapter 4 showed that learner productions differed from native Korean speech in some aspects of VOT and F0. For example, English learners of Korean appeared to rely more on VOT in distinguishing lenis and aspirated stops, in which F0 is a primary cue to distinguishing this contrast in native Korean. While Chapter 4 documented acoustic differences between learner and native speaker patterns, whether these differences lead to perceptual consequences, such as perceived accentedness, remains an empirical question.

With respect to vowels, there were not many robust differences between learners and native speakers, although some vowel categories such as F1 of [o] or [ʌ] were significantly different. However, the present chapter goes beyond group comparisons by examining how specific acoustic dimensions, such as F1 and F2 values, related to accentedness judgments. This would make it possible to capture more fine-grained information about whether particular ranges of vowel quality contribute to listeners' perception of foreign accent. In addition, Korean has some phonological rules such as resyllabification, assimilation, and [h]-induced aspiration, which may be difficult for learners to acquire naturally with Hangul orthography alone. Such segmental and phonological deviations may therefore be noticeable to Korean listeners. At the

same time, suprasegmental fluency-related features such as speech rate and pausing may also influence accent perception more broadly (Kang, 2010; Trofimovich & Baker, 2006).

In addition to accentedness judgments, the present chapter also examines Korean listeners' language attitudes toward English-accented Korean speakers. Accordingly, learners' production patterns are related to listeners' evaluations of status (e.g., intelligence, competence) and solidarity (e.g., trustworthiness, in-group membership), the two dimensions that are central in the conceptualization of language attitude, as discussed in Chapter 3. The present chapter also examined how Korean listeners' judgments of comprehensibility are related to learners' production patterns. Although comprehensibility is not technically a social evaluative dimension such as status and solidarity, it was assessed together with status and solidarity using the same questionnaire format (see Appendix C), and Chapter 3 further showed that comprehensibility was positively correlated with both status and solidarity, and therefore it is examined here as part of a broader set of listener evaluations.

With respect to comprehensibility, a substantial body of research has shown that listeners' perceived ease or difficulty of understanding is associated with a range of linguistic and speech features, including speech rate, prosody, and grammatical accuracy (Isbell et al., 2024; Munro & Derwing, 2001; Shimada, 2024; Trofimovich et al., 2022; Trofimovich & Isaacs, 2012). By contrast, relatively few studies have directly examined which specific acoustic features shape social evaluations such as status and solidarity. Nevertheless, existing work suggests that accent plays an important role in such evaluations (Carlson & McHenry, 2006; Fuertes et al., 2012; Lev-Ari & Keysar, 2010; Subtirelu, 2015). For example, Fuertes and colleagues (2012) found that standard accents are generally evaluated more positively in terms of status, solidarity, and dynamism (e.g., liveliness and activity), while Lev-Ari and Keysar (2010)

showed that foreign-accented speech can reduce perceived credibility, possibly because processing difficulty biases listeners' judgments. However, because these studies focus primarily on accentedness or accent categories rather than the acoustic properties that underlie these evaluations, it remains unclear whether more fine-grained acoustic features also contribute to social evaluations.

Previous studies have shown that Korean listeners' accentedness and comprehensibility judgments are related to linguistic properties of L2 Korean speech. Isbell and colleagues (2024) examined spontaneous monologic speech and showed that various variables such as phonological, lexical, morphosyntactic errors, and fluency-related variables predicted accentedness and comprehensibility, while Jung and Holliday (2026) examined isolated Korean words produced by L1 speakers of English, Chinese, Japanese and Vietamese, and found that foreign accent ratings were strongly predicted by F0 and F1 values. The present chapter builds on these studies by examining controlled sentence-level American English-accented Korean speech, allowing for a more targeted analysis of acoustic features while maintaining a connected speech context. Accordingly, this chapter investigates whether particular acoustic features of American English-accented Korean are associated not only accentedness and comprehensibility, but also with listeners' evaluations of status and solidarity. In doing so, it extends prior research on Korean foreign accent perception to the acoustic correlates underlying both perceptual and social-evaluative dimensions.

5.1.1 Research questions

(1) **Research Question 1**: Which speech-related features are associated with accentedness judgments of American English-accented Korean?

(2) **Research Question 2**: How are speech-related features associated with listeners' evaluations of status, solidarity, and comprehensibility?

5.2 Methodology

5.2.1 Participants

A total of 25 native Korean listeners (13 men, 12 women) living in South Korea participated in the study. Participants were recruited through word of mouth, the researcher's personal network of friends and acquaintances, and Everytime, a social networking and academic platform widely used by university students in South Korea. All participants reported no history of hearing or language disorders. They were born and raised in South Korea and had little to no experience living abroad before the age of 18. Participants' age ranged from 20 to 52 years ($M = 30.68$, $SD = 7.52$).

5.2.2 Audio stimuli

Among 60 sentences recorded by both native Korean speakers and American English learners of Korean (see Section 4.2), 10 sentences were selected for the present study. These sentences were chosen to ensure the inclusion of word-initial aspirated and fortis stops, as tokens of these categories were less frequent than those of lenis stops. In addition, to examine the effect of phonological patterns on accentedness judgments, sentences containing recurrent phonological patterns observed in learner productions were included. These 10 sentences are indicated in bold in Appendix A.

All productions of the 10 selected sentences by the 11 learner speakers were included. In addition, productions of the same sentences by three native Korean speakers were included as

anchors to provide a native baseline for accentedness judgments. The total stimulus set consisted of 140 tokens (14 speakers x 10 sentences). All the audio stimuli were normalized for intensity to 67 dB using a Praat script (version 6.4.47; Boersma & Weenink, 2025), which was lower than originally planned 70 dB due to a clipping issue.

5.2.3 Procedure

The experiment was conducted online using Gorilla Experiment Builder (https://gorilla.sc), and participants received the experiment link directly from the researcher. They were asked to complete the experiment in a quiet place using their own headphones or earphones. The entire experiment took 20 to 30 minutes. Upon clicking the link, participants first read the consent form and indicated their consent before proceeding. Gorilla then automatically presented a pop-up window to ensure that the audio was functioning properly.

The experiment then moved on to the accentedness judgment task. Audio stimuli were presented in random order, and a 9-point Likert scale appeared in the center of the screen, ranging from 0 (sounding like a native Korean speaker) to 8 (sounding heavily accented) and participants rated the degree of accentedness of each stimulus.

Participants then indicated which features of the audio led them to perceive a foreign accent, and completed a survey consisting of status, solidarity, and comprehensibility questions. Each construct was assessed using four items on a 7-point Likert scale, ranging from 1 (not at all) to 7 (very much). An example of the survey screen is shown in Figure 5.1. In this part of the experiment, each page presented one speaker along with 10 sentences in a row. Participants could answer the survey items (see Appendix C) either while listening to the sentences or after all of the stimuli had finished playing. The order of speakers was also randomized. Finally,

participants completed a demographic and language background questionnaire, which included questions about their Korean dialect, second language background, experience living abroad, interactions with foreigners speaking Korean, and exposure to media featuring foreigners speaking Korean. They were also asked about their views on the importance of accent and pronunciation in second language learning.

Figure 5.1
Example screen from the language attitude survey



5.2.4 Acoustic measurement

Acoustic predictors were derived from the measurements described in Chapter 4. Based on the acoustic and phonological analyses presented in Chapter 4, the present chapter included segmental features such as vowel quality (F1 and F2) and stop consonant VOT and F0, as well as

fluency-related features, including speech rate, vowel duration, pause duration, pause frequency and phonological realization patterns as predictors. Pauses were coded as either between-word or within word pauses, and any pause longer than 100 ms was counted (Idemaru et al., 2019; Trofimovich & Baker, 2006). These variables were examined in relation to accentedness judgments and language attitudes. Further details on the measurement and coding procedures are provided in Chapter 4 (see Section 4.2.3).

5.3 Analysis

5.3.1 Acoustic features associated with accentedness judgments

The relationship between acoustic features and accentedness judgments was examined using generalized additive models (GAMs) fitted in R (R Core Team, 2026) with the `gam` function in the *mgcv* package (Wood, 2025). GAMs were adopted because the relationships between acoustic predictors and accent judgments were not assumed to be strictly linear. For variables such as VOT, F0 and formant values, perception may be relatively stable within a more native-like range but change more sharply once values move away from that range. GAMs are suited well to modeling such potentially non-linear patterns. Because accentedness ratings were collected on a 9-point Likert scale, ordinal GAMs were fitted using an ordered categorical family (`ocat`).

Since accentedness ratings were obtained separately for each sentence, predictors were aligned at the speaker-sentence level, but not all predictors were available for all observations. For example, segment-specific acoustic measurements such as VOT, F0, F1, and F2 were only available when the relevant stop category or vowel was present in a given sentence, which

resulted in missing values for some predictors across items. For this reason, separate models were fitted for different sets of predictors rather than including all predictors in a single model.

An initial model included predictors that were available across all stimulus items. This model included speech rate, vowel duration as continuous predictors and the presence or absence of deviant phonological patterns as binary variables. Continuous predictors were modeled with smooth terms so that their relationships with accentedness judgments could show potentially non-linear relationships. Random effect smooths were included for *Participant* and *Sentence* and a *Speaker*-specific factor smooth was also included for *Speech Rate* to allow the effect of *Speech Rate* to vary across speaker.

Additional models were fitted to examine other variables separately. For example, pause-related variables such as pause duration between words (*bp_dur*) and the number of pauses between words (*bp_n*) were tested in the same model. Likewise, phrase-initial lenis stop measures were examined in one model, and word-initial and word-medial stop measures (VOT and F0) were examined in other models. Vowel measures (F1, F2, and duration of each vowel) were also examined in separate models.

All models were estimated using restricted maximum likelihood (REML). Model convergence was assessed using the `gam.check()` function and all models showed full convergence.

5.3.2 Acoustic features associated with language attitudes

Before modeling the language attitude ratings, the internal consistency reliability of the status, solidarity, and comprehensibility scales was assessed using Cronbach's alpha (Cronbach, 1951). Alpha coefficients were calculated in R (R Core Team, 2026) using the `alpha` function

in the *psych* package (Revelle, 2026). This step was used to check whether the items within each attitudinal dimension showed consistency to justify combining them into composite scores. For the comprehensibility scale, the item "effortful to understand" was reverse-coded so that higher values indicated easier understanding. Following conventional guidelines, alpha values between .70 and .95 were considered acceptable (Tavakol & Dennick, 2011).

Acoustic predictors were based on the measurements described in Chapter 4. The relationship between acoustic features and language attitude ratings was examined using generalized additive models (GAMs) fitted in R with the `gam` function in the *mgcv* package (Wood, 2025) for the same reason stated earlier. Because language attitude ratings were collected on a 7-point Likert scale, ordinal GAMs were fitted using an ordered categorical family (`ocat`).

The language attitude data included three constructs, status, solidarity, and comprehensibility. These three language attitude dimensions were modeled together in one model, with *Construct* included as a fixed effect and dimension-specific smooths fitted for continuous predictors using `by = Construct`. This made it possible to examine the effect of each predictor on each language attitude dimension.

Unlike the models for accentedness ratings, the language attitude models did not involve missing variables in the dataset because language attitude ratings were collected only at the speaker level. However, all predictors were not included in a single model because the total number of acoustic variables was large, and including all of them with construct-specific smooths would have resulted in an overly complex model. Therefore, separate models were fitted for different sets of predictors.

One model included global speech features such as speech rate, vowel duration and binary variables such as the presence or absence of deviant phonological patterns. Other models were fitted for pause-related variables such as pause duration and pause frequency, and additional models were fitted for phrase-initial lenis stop measures, word-initial and word-medial stop measures (VOT and F0), and vowel measures (F1, F2, and duration of each vowel).

All continuous variables were modeled with smooth terms. Random effect smooths for *Participant* and *Question* (i.e., language attitude survey items) were included, and a Speaker-specific factor smooth term for *Speech Rate* was added in the models to allow speaker-level variation in the effect of *Speech Rate*. All models were estimated using restricted maximum likelihood (REML) and model convergence was assessed using the `gam.check()` function.

5.3.3 Listeners' subjective responses

After completing accentedness ratings and language attitude survey, participants were asked to provide open-ended responses about features that led them to perceive foreign accents in speech. These responses were analyzed qualitatively to identify recurring patterns across listeners.

All responses were first reviewed and coded into thematic categories (Anderson-Hsieh et al., 1992; Barkhuizen, 2019). The categories include fluency-related features, segmental features (consonants and vowels), phonological patterns, and prosodic features. Fluency-related features included comments on speech rate, pausing, and overall smoothness of delivery (e.g., "slow", "reading word by word"). Segmental features included references to consonant and vowel productions. Phonological patterns included deviations from target forms such as

resyllabification and [h]-induced aspiration. Prosodic features included comments on intonation or pitch-related features (e.g., “unnatural intonation”, “inaccurate pitch”).

After coding, the responses were summarized by category in order to examine which types of features were most salient to listeners when describing foreign-accented Korean.

5.4 Results

5.4.1 Accentedness judgments

5.4.1.1 Fluency-related features

Fluency-related features showed partial associations with accentedness judgments, indicating that native Korean listeners were sensitive to some temporal features of speech when evaluating foreign accents. Among the non-smooth predictors (Table 5.1), phonological errors significantly increased accentedness ratings, whereas hyperarticulated [h] reduced perceived accentedness. In contrast, neither between-word nor within-word pause frequency showed a reliable effect.

Table 5.1
Effects of non-smooth predictors on accentedness judgments

Predictors	Estimate	SE	p-value	Prediction
Phonological error	0.33	0.10	0.001**	Accentedness ↑
Hyperarticulated [h]	-0.29	0.12	0.01*	Accentedness ↓
Between-word pause frequency	-0.02	0.09	0.75	No effect
Within-word pause frequency	0.41	0.22	0.06	No effect

Note. * p < .05, ** p < .01, *** p <.001. Accentedness ↑ indicates a positive relationship with accentedness ratings, and Accentedness ↓ indicates a negative relationship with accentedness ratings.

The smooth terms (Table 5.2) further showed effects of continuous temporal variables. Speech rate showed an overall negative relationship with accentedness in that slower speech was associated with higher accentedness ratings. Although the estimated smooth term was linear (edf =1), the predicted accentedness values exhibited a non-linear pattern with a steeper decrease at lower speech rates and a leveling-off at higher speech rate (Figure 5.2). This pattern likely reflects speaker-specific variation in the speech rate effect rather than a non-linear global effect.

Vowel duration showed a significant positive linear relationship with accentedness, indicating that longer vowel durations were associated with higher accentedness ratings. Together, these results suggest that slower speech contributes to stronger perceptions of foreign accent.

Pause duration measures did not show reliable effects. However, Figure 5.2 shows that utterances containing pauses were generally associated with relatively high predicted accentedness ratings. This pattern suggests that while differences in pause duration and frequency itself did not significantly affect accentedness, pauses tended to occur in speech that was judged as relatively highly accented.

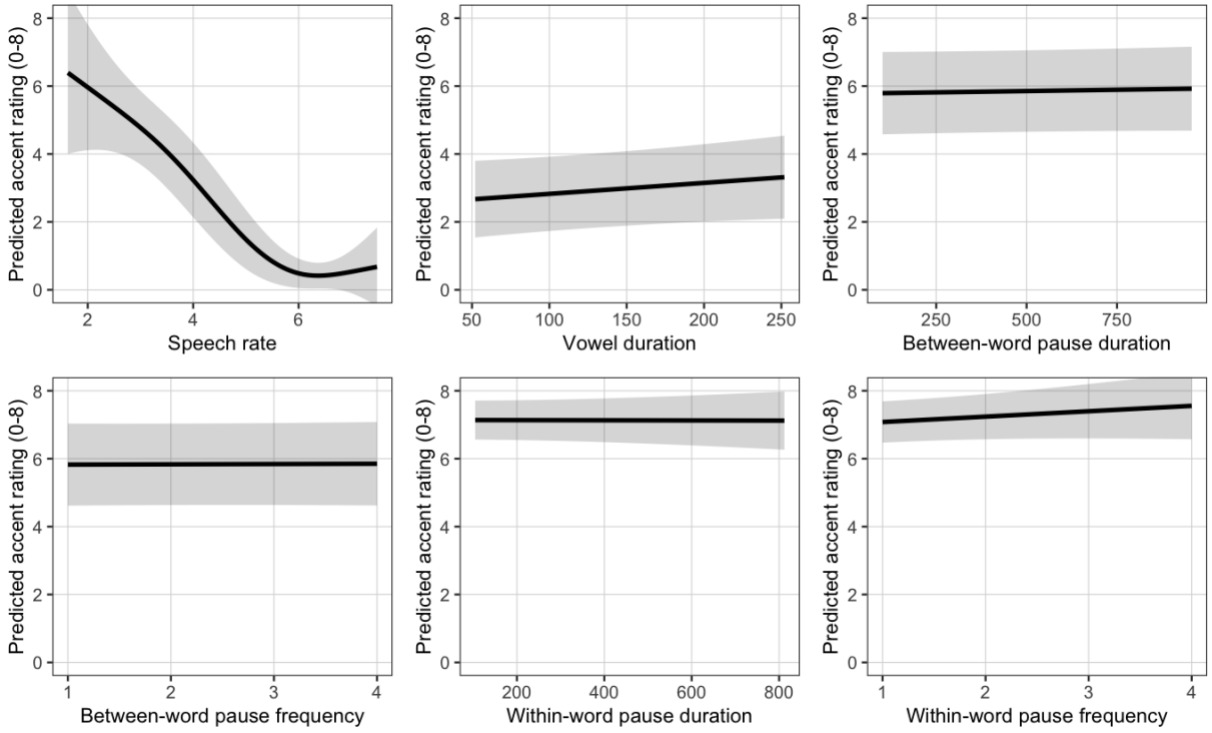
Overall, accentedness ratings were primarily associated with phonological errors, hyperarticulated [h] and some fluency-related features such as speech rate and vowel duration, whereas pause-related measures did not show significant effects.

Table 5.2
Effects of fluency-related predictors on accentedness judgments

Predictors	edf	Chi.sq	p-value	Prediction
Speech rate	1.00	34.75	< .001***	Accentedness ↓
Vowel duration	1.00	6.23	0.01*	Accentedness ↑
Between-word pause duration	1.00	0.02	0.89	No effect
Within-word pause duration	1.00	0.044	0.83	No effect

Note. edf = estimated degrees of freedom. edf = 1.00 indicates a linear effect and values greater than 1 indicate a non-linear effect.

Figure 5.2
GAM smooths for fluency-related predictors on accentedness judgments



5.4.1.2 Stop features

Stop-related acoustic features such as VOT and F0 showed some associations with accentedness across laryngeal categories and prosodic positions (Table 5.3). In phrase-initial position, lenis VOT showed a significant positive linear relationship with accentedness, such that longer VOT was judged as more accented. In word-initial position, aspirated VOT showed a marginal positive linear effect on accentedness, also with longer VOT being judged as more accented.

Lenis VOT in the word-initial position was in general rated with lower accent scores; however, there was a significant non-linear effect. It appears VOT longer than around 150 ms was rated increasingly more accented, albeit very slightly. F0 associated with stops did not show relationship with accent ratings, except for fortis F0 in the word-medial position. Higher fortis F0 in this position was rated more accented although the effect was small.

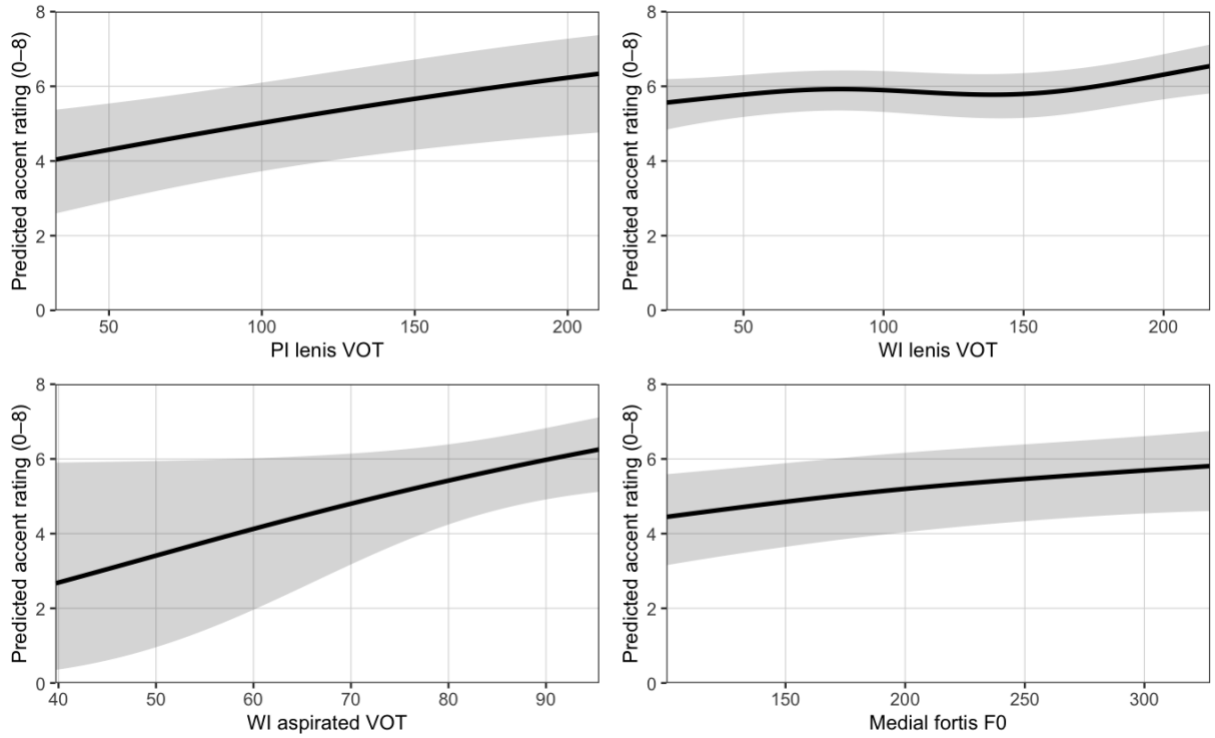
Figure 5.3 shows that word-initial lenis VOT and word-medial fortis F0 were relatively flat for the predicted accent ratings across their ranges. By contrast, phrase-initial lenis VOT and word-initial aspirated VOT showed more upward trends, with predicted accentedness ratings increasing more steeply as VOTs increased.

Table 5.3
Effects of stop-related predictors on accentedness judgments

Predictors	edf	Chi.sq	p-value	Prediction
PI lenis VOT	1.00	7.00	< .01**	Accentedness ↑
PI lenis F0	1.41	1.30	0.56	No effect
WI lenis VOT	3.12	12.36	0.04*	Non-linear effect
WI lenis F0	1.00	1.42	0.23	No effect
WI aspirated VOT	1.00	3.80	0.05	Marginal effect ↑
WI aspirated F0	1.00	0.66	0.41	No effect
WI fortis VOT	2.99	6.88	0.13	No effect
WI fortis F0	1.00	0.96	0.32	No effect
Medial lenis VOT	2.54	6.21	0.10	No effect
Medial lenis F0	1.86	1.67	0.52	No effect
Medial aspirated VOT	1.00	0.56	0.45	No effect
Medial aspirated F0	1.00	0.23	0.63	No effect
Medial fortis VOT	2.95	5.04	0.18	No effect
Medial fortis F0	1.33	13.63	< .001***	Accentedness ↑

Note. PI = phrase-initial, WI = word-initial, and Medial = word-medial.

Figure 5.3
GAM smooths for significant stop-related predictors on accentedness judgments



5.4.1.3 Vowel features

Vowel-related acoustic features (F1 and F2) also showed significant associations with accentedness across vowel categories and spectral dimensions (Table 5.4). Because the number of vowel tokens and vowel-specific measures was large, only the predictors that showed statistical significance are reported here.

Several vowel formant measures showed significant relationships with accentedness judgments. F2, which corresponds to the front-back dimension of vowel quality, was significantly associated with accentedness for mostly front and mid vowels, [i], [e], [ɛ], [ɨ], and a back vowel [ʌ]. For the two vowels [i] and [ʌ], both F1, corresponding to the high-low dimension, and F2 showed significant effects. As shown in Figure 5.4, predicted accent ratings

were relatively raised for several front-vowels, particularly F1 and F2 of [i], F2 of [e], and F2 of [ɛ]. This pattern indicates that listeners were especially sensitive to formant variation in the fron-vowel space, although the direction of the effects differed.

More specifically, a more fronted tongue position (higher F2) of [i] and a more backed position (lower F2) of [e] were judged as more accented, suggesting that a fronted [i] needed to be backed and [e] needed to be fronted to sound less accented. For [ɛ], a vowel that is merged with [e], F2 showed non-linear relationship. Particularly, very high F2 values above approximately 2000 Hz were associated with increasingly higher accentedness ratings.

While accented rating for [ɨ] and [ʌ] were generally lower, [ɨ] was rated more accented when F2 value was either higher or lower than around 1800 Hz. For [ʌ], higher accentedness ratings were associated with lower F1 and lower F2, which corresponds to a higher and more backed position of it, was perceived as more accented.

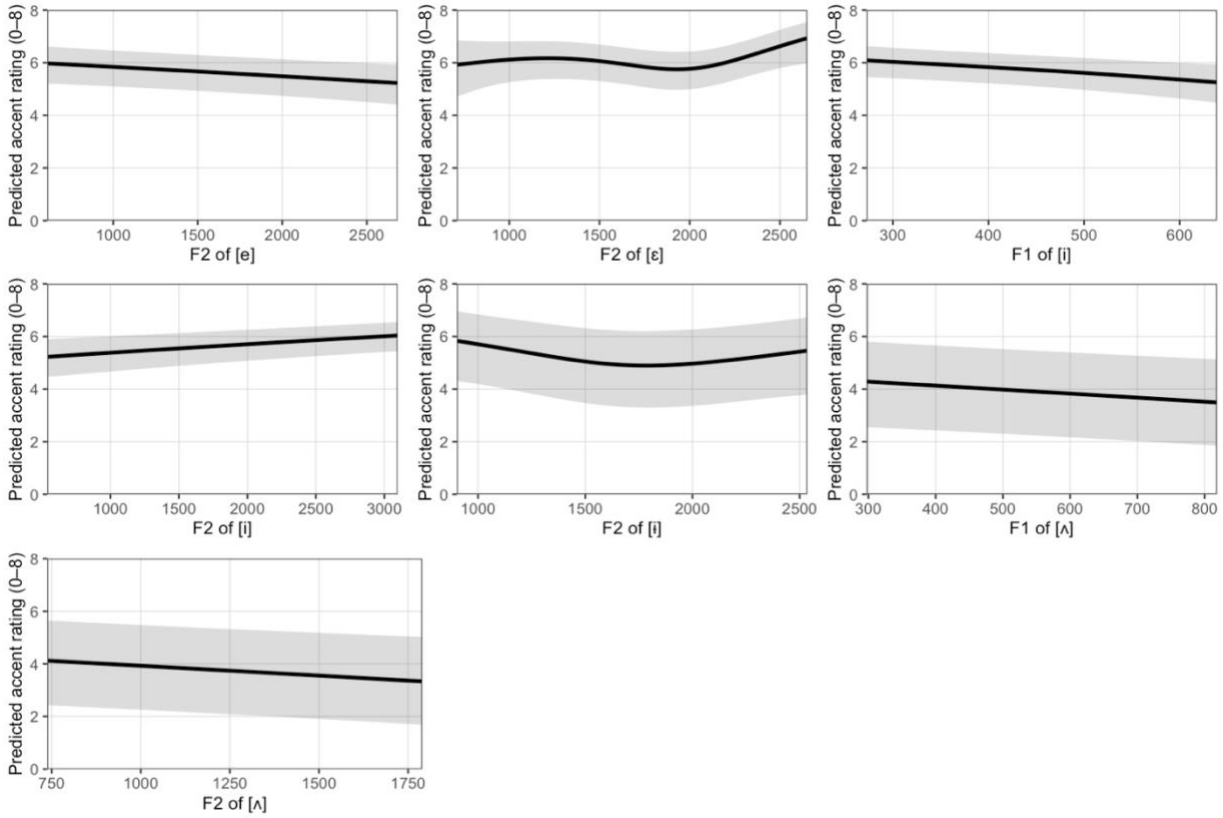
Overall, the effects of vowel-related predictors appeared to vary across vowel categories and formants with both linear and non-linear relationships with accent judgments.

Table 5.4
Effects of vowel-related predictors on accentedness judgments

Predictors	**edf**	**Chi.sq**	**p-value**	**Prediction**
F2 of [e]	1.00	7.00	< .01**	Accentedness ↓
F2 of [ɛ]	3.18	9.18	0.03*	Non-linear effect
F1 of [i]	1.00	10.05	0.01**	Accentedness ↓
F2 of [i]	1.00	14.28	0.001**	Accentedness ↑
F2 of [ɨ]	2.64	14.19	< .01**	Non-linear effect
F1 of [ʌ]	1.00	4.54	0.03*	Accentedness ↓
F2 of [ʌ]	1.00	4.62	0.03*	Accentedness ↓

In summary, accentedness ratings increased with phonological errors, slower speech rate, and longer vowel duration, while hyperarticulated [h] reduced perceived accentedness. At the segmental level, VOT was the primary stop-related cue along with word-medial fortis F0, and vowel formant variation mostly in front vowels, but also in mid and back vowels, contributed to accentedness judgments.

Figure 5.4
GAM smooths for significant vowel-related predictors on accentedness judgments



5.4.2 Language attitudes

Prior to reporting the results, internal consistency for the language attitude scales was high, with Cronback's alpha values of .97 for status, .94 for solidarity, and .92 for comprehensibility.

5.4.2.1 Fluency-related features

Fluency-related features showed different patterns across the three language attitude constructs (status, solidarity, and comprehensibility; Table 5.5 and 5.6).

Status ratings (e.g., judgments of the speaker's education and competence) were affected only by hyperarticulated [h]. Specifically, it was associated with lower status ratings, which suggests that speech containing this feature sounded less educated, proficient, and competent. Solidarity ratings (e.g., judgments of the speaker's friendliness and warmth) were significantly reduced by higher within-word pause frequency. Comprehensibility ratings were significantly reduced by both phonological errors and hyperarticulated [h]. However, between-word pause frequency did not significantly affect any of the three constructs.

Table 5.5
Effects of non-smooth predictors on language attitude ratings

Construct	Predictors	Estimate	SE	p-value	Prediction
Status	Phonological error	-0.14	0.71	0.84	No effect
	Hyperarticulated [h]	-1.88	0.54	< .001***	Rating ↓
	BW pause frequency	-0.18	0.26	0.48	No effect
	WW pause frequency	-1.10	0.75	0.14	No effect
Solidarity	Phonological error	-0.89	0.89	0.31	No effect
	Hyperarticulated [h]	-0.97	0.56	0.08	No effect
	BW pause frequency	-0.01	0.25	0.96	No effect
	WW pause frequency	-2.58	0.75	< .001***	Rating ↓
Comprehensibility	Phonological error	-1.57	0.63	0.01*	Rating ↓
	Hyperarticulated [h]	-1.39	0.54	< .01**	Rating ↓
	BW pause frequency	-0.30	0.51	0.56	No effect
	WW pause frequency	-3.22	1.89	0.09	No effect

Note. BW = between-word, WW = within-word. Rating ↑ indicates a positive relationship with language attitude ratings, and Rating ↓ indicates a negative relationship with language attitude ratings.

The smooth terms (Table 5.6) further showed the effects of fluency-related predictors on language attitude ratings. For status, only vowel duration showed a significant effect with longer vowel duration reducing status ratings.

For solidarity, vowel duration also showed a significant effect, but the relationship was non-linear. Solidarity ratings appeared to be highest when vowel was around 100 ms, whereas both shorter and longer vowel durations were associated with lower solidarity ratings. Within-word pause duration also significantly increased solidarity ratings.

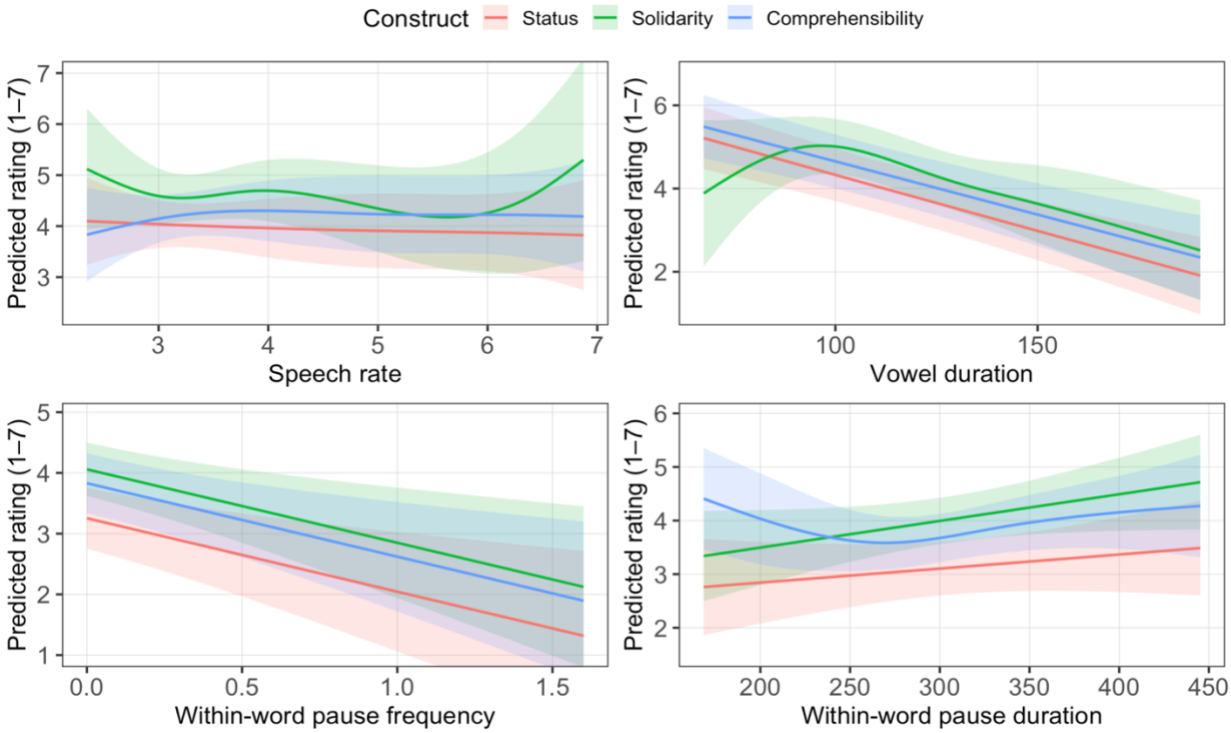
For comprehensibility, both speech rate and vowel duration were significant predictors, Faster speech rate was associated with higher comprehensibility ratings, whereas longer vowel duration reduced comprehensibility ratings. This suggests that a faster speaking tempo enhanced comprehensibility, while vowel-specific slower tempo reduced perceived comprehensibility.

In sum, vowel duration was the most consistent predictor across all three constructs, although the direction and shape of its effect varied by construct. Speech rate was specific to comprehensibility and within-word pause duration was specific to solidarity.

Table 5.6
Effects of fluency-related predictors on language attitude ratings

Construct	Predictors	edf	Chi.sq	p-value	Prediction
Status	Speech rate	2.39	4.64	0.15	No effect
	Vowel duration	1.00	20.75	< .001***	Rating ↓
	BW pause duration	1.00	2.59	0.11	No effect
	WW pause duration	1.00	0.11	0.74	No effect
Solidarity	Speech rate	1.00	0.03	0.85	No effect
	Vowel duration	2.35	23.42	< .001***	Non-linear effect
	BW pause duration	1.00	0.35	0.55	No effect
	WW pause duration	1.00	4.58	0.03*	Rating ↑
Comprehensibility	Speech rate	1.00	4.24	< .001***	Rating ↑
	Vowel duration	1.00	18.09	< .001***	Rating ↓
	BW pause duration	1.43	4.74	0.09	No effect
	WW pause duration	1.47	0.25	0.70	No effect

Figure 5.5
GAM smooths for fluency-related predictors on language attitude ratings



5.4.2.2 Stop features

In the subsection on stop-related features for language attitudes, only statistically significant effects are reported due to the large number of predictors (Table 5.7). Stop-related predictors showed significant associations with all three language attitude dimensions, although the patterns differed across constructs, prosodic positions, and laryngeal categories.

Interestingly, status ratings were significantly associated with F0s, but not with VOTs. Although some of these effects were non-linear, lower word-initial lenis F0 and higher word-initial aspirated F0 (red lines in the WI lenis F0 and WI aspirated F0 panels) were associated with higher status ratings. In word-medial position, higher lenis F0 was associated with higher status ratings, whereas higher fortis F0 was associated with lower status ratings.

For solidarity, several predictors including both F0 and VOT were significant. Lenis stop, in particular, was heavily associated with judgment of solidarity. Whereas the influence of lenis VOT and F0 in the phrase-initial position on solidarity (green lines in the PI lenis VOT and PI lenis F0 panels) were relatively small, whereas the corresponding effects in word-initial position (green lines in the WI lenis VOT and WI lenis F0 panels) were more pronounced. In particular, longer word-initial lenis VOT was associated with higher solidarity rating. Word-initial lenis F0 also showed a non-linear pattern, with F0 values around 190 Hz receiving relatively high solidarity ratings, while both lower and higher F0 values were associated with lower ratings.

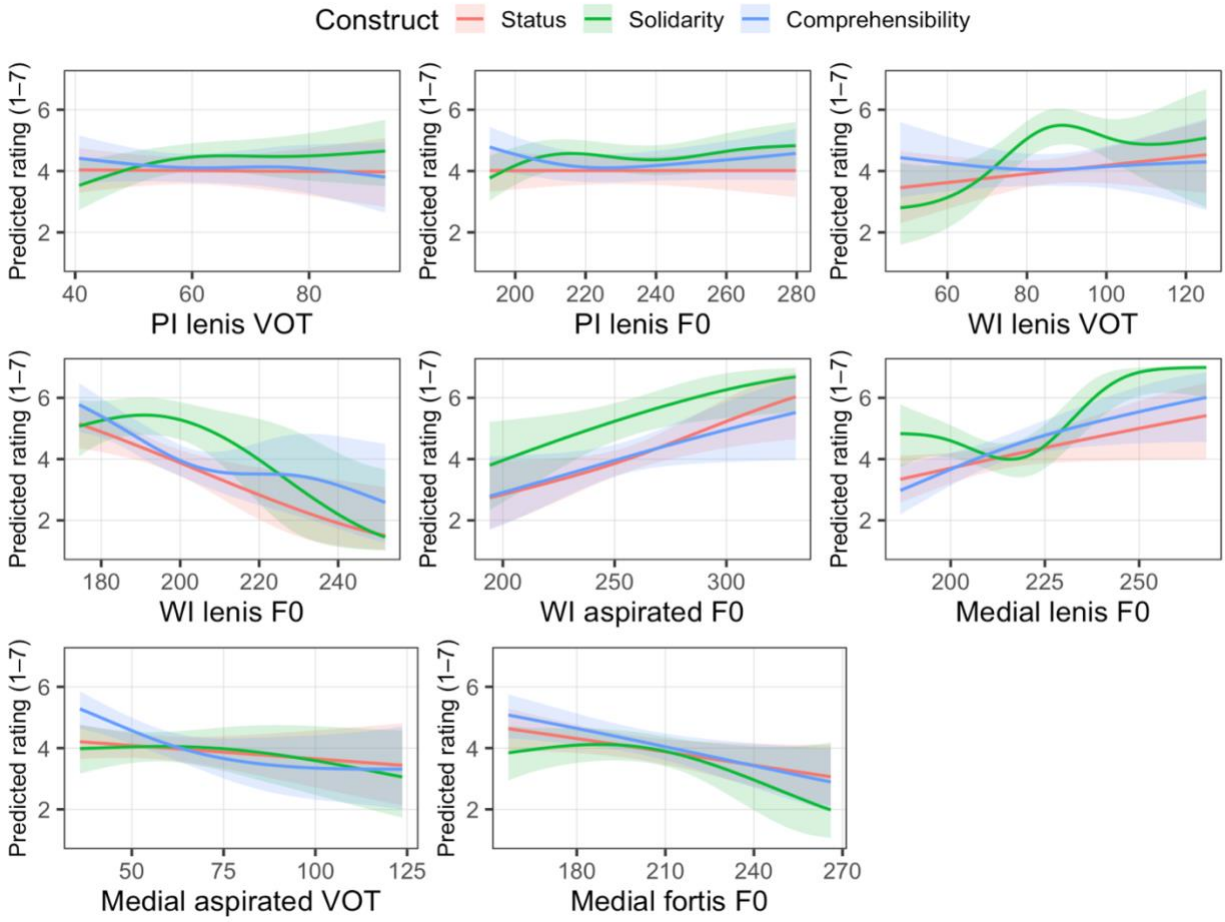
Similar to status, comprehensibility ratings were mainly associated with F0s. Lenis F0 was significant across all three positions. Its effect was relatively small in phrase-initial position (blue line in the PI lenis F0 panel), but more pronounced in word-initial and word-medial positions (blue lines in the WI lenis F0 and Medial lenis F0 panels). Lower word-initial lenis F0 and higher word-medial lenis F0 were associated with higher comprehensibility ratings. In addition, higher word-initial aspirated F0 was associated with higher comprehensibility, whereas higher word-medial fortis F0 was associated with lower comprehensibility. Word-medial aspirated VOT also showed a significant non-linear effect with shorter VOTs generally associated with higher comprehensibility ratings.

Table 5.7
Effects of stop-related predictors on language attitude ratings

Construct	Predictors	edf	Chi.sq	p-value	Prediction
Status	WI lenis F0	1.00	9.02	< .01**	Non-linear effect
	WI aspirated F0	1.87	7.81	0.02*	Non-linear effect
	Medial lenis F0	1.00	3.91	0.04*	Rating ↑
	Medial fortis F0	1.00	3.88	0.04*	Rating ↓
Solidarity	PI lenis VOT	3.00	17.47	0.001**	Non-linear effect
	PI lenis F0	3.59	17.75	0.01*	Non-linear effect
	WI lenis VOT	3.68	14.93	< .01**	Non-linear effect
	WI lenis F0	2.53	11.03	< .01**	Non-linear effect
	WI aspirated F0	1.00	5.88	0.01*	Rating ↑
	Medial lenis F0	3.32	48.57	< .001***	Non-linear effect
Comprehensibility	PI lenis F0	2.96	18.48	< .001***	Non-linear effect
	WI lenis F0	2.89	15.09	< .01**	Non-linear effect
	WI aspirated F0	1.00	3.99	0.04*	Rating ↑
	Medial lenis F0	1.88	11.05	< .01**	Non-linear effect
	Medial aspirated VOT	2.37	17.36	< .01**	Non-linear effect
	Medial fortis F0	1.00	6.70	< .01**	Rating ↓

It is also notable that word-initial lenis F0, word-initial aspirated F0, and word-medial F0 were significant across all three dimensions. In sum, the status dimension was associated primarily with F0s rather than VOTs. The solidarity dimension was generally associated with lenis-related measures, with an additional effect of word-initial aspirated F0. Comprehensibility was associated mainly with F0-related measures, with an additional effect of word-medial aspirated VOT. Overall, stop-related features showed both linear and non-linear associations across the three language attitude dimensions.

Figure 5.6
GAM smooths for stop-related predictors on language attitude ratings



5.4.2.3 Vowel features

Vowel-related acoustic features (F1 and F2) also showed significant associations with language attitude ratings across the three constructs (Table 5.8). Because the number of vowel-specific measures was large, only the predictors that showed statistical significance are reported here.

For status, significant effects were observed for F1 of [ɛ] and [ɨ], and F2 of [a], [i] and [ʌ]. Among these, higher F2 of [a], which is a more fronted realization, and higher F1 of [ɛ],

which is in a lower tongue position, was associated with lower status ratings. In contrast, higher F2 of [ɪ], which corresponds to a more fronted position, increased status ratings. The effects of F1 of [ɨ] and F2 of [ʌ] showed non-linear relationships; however, Figure 5.7 suggests that a lower tongue position for [ɨ] and a more backed position for [ʌ] tended to reduce status ratings.

For solidarity, a larger number of vowel segments showed significant effects. Higher F1 values for [a], [ɪ] and [u], corresponding to lower tongue positions, were associated with lower solidarity ratings. By contrast, more fronted realization of [ɪ] increased solidarity ratings. Several other predictors showed non-linear relationships, including F1 of [e] and [ɨ], and F2 of [a], [ɨ], [u], [o] and [ʌ].

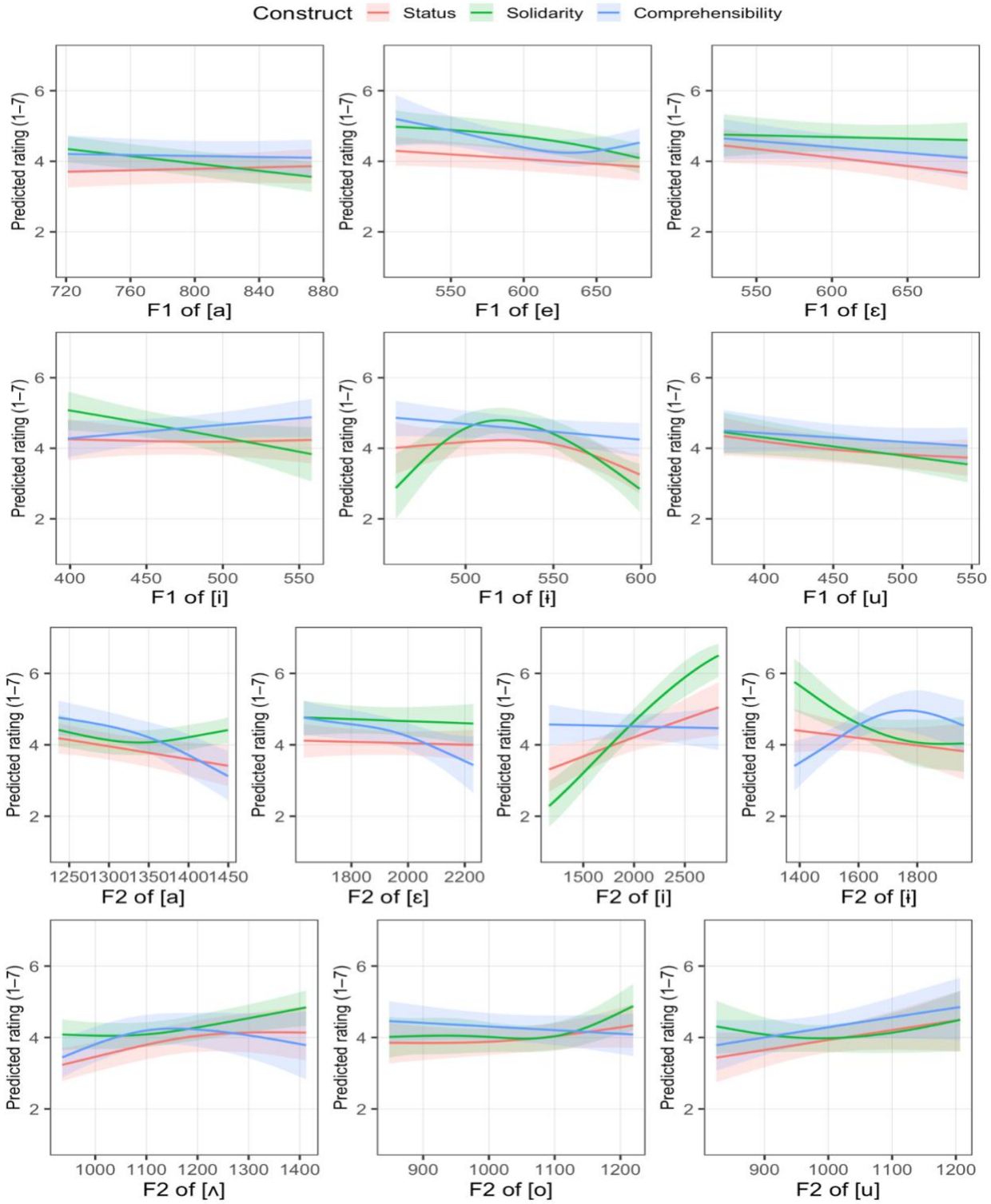
For comprehensibility, all significant vowel predictors showed non-linear effects including F1 of [e] and F2 of [a], [ɛ], [ɨ], and [ʌ]. Although these effects were non-linear, the patterns shown in Figure 5.7 suggest that comprehensibility ratings tended to decrease when [a] and [ɛ] were produced with more fronted tongue positions and when [ʌ] was produced with a higher tongue position.

In sum, vowel-related predictors showed both linear and non-linear relationships with language attitude ratings, and the effects varied across constructs, vowel categories, and spectral dimensions. Compared with status and comprehensibility, solidarity ratings were associated with more vowel segments.

Table 5.8
Effects of vowel-related predictors on language attitude ratings

Construct	Predictors	edf	Chi.sq	p-value	Prediction
Status	F2 of [a]	1.00	5.48	0.01*	Rating ↓
	F1 of [ɛ]	1.00	4.17	0.04*	Rating ↓
	F2 of [i]	1.00	7.44	< .01**	Rating ↑
	F1 of [ɨ]	2.42	9.12	0.01*	Non-linear effect
	F2 of [ʌ]	2.10	14.73	< .001***	Non-linear effect
Solidarity	F1 of [a]	1.00	15.48	< .001***	Rating ↓
	F2 of [a]	2.29	15.49	0.03*	Non-linear effect
	F1 of [e]	1.63	14.90	< .01**	Non-linear effect
	F1 of [i]	1.00	4.91	0.02*	Rating ↓
	F2 of [i]	1.00	40.19	< .001***	Rating ↑
	F1 of [ɨ]	3.30	43.31	< .001***	Non-linear effect
	F2 of [ɨ]	2.27	9.90	0.01*	Non-linear effect
	F1 of [u]	1.00	4.91	0.02*	Rating ↓
	F2 of [u]	2.44	7.82	0.03*	Non-linear effect
	F2 of [o]	2.69	10.15	0.01*	Non-linear effect
	F2 of [ʌ]	2.16	12.57	< .01**	Non-linear effect
Comprehensibility	F2 of [a]	2.01	24.97	< .001***	Non-linear effect
	F1 of [e]	2.65	10.68	0.01*	Non-linear effect
	F2 of [ɛ]	1.99	8.51	0.02*	Non-linear effect
	F2 of [ɨ]	2.79	16.18	0.001**	Non-linear effect
	F2 of [ʌ]	2.50	17.14	< .001***	Non-linear effect

Figure 5.7
GAM smooths for vowel-related predictors on language attitude ratings



5.4.3 Listeners' subjective responses

The qualitative analysis of open-ended responses showed that listeners most frequently referred to fluency-related and segmental features when describing the aspects of speech that led them to perceive it as foreign-accented. Because individual listeners could mention more than one feature in their responses, the total number of responses across categories does not equal the total number of listeners. As shown in Table 5.9, segmental features were mentioned by 12 out of the 25 listeners, which is the most frequently reported category. This was followed by fluency-related features, which were mentioned by 10 listeners. Phonological patterns were reported by 5 listeners, while prosodic features were mentioned less often by 3 listeners. An additional 3 listeners did not provide a specific description and instead gave responses such as "I don't know."

Segmental features were the most salient category in the listeners' comments. Many listeners referred to consonant production, particularly the distinction between Korean stop categories. For example, some comments noted that "[k] and [k^h] were not distinguished," while other referred to other contrasts such as "[p] sounded like [p^h]" or "[k*] sounded like [k^h]." Some listeners also commented on the presence of a "strong English [r] sound," and responded that "[ɯi] sounded like [i]." These responses suggest that deviations of consonants and vowels were particularly noticeable to listeners and were consciously perceived as sources of foreign accent.

Fluency-related features were also frequently mentioned. Listeners described the accented speech as "slow," "reading word-by-word," or containing "many pauses." These comments indicate that temporal and delivery-related features of speech were highly salient in listeners' judgments.

Phonological patterns were mentioned by a smaller but still notable number of listeners. These comments included references to non-target realizations such as [tal.kɨl] instead of [ta.kɨl] and [tɕʌkʰ.iŋ] instead of [tɕʌ.giŋ], as well as comments such as "final [l] in [kʌl.lyʌ.t*a] was weak." They are in turn related to phonological patterned examined in Chapter 4 such as resyllabification and coda [l] before onset [l]. Such comments suggest that some listeners noticed deviated phonological patterns, although these were less frequently reported than segmental and fluency-related features.

Prosodic features were the least frequently mentioned category. Only three listeners referred to features such as "unnatural intonation" or "inaccurate pitch." But the comments did not include specific phrases. This suggests that prosodic deviations were also perceptible to some listeners, but were less salient overall than segmental and fluency-related features.

In sum, the open-ended responses indicate that listeners primarily associated foreign-accented Korean with segmental deviations and reduced fluency. Phonological patterns and prosodic features were mentioned less frequently in listeners' explicit comments. These qualitative findings complement the acoustic analyses by showing which aspects of learner speech were most noticeable to listeners.

Table 5.9
Summary of listeners' subjective responses about features perceived as foreign accents

Categories	Selected comments	Number of listeners (25)
Fluency-related features	"slow", "reading word-by-word", "many pauses"	10
Segmental features	"[k] and [k^h] were not distinguished", "[p] → [p^h]", "[k*] → [k^h]", "strong English [r] sound"	12
Phonological patterns	"[tal.kɨl] instead of [ta.kɨl]", "[tɕʌk^h.ɨŋ] instead of [tɕʌ.gɨŋ]", "final [l] was weak"	7
Prosodic features	"unnatural intonation" "inaccurate pitch"	3
No response	"I don't know"	3

5.5 Discussion

5.5.1 Accentedness judgments

The present chapter examined how speech-related features of American English-accented Korean are associated with native Korean listeners' accentedness judgments and language attitude evaluations. By examining both the acoustic analyses and listeners' subjective responses, this chapter provides a more detailed account of which speech features were related to different types of listener judgments.

With respect to accentedness judgments, phonological errors emerged as a robust predictor of higher accentedness ratings, suggesting that deviations from target phonological patterns are highly salient to native listeners. In contrast, hyperarticulated [h] was associated with lower accentedness ratings. Given that intervocalic [h] is frequently reduced or deleted in native Korean speech, hyperarticulation was initially expected to increase perceived foreign accent. However, the present results suggest that it may instead have been interpreted as a cue to careful

or precise pronunciation, potentially leading to reduced perceived accentedness. This raises the possibility that listeners may associate articulatory clarity with lower accentedness or higher proficiency, although this explanation remains speculative and needs further investigation.

Pause-related measures, including both frequency and duration of within-word and between-word pauses, did not show statistically significant effects on accentedness judgments. While Figure 5.2 indicates that speech containing pauses was generally associated with relatively high accentedness ratings, the lack of significant effects suggests that increases in pause frequency or duration did not further increase or decrease listeners' ratings. In other words, the presence of pauses may have contributed to the perception of accentedness, but once pauses were present, differences in their frequency or duration did not appear to modulate the degree of perceived accentedness.

By contrast, other temporal features showed clearer effects. Slower speech rate and longer vowel durations were both associated with higher accentedness ratings, indicating that reduced fluency contributes to perceptions of stronger foreign accent. These findings are consistent with previous research demonstrating that fluency-related temporal features, particularly speech rate, are strongly associated with accent perception (Maastricht et al., 2017; Munro & Derwing, 2001; Trofimovich & Baker, 2006).

Segmental acoustic features such as stop consonant and vowel measures were also associated with accentedness judgments, although their effects were specific to particular stop categories and prosodic positions. For stop consonants, the effects were mostly observed in VOT for lenis and aspirated stops in non-medial positions, as well as in F0 for fortis stops in word-medial position. Thus, while Chapter 4 showed that learners produced lower F0 for word-initial aspirated stop and longer VOTs for some stop categories than native speakers, the present results

suggest that these group-level differences did not necessarily translate into a direct effect on accentedness ratings and no single segmental cue consistently explained listeners' accented judgments. Rather, the effect of stop-related cues for accent perception depended on the specific stop category and its position in the sentence. Interestingly, the effect of VOT appeared to be relatively specific to accent perception. As discussed below, evaluations of status, solidarity, and comprehensibility showed more robust associations with F0 measures, while VOT had relatively little impact on these dimensions.

For vowels, several formant measures were also significantly associated with accentedness. In particular, more backed [e] and [ʌ] and higher tongue positions for [i] and [ʌ] were associated with higher accentedness ratings. At the same time, the tongue backness of [ɛ] and [ɨ] showed non-linear effects. However, these effects did not lead to a single general pattern across the vowel system. For example, F1 of [o] was one of the significant group differences between learners and native speakers in Chapter 4, but this feature was not found to be significantly associated with accentedness in the present chapter. In addition, some effects were linear, meaning that accentedness ratings increased or decreased steadily as the formant value changed, whereas others were non-linear, meaning that only particular ranges of the vowel space were associated with stronger accentedness ratings. Thus, the vowel results suggest that listeners were sensitive to specific ranges of variation in specific vowel categories, rather than to a broad native-learner difference in vowel production.

The relatively mixed pattern of acoustic predictors in the present chapter may also reflect differences in stimulus type and analysis level. Jung and Holliday (2026) found that deviations in F0 and F1 from L1 Korean norms were strong predictors of foreign accent ratings in isolated Korean words. In the present chapter, however, the effects of F0 and vowel formants

were less consistent. One possible explanation is that isolated word stimuli allow listeners to attend more directly to fine-grained acoustic properties of individual segments, whereas connected sentence-level speech contains more speech-related factors, including prosodic context, speech rate, and phonological errors. As a result, fine-grained acoustic cues such as F0 and F1 may have been less clearly isolated in the present sentence-level stimuli. This difference suggests that acoustic predictors of accentedness may vary depending on the granularity of the speech sample and the extent to which listeners can focus on individual segmental cues.

5.5.2 Language attitudes

Language attitude evaluations of status, solidarity, and comprehensibility showed different patterns from accentedness judgments, and it suggests that these evaluative dimensions draw on different aspects of the speech signal. Nonetheless, some predictors were shared across all three constructs, while others were specific to individual dimensions.

Several acoustic features were shared predictors across the three language attitude constructs. Vowel duration was significant for all three constructs, with longer vowel duration predicting lower ratings of status (e.g., competence), solidarity (e.g., friendliness) and comprehensibility, as well as higher accentedness as shown above. This pattern suggests that longer vowel duration was generally associated with less favorable listener evaluations in Korean. Interestingly, for status and solidarity dimensions, speech rate did not have a significant effect. This suggests that social evaluations such as status and solidarity were not simply based on overall speech tempo, but were more sensitive to particular temporal features of speech.

In addition, F0 of stop sounds, rather than VOT, generally contributed to the ratings of status, solidarity, and comprehensibility. This makes an interesting contrast to the perception of

accentedness, where VOT, not F0, was a more prominent predictor. Although the patterns varied across constructs, an emerging pattern was that lower F0 of word-initial lens stops and higher F0 of word-initial aspirated stops generally predicted higher ratings of the three constructs. This pattern may reflect the different perceptual roles of VOT and F0. VOT is a primary cue in English stop contrasts, and extended VOT may therefore be a salient feature of English-accented Korean. For this reason, VOT deviations may be more directly linked to accentedness judgments. By contrast, F0 is a pitch-related cue, which may contribute more to listeners' broader impressions of how natural, competent, or friendly the speech sounds.

Further, the opposite patterns for word-initial and word-medial lenis F0 suggest that listeners' responses were not driven by a general preference for higher or lower pitch of a specific stop category. Rather, listeners may have evaluated F0 in relation to its prosodic position and the expected pitch pattern of Korean. In particular, the favorable evaluations associated with higher word-medial lenis F0 may be related to phrasal tone or local pitch movement within the utterance, rather than to the stop category alone. However, because the present study did not directly model phrasal tone or intonational structure, this interpretation remains tentative and should be examined in future research.

Vowel formants were also associated with all three constructs, but their interpretation is less straightforward. Rather than discussion specific vowel quality effects on language attitudes, these results indicate that listeners' social evaluations and perceived comprehensibility were sensitive to fine-grained vowel variation. This was especially apparent for the solidarity dimension, which was associated with a wide range of vowel predictors. One possible interpretation is that, because solidarity reflects in-group preference or perceived social

closeness, listeners may be especially sensitive to subtle differences in speech when making judgments on solidarity.

With respect to status, the predictors including hyperarticulated [h], vowel duration, F0 following stops, and some vowel formants contributed to the ratings. Especially, the presence of hyperarticulated [h] was associated with lower rating of status and this effect was also found in the comprehensibility dimension. This pattern is noteworthy because hyperarticulated [h] was associated with lower accentedness ratings, but it predicted less favorable status evaluations. In this sense, the present results suggest that features reducing perceived foreign accent do not necessarily lead to more positive status evaluations. One possible interpretation is that hyperarticulated [h] may have been perceived as a cue to careful pronunciation, which reduced perceived accentedness but at the same time made the speech sound less socially fluent and competent. If so, listeners may have perceived this kind of careful articulation not as a sign of higher competence, which in turn lowered status evaluations.

For the rating of solidarity dimension, the predictors including within-word pause frequency and duration, vowel duration, mostly lenis F0 and VOT, and all vowels contributed to the ratings. Among these predictors, the pause-related features were uniquely associated with the solidarity dimension. Pauses that break up a word seem to affect a sense of closeness or friendliness, but not status or comprehensibility. However, the direction of the effects was inconsistent: a higher frequency of within-word pauses lowered solidarity ratings, whereas longer within-word pauses were associated with higher solidarity ratings. Although these factors are in different directions, they may suggest that listeners responded differently to the number of within-word pauses and to the duration of within-word pauses when evaluating solidarity. A greater number of within-word pauses may have made the speech sound more fragmented and

less fluent, thereby reducing impressions of social closeness. By contrast, longer within-word pauses, when they are not frequent, may have signaled careful speech or impressions of sincerity, which could have contributed to higher solidarity ratings.

Finally, comprehensibility showed significant relationships with phonological error, hyperarticulated [h], speech rate, vowel duration, F0 following mostly lenis stops, and some vowel formants predicted the ratings of comprehensibility. It is noted that this profile is similar to that of accentedness judgments, which also included phonological error, hyperarticulated [h], speech rate, vowel duration, and some vowel formants. Although the present chapter did not directly examine the correlation between accentedness and comprehensibility, this overlap is compatible with Isbell et al.'s (2024) finding that the two dimensions were highly correlated in L2 Korean speech. In particular, phonological errors and speech rate were associated with comprehensibility and accentedness, but not with status or solidarity. This suggests that these two features may be more directly related to how clear or easy the speech sounded, rather than to broader social evaluations of the speaker. Faster speech rate predicted higher comprehensibility ratings, whereas phonological errors and longer vowel durations reduced comprehensibility. This pattern is consistent with previous studies showing that comprehensibility is closely related to fluency-related temporal features (O. Kang, 2010; Munro & Derwing, 2001; Trofimovich et al., 2022; Trofimovich & Baker, 2006). Segmental features were also associated with comprehensibility. In particular, comprehensibility was associated mainly with F0 values, with an additional effect of word-medial aspirated VOT and several vowel-related measures. This suggests that comprehensibility, like status, depended less on segmental deviations than on fluency-related properties.

As the final note on this topic, some non-linear effects observed across several fluency-related and segmental predictors suggest that listeners' evaluations did not always vary linearly with acoustic difference. Instead, native listeners may have internalized expected acoustic ranges for specific features, such as VOT and F0 in certain stop categories. Therefore, values falling within these ranges were treated similarly to native speech, whereas values falling outside them were evaluated differently. This interpretation helps explain why some variables showed significant non-linear effects rather than simple linear relationships.

The qualitative responses provide an additional perspective on these results. Native listeners most frequently referred to segmental features, followed by fluency-related features, when describing what made the speech sound foreign-accented. Particularly, many comments mentioned stop consonant contrasts and their deviations and it suggests that segmental features were consciously salient to listeners. However, these explicit responses were not fully aligned with the statistical results for accentedness. Although some segmental predictors were associated with accentedness, the overall results do not suggest that stop-related measures were the strongest or most consistent predictors of accentedness judgments. This instead suggests that the features listeners consciously identify as "foreign" are not always the same as those that are associated with their ratings in statistical models, and that supplementary qualitative analysis may be helpful in some cases.

One of the noteworthy comments was the word /dalg-eul/, which three listeners identified as sounding foreign when produced as [tal.kɨl], even though this pronunciation is prescriptively correct and was also produced by one of the native speaker anchors. This suggests that listeners' judgments of foreign accent may not always align with prescriptive phonological norms. Instead, they may reflect expectations that are shaped by patterns that are more familiar

and more frequently encountered in everyday speech. In other words, a pronunciation can be phonologically correct but still sound marked or less natural if it deviates from the listener's familiar expectations. This observation further supports the interpretation that listener judgments are shaped not only by prescriptive correctness, but also by probabilistic expectations about what sounds natural or native-like.

Overall, the results of this chapter indicate that accentedness judgments and language attitude evaluations were shaped by different sets of speech features. While accentedness was more associated with phonological deviations and fluency-related features, language attitudes drew on more different combinations of temporal and segmental properties. Importantly, the present chapter also suggests that fine-grained segmental features such as VOT, F0 and vowel formants, were associated not only with accentedness but also with social evaluative dimensions such as status and solidarity. In addition, the non-linear effects and the mismatch between qualitative responses and statistical results indicate that listener evaluation is shaped not only by explicit awareness of salient features, but also by implicit expectations about what is perceived as natural or native-like speech.

5.6 Conclusion

This chapter examined which speech-related features of American English-accented Korean were associated with native Korean listeners' accentedness judgments and language attitude evaluations. The present chapter showed that perception of foreign accent and language attitude evaluations were shaped by different sets of speech properties. Accentedness judgments were mostly associated with phonological deviations and fluency-related features such as speech rate and vowel duration, whereas language attitude dimensions showed more differentiated

patterns in that comprehensibility was most clearly related to fluency-related features, status was associated more with temporal and F0-related cues, and solidarity was associated with a broader range of temporal and segmental properties.

The results also showed that fine-grained segmental features such as VOT, F0, and vowel formants were associated not only with accentedness but also with social evaluative dimensions such as status, solidarity, and comprehensibility. At the same time, several predictors showed non-linear effects, which suggests that listener evaluations may depend not simply on whether speech features of learner productions differ from native speech, but also on whether particular acoustic values fall within or outside perceptually expected native-like ranges. This indicates the importance of examining acoustic features at a more fine-grained level than just group-level comparisons between native and learner productions.

Finally, the qualitative responses showed that listeners consciously identified segmental and fluency-related features as important sources of foreign accent, especially stop-related deviations. However, these explicit responses were not fully aligned with the statistical results, which means that the features listeners consciously notice are not necessarily the same as those that strongly shape their ratings. In summary, these findings show that evaluations of foreign-accented Korean are multidimensional and are shaped by both explicit awareness of salient deviations and implicit expectations about what sounds native-like. Therefore, the present chapter extends previous work on foreign-accented speech perception by showing that acoustic correlates of foreign-accented Korean speech are relevant not only to accentedness judgments, but also to listener evaluations such as status, solidarity, and comprehensibility.

VI. CONCLUSION

This dissertation examined how native Korean listeners perceive and evaluate foreign-accented Korean speech. Specifically, it investigated how social expectations associated with race, ethnicity, and nationality, as well as expectation mismatch between speaker identity and accent, influence intelligibility, accentedness judgments, and language attitudes. It also sought to explore acoustic correlates of perceived accents and social evaluations by examining American English learners production of Korean stops, vowels, and phonological patterns. Together, this dissertation provides a multidimensional account of foreign-accented Korean speech perception by integrating social and acoustic dimensions.

6.1 Summary of the findings

The main research question of Chapter 2 was whether social information about a speaker influences native Korean listeners' intelligibility and accentedness judgments of foreign-accented Korean. The findings showed that social information had limited effects on intelligibility in that sentence-level transcription accuracy did not reliably differ across guise conditions, and only one target word contrast between White and Chinese guises showed a possible race-based socioeconomic interpretation. In contrast, accentedness judgments showed evidence of expectation mismatch, but no bias based on race or ethnicity, in that the Korean guise was rated as more accented than the Baseline condition which provided no speaker information other than the speaker was a learner. Thus, unlike prior studies on accented English (Babel & Russell, 2015; Kutlu et al., 2022; McGowan, 2015), Chapter 2 demonstrated that social information such as race and ethnicity about the speaker did not systematically affect how

accurately Korean listeners understood the speech, but it influenced how accented the speech was when the speaker identity had a mismatch with the expectation of foreign accent.

Chapter 3 examined whether the same social information examined in Chapter 2 shaped language attitudes toward foreign-accented Korean speakers. The findings showed that social information about the speaker's race and ethnicity influenced listeners' evaluations, but not in a way that aligned with previously reported status-solidarity patterns in the field. The Korean guise received less favorable affective evaluations, again suggesting an effect of expectation mismatch. In the meantime, the Korean-American guise did not receive the expected solidarity advantage associated with shared ethnicity, nor did its association with U.S. nationality lead to higher status evaluations. This suggests that shared ethnicity and nationality-based high-status cues may not necessarily confer a social advantage for the Korean-American guise when paired with foreign-accented speech. This pattern contrasts with previous research (Coupland & Bishop, 2007; Dragojevic et al., 2013; Fiske et al., 2007; Fuertes et al., 2012) suggesting that shared in-group membership enhances solidarity judgments and high-status social cues increase status judgments.

Chapter 4 investigated how American English learners of Korean differ from native Korean speakers in their production of Korean stops, vowels, and connected speech phonological patterns. The findings showed relatively limited differences in vowel production, but clearer differences in stop production and realization of phonological rules. To produce stop contrasts, learners relied on VOT and F0, whereas native Seoul Korean speakers relied on F0, especially in distinguishing lenis and aspirated stops. Learner productions also showed recurrent non-native patterns in phonological realization, suggesting that accented Korean involves not only segmental acoustic differences but also differences in phonological patterns.

Lastly, Chapter 5 explored which speech-related features are associated with accentedness judgments and language attitude evaluations. The findings showed that accentedness was associated with several features including phonological errors, temporal features, and segmental features. Language attitude ratings were also associated with acoustic and phonological features, but different evaluative dimensions were related to different cues. Further, although features associated with higher accentedness did not always lead to less favorable social evaluations, this chapter showed that specific speech-related features contribute to both accentedness judgments and social evaluations.

In sum, the findings suggest that native Korean listeners' responses to foreign-accented Korean speech are shaped by both acoustic features and social expectations. Thus, their judgments reflect the combined influence of the speech signal, social expectations about the speaker, and social meanings specific to the Korean context.

6.2 Contributions

This dissertation makes several contributions to research on foreign-accented speech perception, sociophonetics, and language attitudes. First, it extends research on RLS and exemplar account of speech perception to the Korean context. Much of the previous work on foreign-accented speech perception has been conducted in English-speaking contexts or with Indo-European languages, where racial minority, foreignness, non-nativeness, and lower socioeconomic status are often linked. However, South Korea provides a different sociolinguistic context where foreignness is not always associated with lower socioeconomic status, and ethnic in-group membership does not necessarily lead to cultural belonging. By examining Korean listeners' responses to different social information paired with foreign-accented speech, this

dissertation shows how theories developed in Western contexts can be tested and refined in a non-Western, Korean-speaking context.

Specifically, the findings demonstrate that the status and solidarity dimensions of language attitudes may not map exactly onto Korean listeners' responses. Previous research, typically conducted in English-speaking context, repeatedly showed that speakers associated with socioeconomically high status receive favorable ratings on status, while in-group speakers receive favorable ratings on solidarity (Coupland & Bishop, 2007; Dragojevic et al., 2013; Fiske et al., 2007; Fuertes et al., 2012). However, in the current study, an ethnically in-group member (i.e., Korean-American guise) did not receive a clear advantage on solidarity, even though a Korean-American individual was ethnically closer to Korean listeners than a White individual. This ethnically in-group member did not receive a nationality-based advantage on status either, even though the U.S. nationality was expected to be associated with higher status. These findings challenge the idea that in-group membership universally leads to positive solidarity evaluations and that higher socioeconomic status leads to positive status evaluations.

Chapter 4 contributes to research on Korean second language phonetics by documenting acoustic and phonological characteristics of American English-accented Korean. The findings showed that learner speech differed from native Korean speech not only in acoustic dimensions, but also in cue use and phonological patterns in connected speech. In particular, learners' stronger reliance on VOT relative to F0 suggests that American English learners may carry over English-like cue-weighting patterns into Korean stop production, consistent with prior studies such as (H. Kim, 2023; M.-R. C. Kim & Lotto, 2002; Kong et al., 2022). While many studies focused on phonetic features, few studies have investigated learners' realization of Korean phonological rules. The current study offers a discovery that accentedness in Korean speech may

arise from difficulty applying Korean phonological rules in connected speech, not simply from difficulty producing individual sounds.

Finally, this dissertation contributes to the literature by exploring speech-related features associated with both accentedness and language attitudes. Previous research has often shown that accented speech affects social evaluation, but fewer studies have examined which specific acoustic features are associated with status, solidarity, and comprehensibility. The present findings show that listeners' social evaluations are influenced not only by social information or overall accentedness, but also by fine-grained speech features.

6.3 Limitations and future directions

Several limitations of this dissertation should be acknowledged. First, the examination of intelligibility (Chapter 2) and language attitudes (Chapter 3) had a relatively small sample size, with 25 participants in each experimental condition. Although the study was able to show several meaningful patterns, the sample size may have limited the possibility to find stronger effects. Some effects may not have emerged clearly because the study did not have enough power to detect subtle differences across conditions. Future research could replicate the study with a larger sample to determine whether the observed patterns are stable across a bigger population of native Korean listeners.

Second, the findings may depend on the specific accented speech used in the study. Chapters 2 and 3 used moderately accented Korean speech. If the speech had been less accented or more heavily accented, the results may have differed. This is especially relevant for the Korean and Korean-American guises. Although the Korean guise showed reliably higher accentedness ratings than the Baseline condition, the effects may have been more clearly

observed between the Korean guise and the other guises if the speech had been more strongly accented. On the other hand, if the Korean-American guise had been paired with less-accented Korean speech, listeners might have evaluated the speaker more favorably in terms of solidarity or status. Future research could manipulate accent heaviness directly to examine how social information interacts with different degrees of accentedness.

Regarding the Korean-American guise, another future direction is to examine whether speaker identity changes how native Korean listeners use acoustic cues in perceiving accented Korean speech. Building on work (Hay & Drager, 2010; Niedzielski, 1999) showing that social information can shift phonetic categorization and cue use, future research could present the same Korean stop continuum while describing the speaker as Korean-American or American. If Korean listeners expect Korean-American speakers to sound more similar to native Korean speakers, they may rely more strongly on native Korean cue-weighting patterns or show sharper category boundaries. By contrast, when the same speech is presented as American, listeners may tolerate non-native cue patterns more readily. This approach would clarify whether Korean-American identity creates an expectation of linguistic in-groupness, and whether that expectation changes the perception of foreign-accented Korean.

In addition, the dissertation examined a limited set of social categories, in that the guises used in this dissertation all involved national identities that may be associated with relatively high socioeconomic status in the Korean context. Although this design allowed the study to test race-based socioeconomic expectations, ethnic in-group membership, and expectation mismatch, it did not fully examine how Korean listeners respond to foreign-accented Korean when the speaker is associated with different levels of social status or different racial backgrounds. Future research could therefore include speaker descriptions that vary more explicitly in socioeconomic

status or race, while holding accent constant. This would clarify whether the patterns observed in this dissertation are specific to the high-status national identities tested here, or whether they reflect broader status-based and racial ideologies about accented Korean.

Finally, the acoustic analysis was constrained by the stimulus materials. In Chapter 4, aspirated and fortis stops did not occur in phrase-initial position, and the number of tokens across the three stop categories was not fully balanced across sentences. These limitations made it difficult to fully compare lenis, aspirated, and fortis stops across prosodic positions. Future studies should use stimuli specifically designed to balance stop categories across phrase-initial, word-initial, and word-medial positions. This would allow a more systematic analysis of how learners produce Korean laryngeal contrasts across prosodic contexts.

6.4 Conclusion

Overall, this dissertation shows that foreign-accented Korean speech perception is shaped by both acoustic properties and social expectations. The findings do not simply support that social expectations affect listener judgments, but instead, intelligibility, accentedness, and language attitudes were influenced in different ways. Social information had limited effects on intelligibility, but expectation mismatch shaped accentedness judgments and affective evaluations. Acoustic features also contributed to accentedness and language attitudes, but their effects differed across evaluative dimensions.

These findings have cautious implications for L2 Korean speakers. The results suggest that social information about a speaker does not necessarily lead to serious communication difficulty and even accentedness was not reliably correlated with intelligibility, suggesting that

accented speech is not necessarily difficult to understand. In this sense, achieving a native-like accent should not be treated as a necessary condition for successful communication in Korean.

Further, by examining foreign-accented Korean in the South Korean context, this dissertation shows that theories of accented speech perception must attend to local social meanings. Foreignness is not always associated with lower socioeconomic status, ethnic in-group membership does not necessarily guarantee solidarity, and Korean identity paired with foreign-accented speech can create expectation mismatch. Understanding accented speech perception therefore requires attention not only to learners' productions, but also to listeners' expectations and the specific sociolinguistic context where communication takes place.

APPENDICES

Appendix A

List of sentence stimuli

A list of 60 Korean sentences used for the pilot and production studies with an asterisk (*) marking the 30 selected sentences. Sentences in bold indicate the items used in Chapter 5. English translations are provided below each sentence.

1. 직원들이 사장의 성격을 닮았다.

   The employees have personalities similar to the boss.

2. 그는 친구들의 생각을 잘 읽는다.

   He can read his friends' thoughts well.

*3. 나는 자연 속에서 마음을 비웠다.

   I cleared my mind while being in nature.

4. 어머니가 가을에 산을 다니신다

   My mother goes hiking in the mountains in the fall.

5. 나는 회의가 끝나고 마음이 무거웠다.

   I felt heavy-hearted after the meeting.

6. 잔치에 음식과 과일이 충분하다.

   There is plenty of food and fruit at the feast.

*7. 그의 이상한 행동이 관객들을 웃겼다.

   His strange behavior made the audience laugh.

8. 나는 그의 안타까운 사정을 들었다.

   I heard about his unfortunate situation.

*9. 우리 아들과 딸은 힘들게 자랐다.

   Our son and daughter had a difficult childhood.

10. 나는 케이크를 먹다가 그에게 걸렸다.

   He caught me eating cake.

*11. 그가 동네 여학생들 앞에서 넘어졌다.

   He fell in front of the neighborhood girls.

12. 학생들이 소설의 내용을 이해했다.

The students understood the novel's content.

***13. 나는 감독의 지도 방식에 적응했다.**

I got used to the director's coaching style.

14. 이 나라의 우체국 업무는 빠르다.

Postal services in this country are fast.

*15. 아저씨가 귀여운 강아지에게 물렸다.

The older guy got bitten by a cute puppy.

*16. 회사가 제품의 문제점을 고쳤다.

The company fixed the product defects.

*17. 마을에 위험한 짐승들이 내려왔다.

Dangerous animals came down to the village.

***18. 할머니가 부엌에서 닭을 끓였다.**

My grandmother cooked chicken in the kitchen.

19. 나는 다양한 종류의 술을 담근다.

I make different types of liquor.

*20. 기술자가 자동차 바퀴를 갈았다.

The technician changed the car tire.

*21. 나는 요리사를 도와 야채를 씻었다.

I helped the chef wash the vegetables.

22. 나는 엄마의 가방을 몰래 열었다.

I secretly opened my mom's bag.

*23. 젊은이들이 우리 아파트를 지킨다.

Young people are protecting our apartment.

24. 아이가 웬일로 얌전하게 공부했다.

The child studied quietly for some reason.

25. 아버지는 옛날 사고 방식을 바꾸셨다.

My father changed his old-fashioned way of thinking.

26. 부모와 자식 간에 대화가 중요하다.

Communication between parents and children is important.

27. 은행의 업무 분위기는 딱딱하다.

The work environment in the bank is strict.

28. 아내는 남편과 성격이 비슷하다.

The wife has a personality similar to her husband's.

29. 그녀는 직장에 많은 것을 걸었다.

She put a lot on the line for her job.

*30. 이 영화는 어린이들에게 해롭다.

This movie is harmful for children

31. 나는 훌륭한 건축가와 결혼했다.

I married a great architect.

32. 신부 아버지가 결혼식장에 도착했다.

The bride's father arrived at the wedding venue.

*33. 선생님과의 입시 상담이 끝났다.

The entrance exam consultation with the teacher has ended.

34. 그는 고기가 익지 않아서 뱉었다.

He spat out the meat because it was undercooked.

*35. 나는 유명한 사람의 얼굴을 그린다.

I draw portraits of famous people

*36. 이번에 선수의 기록이 올라갔다.

The athlete set a new personal record this time.

*37. 나는 공연 중에 무대에서 떨어졌다.

I fell off the stage during the performance.

38. 그 사람의 말투와 표정이 무섭다.

His way of speaking and facial expressions are intimidating.

*39. 아이들이 불평 없이 잘 일어난다.

The children wake up without complaining.

40. 아이들이 교실 한 가운데에 앉았다.

The children sat in the center of the classroom.

41. 나는 그 가수의 춤 실력이 아쉽다.

I was disappointed in that singer's dancing skills.

*42. 그녀는 그 친구의 전화번호를 지웠다.

She deleted her friend's phone number.

*43. 그가 같은 반의 조용한 친구를 때렸다.

He hit the quiet classmate in his class.

44. 호랑이가 오랫동안 먹이를 씹었다.

The tiger chewed its food for a long time.

*45. 나는 회의 내용의 대부분을 놓쳤다.

I missed most of the meeting.

46. 나는 식사를 하고 바로 잠들었다.

I fell asleep right after eating.

*47. 나는 추운 베란다에 야채를 심었다.

I planted vegetables on the cold balcony.

48. 행사장에 간식과 음료가 남았다.

There were leftover snacks and drinks at the event.

***49. 종업원들이 설거지를 하고 쉬었다.**

The employees did the dishes and then took a break.

*50. 그는 늘 신선한 오징어를 먹는다.

He always eats fresh squid.

*51. 그는 매년 새로운 곳으로 이사한다.

He moves to a new place every year.

52. 나는 그의 목소리를 들으니 반가웠다.

I felt happy when I heard his voice.

53. 선생님이 학생의 연락에 감동했다.

The teacher was moved by the student's message.

*54. 그는 선생님들의 의견을 모았다.

He gathered the teachers' opinions.

*55. 나는 부모님께 돈을 겨우 갚았다.

I barely managed to repay my parents.

***56. 엄마는 급한 마음에 빠르게 걸었다.**

My mom walked quickly in a hurry.

57. 그는 머리부터 발끝까지 다 젖었다.

He was soaked from head to toe.

58. 나는 그 곳에 필요한 물건을 기부했다.

I donated the necessary items there.

*59. 부장님은 회의실의 유리창을 가렸다.

The manager covered the conference room window.

***60. 나는 집에서 일을 하니 답답하다.**

I feel frustrated working from home.

Appendix B

Examples of transcriptions judged as correct or incorrect

Examples of transcriptions categorized as correct or incorrect. Minor variation such as informal spellings or typographical errors was considered correct, and particles were judged generously. However, transcriptions that change the intended meaning were marked as incorrect.

Stimulus	Transcribed	Judgment	Reasons
비웠다 piwessta	버렸다 peryessta	incorrect	Different meaning
자랐다 calassta	자랬다 calayssta	incorrect	Different meaning
끓였다 kkulyessta	그렸다 kuryessta	incorrect	Different meaning
갈았다 kalassta	간다 kanta	incorrect	Tense mismatch and more than two different letters
지킨다 cikinta	지키고있다 cikikoissta	incorrect	Similar meaning but different syllable counts

심었다 simessta	심는다 simnta	incorrect	Tense mismatch and more than two different letters
마음 maum	맘 mam	correct	Common informal spelling
웃겼다 wuskyessta	웃겼따 wuskyesstta	correct	Common typo and meaning does not change
갚았다 kapassta	가파타 kaphatha	correct	Phonologically transcribed

Appendix C

Language attitude auestionnaire

Participants responded to the following statements after intelligibility and accentedness judgment tasks. All items were presented in Korean during the experiment. English translations are provided here for refence. Responses were recorded on a 7-point Likert scale (1 = Not at all, 7 = Extremely).

Positive affect

1. I feel *interested* in the speaker's language ability or background.
   화자의 언어 능력이나 배경에 "관심이 간다"
2. I feel *commendable* of the speaker's effort to use Korean.
   화자가 한국어를 사용하는 모습이 "대견스럽다"
3. I feel *happy* considering the speaker's effort to speak Korean.
   한국어를 말하려는 화자의 노력을 생각하니 "기쁘다"

Negative affect

4. I feel *frustrated* trying to understand the speaker's sentences.
   화자의 문장을 이해하려다 보니 "답답하다"
5. I feel *bothered* by the effort of understanding what the speaker says.
   화자가 말하는 것을 이해하기 "귀찮다"
6. I feel *irritated* listening to the accented Korean of the speaker.
   화자의 억양이 있는 한국어를 듣기가 "짜증난다"

Status

7. The speaker sounds *intelligent.*
   화자가 “지적인” 사람처럼 느껴진다
8. The speaker sounds *competent.*
   화자가 “유능한” 사람처럼 느껴진다
9. The speaker sounds *well-educated*.
   화자가 “교육을 잘 받은” 사람처럼 느껴진다
10. The speaker sounds *accomplished*.
    화자가 “능숙한” 사람처럼 느껴진다
11. The speaker sounds *smart*.
    화자가 “똑똑한” 사람처럼 느껴진다

Solidarity

12. The speaker sounds *friendly*.
    화자가 “친근한” 사람처럼 느껴진다
13. The speaker sounds *sociable*.
    화자가 “사교적인” 사람처럼 느껴진다
14. The speaker sounds *trustworthy*.
    화자가 “신뢰할 수 있는” 사람처럼 느껴진다
15. The speaker sounds *warm*.
    화자가 “따뜻한” 사람처럼 느껴진다
16. The speaker sounds *nice*.
    화자가 “괜찮은” 사람처럼 느껴진다

Comprehensibility

17. It was *easy to understand* the speaker’s sentences.
    화자의 문장을 이해하기 쉬웠다
18. The speaker’s pronunciation and way of speaking were *clear.*
    화자의 발음이나 말하는 방식이 명확했다
19. It was *effortful to understand* the speaker.

화자의 말을 이해하는 데 노력이 필요했다

20. The speaker's speech sounded *natural and smooth*.

화자의 말이 자연스럽고 매끄럽게 느껴졌다

Appendix D

Demographic and language background questionnaire

1. Gender
    - Male
    - Female
    - I would rather not say
    - Other
2. Age
4. Do you have any hearing or language-related disorder?
    - Yes
    - No
5. Education level
    - High school or equivalent
    - Associate degree (2-3 years; 전문대학)
    - Bachelor's degree (4 years)
    - Master's degree
    - Doctoral degree
    - Other
6. What is your native language?
7. If you have learned languages other than English, please provide the following information:
    - Language
    - Proficiency level
    - Age of acquisition
8. Where do you currently live?

9. If you have lived in another country for more than 6 months, please provide the following information:

- Country
- Language(s) spoken in that country or region
- Age at arrival
- Length of residence

10. If you primarily speak a dialect, which region does it originate from?

- Seoul or Metropolitan Area
- Chungcheong
- Gangwon
- Jeolla
- Gyeongsang
- Jeju

9. Do you have any foreign friends learning Korean?

- Yes
- No

10. If yes, how often do you talk with them in Korean?

- Never
- Rarely
- Sometimes
- Often
- Always

11. Have you ever watched TV shows or YouTube videos featuring foreigners speaking Korean? (e.g., “Abnormal Summit”)

- Yes
- No

12. If you have watched broadcast or YouTube videos featuring foreigners speaking Korean, how much have you watched?

- Once or twice
- Several times

- Sometimes
- Often
- Always

13. How important do you think accent or pronunciation is when learning or using a foreign language?

- Not important at all
- Slightly important
- Moderately important
- Very important
- Extremely important

REFERENCES

Ahn, J. (2023). An Influencer and Online Nationalism: The Case of the "Kimchi" Dispute on YouTube. *International JOURNAL OF CONTENTS*, *19*(3), 85–99. https://doi.org/10.5392/IJoC.2023.19.3.085

Ahn, J.-H. (2015). Desiring biracial whites: Cultural consumption of white mixed-race celebrities in South Korean popular media. *Media, Culture & Society*, *37*(6), 937–947. https://doi.org/10.1177/0163443715593050

Allport, G. W. (1954). The Nature of Prejudice. *Addison-Wesley*, 537.

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the Tribes of Fluency to Form a Metacognitive Nation. *Personality and Social Psychology Review*, *13*(3), 219–235. https://doi.org/10.1177/1088868309341564

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The Relationship Between Native Speaker Judgments of Nonnative Pronunciation and Deviance in Segmentais, Prosody, and Syllable Structure. *Language Learning*, *42*(4), 529–555. https://doi.org/10.1111/j.1467-1770.1992.tb01043.x

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Babel, M., & Russell, J. (2015). Expectations and speech intelligibility. *The Journal of the Acoustical Society of America*, *137*(5), 2823–2833. https://doi.org/10.1121/1.4919317

Baese-Berk, M. M., McLaughlin, D. J., & McGowan, K. B. (2020). Perception of NON-NATIVE speech. *Language and Linguistics Compass*, *14*(7), e12375. https://doi.org/10.1111/lnc3.12375

Barkhuizen, G. (2019). Core dimensions of narrative inquiry. In J. McKinley & H. Rose (Eds.), *The Routledge Handbook of Research Methods in Applied Linguistics* (1st ed., pp. 188–198). Routledge. https://doi.org/10.4324/9780367824471-17

Bassetti, B. (2008). 10. Orthographic Input and Second Language Phonology. In T. Piske & M. Young-Scholten (Eds.), *Input Matters in SLA* (pp. 191–206). Multilingual Matters. https://doi.org/10.21832/9781847691118-013

Bassetti, B. (2024). Orthographic Effects in the Phonetics and Phonology of Second Language Learners and Users. In M. Amengual (Ed.), *The Cambridge Handbook of Bilingual Phonetics and Phonology* (1st ed., pp. 699–720). Cambridge University Press. https://doi.org/10.1017/9781009105767.032

Baugh, J. (1983). *Black street speech: Its history, structure, and survival*. University of Texas Press.

Beckman, J., Helgason, P., McMurray, B., & Ringen, C. (2011). Rate effects on Swedish VOT: Evidence for phonological overspecification. *Journal of Phonetics*, *39*(1), 39–49. https://doi.org/10.1016/j.wocn.2010.11.001

Bergeron, A., & Trofimovich, P. (2017). Linguistic Dimensions of Accentedness and Comprehensibility: Exploring Task and Listener Effects in Second Language French. *Foreign Language Annals*, *50*(3), 547–566. https://doi.org/10.1111/flan.12285

Bestelmeyer, P. E. G., Belin, P., & Ladd, D. R. (2015a). A Neural Marker for Social Bias Toward In-group Accents. *Cerebral Cortex*, *25*(10), 3953–3961. https://doi.org/10.1093/cercor/bhu282

Bestelmeyer, P. E. G., Belin, P., & Ladd, D. R. (2015b). A Neural Marker for Social Bias Toward In-group Accents. *Cerebral Cortex*, *25*(10), 3953–3961. https://doi.org/10.1093/cercor/bhu282

Boersma, P., & Weenink, D. (2024). Praat: Doing phonetics by computer, 2024. *Place: Amsterdam*.

Brown, V. A., McLaughlin, D. J., Strand, J. F., & Van Engen, K. J. (2020). Rapid adaptation to fully intelligible nonnative-accented speech reduces listening effort. *Quarterly Journal of Experimental Psychology*, *73*(9), 1431–1443. https://doi.org/10.1177/1747021820916726

Bürkner, P.-C. (2017). **brms**: An *R* Package for Bayesian Multilevel Models Using *Stan*. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01

Bybee, J., & Torres, R. (2008). Phonological and Grammatical Variation in Exemplar Models. *Studies in Hispanic and Lusophone Linguistics*, *1*(2), 399–414. https://doi.org/10.1515/shll-2008-1026

Campbell-Kibler, K. (2007). ACCENT, (ING), AND THE SOCIAL LOGIC OF LISTENER PERCEPTIONS. *American Speech*, *82*(1), 32–64. https://doi.org/10.1215/00031283-2007-002

Cargile, A. C., & Giles, H. (1997). Understanding language attitudes: Exploring listener affect and identity. *Language & Communication*, *17*(3), 195–217. https://doi.org/10.1016/S0271-5309(97)00016-5

Carlson, H. K., & McHenry, M. A. (2006). Effect of accent and dialect on employability. *Journal of Employment Counseling*, *43*(2), 70–83. https://doi.org/10.1002/j.2161-1920.2006.tb00008.x

Casasanto, L. S. (2008). Does Social Information Influence Sentence Processing? *Proceedings of the Annual Meeting of the Cognitive Science Society*, *Vol. 30*(No. 30).

Chang, C. B. (2009). The Implementation of Laryngeal Contrast in Korean as a Second Language. *UC Berkeley Phonology Lab Annual Reports*, *5*. https://doi.org/10.5070/P743Q21727

Cho, T., Jun, S.-A., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics*, *30*(2), 193–228. https://doi.org/10.1006/jpho.2001.0153

Choi, J. (2020). How does focus-induced prominence modulate phonetic realizations for Korean word-medial stops? *Phonetics and Speech Sciences*, *12*(4), 57–61. https://doi.org/10.13064/KSSS.2020.12.4.057

Clopper, C. G., Burdin, R. S., & Turnbull, R. (2019). Variation in /u/ fronting in the American Midwest. *The Journal of the Acoustical Society of America*, *146*(1), 233–244. https://doi.org/10.1121/1.5116131

Collins, K. A., & Clément, R. (2012). Language and Prejudice: Direct and Moderated Effects. *Journal of Language and Social Psychology*, *31*(4), 376–396. https://doi.org/10.1177/0261927X12446611

Coupland, N., & Bishop, H. (2007). Ideologised values for British accents[1]. *Journal of Sociolinguistics*, *11*(1), 74–93. https://doi.org/10.1111/j.1467-9841.2007.00311.x

Cronbach, L. J. (1951). Coefficient Alpha and the Internal Structure of Tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/BF02310555

Crowther, D., Trofimovich, P., Saito, K., & Isaacs, T. (2018). LINGUISTIC DIMENSIONS OF L2 ACCENTEDNESS AND COMPREHENSIBILITY VARY ACROSS SPEAKING

TASKS. *Studies in Second Language Acquisition*, *40*(2), 443–457. https://doi.org/10.1017/S027226311700016X

Díaz-Campos, M., & Killam, J. (2012). Assessing Language Attitudes through a Matched-guise Experiment: The Case of Consonantal Deletion in Venezuelan Spanish. *Hispania*, *95*(1), 83–102. https://doi.org/10.1353/hpn.2012.a469668

Diller, K. (1981). "NATURAL METHODS" OF FOREIGN-LANGUAGE TEACHING: CAN THEY EXIST? WHAT CRITERIA MUST THEY MEET? *Annals of the New York Academy of Sciences*, *379*(1), 75–86. https://doi.org/10.1111/j.1749-6632.1981.tb41999.x

Drager, K. (2010). Sociophonetic Variation in Speech Perception. *Language and Linguistics Compass*, *4*(7), 473–480. https://doi.org/10.1111/j.1749-818X.2010.00210.x

Drager, K. (2011). Speaker Age and Vowel Perception. *Language and Speech*, *54*(1), 99–121. https://doi.org/10.1177/0023830910388017

Dragojevic, M., & Giles, H. (2014). The Reference Frame Effect: An Intergroup Perspective on Language Attitudes: Reference Frame Effect. *Human Communication Research*, *40*(1), 91–111. https://doi.org/10.1111/hcre.12017

Dragojevic, M., & Giles, H. (2016). I Don't Like You Because You're Hard to Understand: The Role of Processing Fluency in the Language Attitudes Process: Processing Fluency and Language Attitudes. *Human Communication Research*, *42*(3), 396–420. https://doi.org/10.1111/hcre.12079

Dragojevic, M., Giles, H., Beck, A.-C., & Tatum, N. T. (2017). The fluency principle: Why foreign accent strength negatively biases language attitudes. *Communication Monographs*, *84*(3), 385–405. https://doi.org/10.1080/03637751.2017.1322213

Dragojevic, M., Giles, H., & Watson, B. M. (2013). Language ideologies and language attitudes: A foundational framework. In *The Social Meanings of Language, Dialect and Accent: International Perspectives on Speech Styles* (pp. 1–25). Peter Lang Publishing.

Dragojevic, M., & Goatley-Soan, S. (2022). Americans' attitudes toward foreign accents: Evaluative hierarchies and underlying processes. *Journal of Multilingual and Multicultural Development*, *43*(2), 167–181. https://doi.org/10.1080/01434632.2020.1735402

Eychenne, J., & Jang, T.-Y. (2018). On the role of contrastivity in the development of the /e ∼ ɛ / merger in Korean. *Journal of East Asian Linguistics*, *27*(2), 111–139. https://doi.org/10.1007/s10831-018-9171-7

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Fuertes, J. N., Gottdiener, W. H., Martin, H., Gilbert, T. C., & Giles, H. (2012). A meta-analysis of the effects of speakers' accents on interpersonal evaluations. *European Journal of Social Psychology*, *42*(1), 120–133. https://doi.org/10.1002/ejsp.862

Garrett, P. (2010). *Attitudes to language*. Cambridge university press.

Giles, H. (1972). The Effect of Stimulus Mildness-Broadness in the Evaluation of Accents. *Language and Speech*, *15*(3), 262–269. https://doi.org/10.1177/002383097201500304

Giles, H., & Marlow, M. L. (2011). Theorizing Language Attitudes Existing Frameworks, an Integrative Model, and New Directions[1]. *Annals of the International Communication Association*, *35*(1), 161–197. https://doi.org/10.1080/23808985.2011.11679116

Giles, H., & Niedzielski, N. (1998). Italian is beautiful, German is ugly. In *Language myths* (pp. 85–93). Penguin.

Giles, H., & Watson, B. M. (Eds.). (2013). *The social meanings of language, dialect and accent: International perspectives on speech styles*. Lang.

Gnevsheva, K. (2018). The expectation mismatch effect in accentedness perception of Asian and Caucasian non-native speakers of English. *Linguistics*, *56*(3), 581–598. https://doi.org/10.1515/ling-2018-0006

Graillot, K. S., & Oh, E. (2025). Role of proficiency and visual cue in Korean listeners' perception of accented speech. *Phonetics and Speech Sciences*, *17*(2), 1–8. https://doi.org/10.13064/KSSS.2025.17.2.001

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Grondelaers, S., Van Hout, R., & Van Gent, P. (2019). Re-evaluating the Prestige of Regional Accents in Netherlandic Standard Dutch: The Role of Accent Strength and Speaker Gender. *Journal of Language and Social Psychology*, *38*(2), 215–236. https://doi.org/10.1177/0261927X18810730

Gunter, K. M., Vaughn, C. R., & Kendall, T. S. (2020). Perceiving Southernness: Vowel categories and acoustic cues in Southernness ratings. *The Journal of the Acoustical Society of America*, *147*(1), 643–656. https://doi.org/10.1121/10.0000550

Han, J.-I., & Kang, H. (2013). Cross-generational Change of /o/ and /u/ in Seoul Korean I: Proximity in Vowel Space. *Phonetics and Speech Sciences*, *5*(2), 25–31. https://doi.org/10.13064/KSSS.2013.5.2.025

Han, J.-I., & Kim. (2014). Acquisition of Korean stops by native speakers of Mandarin Chinese: A one-year longitudinal study. *Korean Journal of Linguistics*, *39*(2), 377–403. https://doi.org/10.18855/LISOKO.2014.39.2.007

Han, M. S., & Weitzman, R. S. (1970). Acoustic Features of Korean /P, T, K/, /p, t, k/ and /p^h , t^h , k^h /. *Phonetica*, *22*(2), 112–128. https://doi.org/10.1159/000259311

Hansen, K., Rakić, T., & Steffens, M. C. (2018). Foreign-Looking Native-Accented People: More Competent When First Seen Rather Than Heard. *Social Psychological and Personality Science*, *9*(8), 1001–1009. https://doi.org/10.1177/1948550617732389

Hansen, K., Steffens, M. C., Rakić, T., & Wiese, H. (2017). When appearance does not match accent: Neural correlates of ethnicity-related expectancy violations. *Social Cognitive and Affective Neuroscience*, *12*(3), 507–515. https://doi.org/10.1093/scan/nsw148

Hay, J., & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics*, *48*(4). https://doi.org/10.1515/ling.2010.027

Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, *34*(4), 458–484. https://doi.org/10.1016/j.wocn.2005.10.001

Hayes-Harb, R., & Barrios, S. (2021). The influence of orthography in second language phonological acquisition. *Language Teaching*, *54*(3), 297–326. https://doi.org/10.1017/S0261444820000658

Heaton, H., & Nygaard, L. C. (2011). Charm or Harm: Effect of Passage Content on Listener Attitudes Toward American English Accents. *Journal of Language and Social Psychology*, *30*(2), 202–211. https://doi.org/10.1177/0261927X10397288

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup Bias. *Annual Review of Psychology*, *53*(1), 575–604. https://doi.org/10.1146/annurev.psych.53.100901.135109

Hillenbrand, J., Getty, L. A., Wheeler, K., & Clark, M. J. (1994). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, *95*(5_Supplement), 2875–2875. https://doi.org/10.1121/1.409456

Hillenbrand, J. M., Clark, M. J., & Houde, R. A. (2000). Some effects of duration on vowel recognition. *The Journal of the Acoustical Society of America*, *108*(6), 3013–3022. https://doi.org/10.1121/1.1323463

Hogg, M. A. (2000). Subjective Uncertainty Reduction through Self-categorization: A Motivational Theory of Social Identity Processes. *European Review of Social Psychology*, *11*(1), 223–255. https://doi.org/10.1080/14792772043000040

Hogg, M. A., & Reid, S. A. (2006). Social Identity, Self-Categorization, and the Communication of Group Norms. *Communication Theory*, *16*(1), 7–30. https://doi.org/10.1111/j.1468-2885.2006.00003.x

Holliday, J. J. (2015). A longitudinal study of the second language acquisition of a three-way stop contrast. *Journal of Phonetics*, *50*, 1–14. https://doi.org/10.1016/j.wocn.2015.01.004

Hwang, Y. J. (2025). Who endorses online hate? The roles of ideology and knowledge in South Korean perceptions of anti-Chinese slurs. *Online Media and Global Communication*, *4*(3), 459–486. https://doi.org/10.1515/omgc-2025-0004

Idemaru, K., Wei, P., & Gubbins, L. (2019). Acoustic Sources of Accent in Second Language Japanese Speech. *Language and Speech*, *62*(2), 333–357. https://doi.org/10.1177/0023830918773118

Isbell, D. R., Lee, J., & Jang, J. (2024). Linguistic influences on comprehensibility and accentedness in second language Korean speech. *International Journal of Applied Linguistics*, *34*(4), 1575–1591. https://doi.org/10.1111/ijal.12580

Isbell, D. R., Park, O.-S., & Lee, K. (2019). Learning Korean pronunciation: Effects of instruction, proficiency, and L1. *Journal of Second Language Pronunciation*, *5*(1), 13–48. https://doi.org/10.1075/jslp.17010.isb

Iverson, P., & Evans, B. G. (2007). Learning English vowels with different first-language vowel systems: Perception of formant targets, formant movement, and duration. *The Journal of the Acoustical Society of America*, *122*(5), 2842–2854. https://doi.org/10.1121/1.2783198

Jadoul, Y., Thompson, B., & De Boer, B. (2018). Introducing Parselmouth: A Python interface to Praat. *Journal of Phonetics*, *71*, 1–15. https://doi.org/10.1016/j.wocn.2018.07.001

Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics*, *34*(4), 485–499. https://doi.org/10.1016/j.wocn.2005.08.004

Jung, H., & Holliday, J. J. (2026). The Perception of Foreign Accent in Isolated Korean Words. *International Journal of Applied Linguistics*, *36*(1), 667–676. https://doi.org/10.1111/ijal.12803

Jung, H. J., & Jeong, H. W. (2016). South Korean Attitude towards China: Threat Perception, Economic Interest, and National Identity. *African and Asian Studies*, *15*(2–3), 242–264. https://doi.org/10.1163/15692108-12341361

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. https://doi.org/10.1121/1.2166607

Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, *38*(2), 301–315. https://doi.org/10.1016/j.system.2010.01.005

Kang, O., & Rubin, D. L. (2009). Reverse Linguistic Stereotyping: Measuring the Effect of Listener Expectations on Speech Evaluation. *Journal of Language and Social Psychology*, *28*(4), 441–456. https://doi.org/10.1177/0261927X09341950

Kang, Y. (2014). A corpus-based study of positional variation in Seoul Korean vowels. *Japanese/Korean Linguistics*, *23*, 1–20.

Kang, Y., Yoon, T.-J., & Han, S. (2015). Frequency effects on the vowel length contrast merger in Seoul Korean. *Laboratory Phonology*, *6*(3–4). https://doi.org/10.1515/lp-2015-0014

Kessinger, R. H., & Blumstein, S. E. (1997). Effects of speaking rate on voice-onset time in Thai, French, and English. *Journal of Phonetics*, *25*(2), 143–168. https://doi.org/10.1006/jpho.1996.0039

Kim, H. (2023). Tracking the Time Course of Phonetic Cue Integration in the Perception of Korean Stop Contrasts by Korean and English Listeners [Ph.D.]. In *ProQuest Dissertations and Theses*. University of Kansas.

Kim, H., & Jongman, A. (2022). The influence of inter-dialect contact on the Korean three-way laryngeal distinction: An acoustic comparison among Seoul Korean speakers and Gyeongsang speakers with limited and extended residence in Seoul. *Language and Speech*, *65*(3), 531–553. https://doi.org/10.1177/00238309211037720

Kim, M. (2004). Correlation between VOT and F0 in the perception of Korean stops and affricates. *Proceedings of Interspeech 2004*, 49–52.

Kim, M.-R. C., & Lotto, A. J. (2002). An Investigation of Acoustic Characteristics of Korean Stops Produced by Non-Heritage Learners. *The Korean Language in America*, *7*, 177–187.

Kim, N. Y. (2008). *Imperial citizens: Koreans and race from Seoul to LA*. Stanford university press.

Kinzler, K. D., & DeJesus, J. M. (2013). Northern = smart and Southern = nice: The development of accent attitudes in the United States. *Quarterly Journal of Experimental Psychology*, *66*(6), 1146–1158. https://doi.org/10.1080/17470218.2012.731695

Kinzler, K. D., Shutts, K., DeJesus, J., & Spelke, E. S. (2009). Accent Trumps Race in Guiding Children's Social Preferences. *Social Cognition*, *27*(4), 623–634. https://doi.org/10.1521/soco.2009.27.4.623

Kong, E. J., Beckman, M. E., & Edwards, J. (2011). Why are Korean tense stops acquired so early?: The role of acoustic properties. *Journal of Phonetics*, *39*(2), 196–211. https://doi.org/10.1016/j.wocn.2011.02.002

Kong, E. J., Kang, S., & Seo, M. (2022). The acoustic cue-weighting and the L2 production-perception link: A case of English-speaking adults' learning of Korean stops. *Phonetics and Speech Sciences*, *14*(3), 1–9. https://doi.org/10.13064/KSSS.2022.14.3.001

Koops, C. (2010). */u/-Fronting is not Monolithic: Two Types of Fronted /u/ in Houston Anglos*.

Koops, C., Gentry, E., & Pantos, A. (2008). The effect of perceived speaker age on the perception of PIN and PEN vowels in Houston, Texas. *University of Pennsylvania Working Papers in Linguistics*, *14(2)*, 12.

Korea Immigration Service. (2024). *KIS Statistics 2024*. http://www.immigration.go.kr/immigration/3361/subview.do?enc=Zm5jdDF8QEB8JTJG

YmJzJTJGaW1taWdyYXRpb24lMkY0ODklMkY1ODAyMTYlMkZhcnRjbFZpZXcuZG8lM0ZwYXNzd29yZCUzRCUyNnJnc0JnbmRlU3RyJTNEJTI2YmJzQ2xTZXElM0QlMjZyZ3NFbmRkZVN0ciUzRCUyNmlzVmlld01pbmUlM0RmYWxzZSUyNnBhZ2UlM0QxJTI2YmJzT3BlbldyZFNlcSUzRCUyNnNyY2hDb2x1bW4lM0QlMjZzcmNoV3JkJTNEJTI2

Kutlu, E. (2023). Now You See Me, Now You Mishear Me: Raciolinguistic accounts of speech perception in different English varieties. *Journal of Multilingual and Multicultural Development*, *44*(6), 511–525. https://doi.org/10.1080/01434632.2020.1835929

Kutlu, E., Tiv, M., Wulff, S., & Titone, D. (2022). Does race impact speech perception? An account of accented speech in two different multilingual locales. *Cognitive Research: Principles and Implications*, *7*(1), 7. https://doi.org/10.1186/s41235-022-00354-0

Labov, W. (1966). *The social stratification of English in New York City*.

Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *The Journal of Abnormal and Social Psychology*, *60*(1), 44–51. https://doi.org/10.1037/h0044430

Lee, H. (2016a). A Comprehensive Study on the Language Attitudes of Native Korean Speakers toward Foreign Accented Korean Language Speakers. *International Network for Korean Language and Culture*, *13*(3), 199–229. https://doi.org/10.15652/ink.2016.13.3.199

Lee, H. (2016b). Consistent sound change between stops and affricates in Seoul Korean within and across individuals: A diachronic investigation. *The Journal of the Acoustical Society of America*, *140*(6), EL491–EL496. https://doi.org/10.1121/1.4971203

Lee, H., Holliday, J. J., & Kong, E. J. (2020). Diachronic change and synchronic variation in the Korean stop laryngeal contrast. *Language and Linguistics Compass*, *14*(7), e12374. https://doi.org/10.1111/lnc3.12374

Lee, Y. (2025). Phonetic aspects of the vowels /o/ and /u/ in Seoul Korean. *Linguistic Research*, *42*(3), 657–684. https://doi.org/10.17250/KHISLI.42.3.202512.007

Lee, Y., & Cho, S. (2021). The Merger of /e/ and /ɛ/ in Seoul Korean Tracing Speakers Born from the 1930s to the 1960s. *Studies in Linguistics*, *61*, 1–16. https://doi.org/10.17002/sil..61.202110.1

Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, *46*(6), 1093–1096. https://doi.org/10.1016/j.jesp.2010.05.025

Lippi-Green, R. (2012). *English with an Accent: Language, ideology, and discrimination in the United States* (2nd ed.). Routledge. https://doi.org/10.4324/9780203348802

Lisker, L., & Abramson, A. S. (1964). A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. WORD, *20*(3), 384–422. https://doi.org/10.1080/00437956.1964.11659830

Lo, A., & Kim, J. (2011). Manufacturing citizenship: Metapragmatic framings of language competencies in media images of mixed race men in South Korea. *Discourse & Society*, *22*(4), 440–457. https://doi.org/10.1177/0957926510395834

Lobanov, B. M. (1971). Classification of Russian Vowels Spoken by Different Speakers. *The Journal of the Acoustical Society of America*, *49*(2B), 606–608. https://doi.org/10.1121/1.1912396

Long, M. H. (1990). Maturational Constraints on Language Development. *Studies in Second Language Acquisition*, *12*(3), 251–285. https://doi.org/10.1017/S0272263100009165

Loureiro-Rodriguez, V., Boggess, M. M., & Goldsmith, A. (2013). Language attitudes in Galicia: Using the matched-guise test among high school students. *Journal of Multilingual and Multicultural Development*, *34*(2), 136–153. https://doi.org/10.1080/01434632.2012.729591

Maastricht, L. V., Zee, T., Krahmer, E., & Swerts, M. (2017). L1 Perceptions of L2 Prosody: The Interplay Between Intonation, Rhythm, and Speech Rate and Their Contribution to Accentedness and Comprehensibility. *Interspeech 2017*, 364–368. https://doi.org/10.21437/Interspeech.2017-1150

Magen, H. S., & Blumstein, S. E. (1993). Effects of speaking rate on the vowel length distinction in Korean. *Journal of Phonetics*, *21*(4), 387–409. https://doi.org/10.1016/S0095-4470(19)30227-X

Mareüil, P. B. D., & Vieru-Dimulescu, B. (2006). The Contribution of Prosody to the Perception of Foreign Accent. *Phonetica*, *63*(4), 247–267. https://doi.org/10.1159/000097308

Mattys, S. L., & Palmer, S. D. (2015). Divided attention disrupts perceptual encoding during speech recognition. *The Journal of the Acoustical Society of America*, *137*(3), 1464–1472. https://doi.org/10.1121/1.4913507

McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., & Sonderegger, M. (2017). Montreal Forced Aligner: Trainable Text-Speech Alignment Using Kaldi. *Interspeech 2017*, 498–502. https://doi.org/10.21437/Interspeech.2017-1386

McCroskey, J. C., Richmond, V. P., & Daly, J. A. (1975). THE DEVELOPMENT OF A MEASURE OF PERCEIVED HOMOPHILY IN INTERPERSONAL

COMMUNICATION. *Human Communication Research*, *1*(4), 323–332. https://doi.org/10.1111/j.1468-2958.1975.tb00281.x

McCullough, E. (2011). Acoustic correlates of perceptual ratings of foreign-accented English. *The Journal of the Acoustical Society of America*, *129*(4_Supplement), 2453–2453. https://doi.org/10.1121/1.3588056

McGowan, K. B. (2015). Social Expectation Improves Speech Perception in Noise. *Language and Speech*, *58*(4), 502–521. https://doi.org/10.1177/0023830914565191

McLaughlin, D. J., Colvett, J. S., Bugg, J. M., & Van Engen, K. J. (2024). Sequence effects and speech processing: Cognitive load for speaker-switching within and across accents. *Psychonomic Bulletin & Review*, *31*(1), 176–186. https://doi.org/10.3758/s13423-023-02322-1

Melguy, Y. V., & Johnson, K. (2021). General adaptation to accented English: Speech intelligibility unaffected by perceived source of non-native accent. *The Journal of the Acoustical Society of America*, *149*(4), 2602–2614. https://doi.org/10.1121/10.0004240

Milroy, J. (2006). The ideology of the standard language. In *The Routledge companion to sociolinguistics* (pp. 133–139). Routledge.

Milroy, J., & Milroy, L. (1999). *Authority in language: Investigating standard English* (3rd ed). Routledge.

Mulac, A., Giles, H., Bradac, J. J., & Palomares, N. A. (2013). The gender-linked language effect: An empirical test of a general process model. *Language Sciences*, *38*, 22–31. https://doi.org/10.1016/j.langsci.2012.12.004

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. https://doi.org/10.1177/002383099303600103

Munro, M. J., & Derwing, T. M. (1995). Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners. *Language Learning*, *45*(1), 73–97. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x

Munro, M. J., & Derwing, T. M. (2001). MODELING PERCEPTIONS OF THE ACCENTEDNESS ANDCOMPREHENSIBILITY OF L2 SPEECH The Role of Speaking Rate. *Studies in Second Language Acquisition*, *23*(4), 451–468. https://doi.org/10.1017/S0272263101004016

National Institute of Korean Language. (2026). *표준어 규정*. [Korean Language Norms]. National Institute of Korean Language. https://korean.go.kr/kornorms/regltn/regltnView.do?regltn_code=0002

Nearey, T. M., & Rochet, B. L. (1994). Effects of Place of Articulation and Vowel Context on VOT Production and Perception for French and English Stops. *Journal of the International Phonetic Association*, *24*(1), 1–18. https://doi.org/10.1017/S0025100300004965

Niedzielski, N. (1999). The Effect of Social Information on the Perception of Sociolinguistic Variables. *Journal of Language and Social Psychology*, *18*(1), 62–85. https://doi.org/10.1177/0261927X99018001005

Oh, E. (2008). Voice onset time of Korean stops as a function of speaking rate. *The Journal of the Acoustical Society of America*, *124*(4_Supplement), 2556–2556. https://doi.org/10.1121/1.3020581

Oppenheimer, D. M. (2006). Consequences of erudite vernacular utilized irrespective of necessity: Problems with using long words needlessly. *Applied Cognitive Psychology*, *20*(2), 139–156. https://doi.org/10.1002/acp.1178

Park, K. (2011). White supremacy hidden behind the current trend of mixed-race in the entertainment business. *Kyunghyang*. https://www.khan.co.kr/article/201108212119105

Park, K. (2014). Foreigners or multicultural citizens? Press media's construction of immigrants in South Korea. *Ethnic and Racial Studies*, *37*(9), 1565–1586. https://doi.org/10.1080/01419870.2012.758860

Pettigrew, T. F. (1998). INTERGROUP CONTACT THEORY. *Annual Review of Psychology*, *49*(1), 65–85. https://doi.org/10.1146/annurev.psych.49.1.65

Polyanskaya, L., Ordin, M., & Busa, M. G. (2017). Relative Salience of Speech Rhythm and Speech Rate on Perceived Foreign Accent in a Second Language. *Language and Speech*, *60*(3), 333–355. https://doi.org/10.1177/0023830916648720

Powesland, P., & Giles, H. (1975). Persuasiveness and Accent-Message Incompatibility. *Human Relations*, *28*(1), 85–93. https://doi.org/10.1177/001872677502800105

Preston, D. R. (2013). Language with an Attitude. In J. K. Chambers & N. Schilling (Eds.), *The Handbook of Language Variation and Change* (1st ed., pp. 157–182). Wiley. https://doi.org/10.1002/9781118335598.ch7

Purnell, T., Idsardi, W., & Baugh, J. (1999). Perceptual and Phonetic Experiments on American English Dialect Identification. *Journal of Language and Social Psychology*, *18*(1), 10–30. https://doi.org/10.1177/0261927X99018001002

R Core Team. (2026). *R: A Language and Environment for Statistical Computing* (Version 4.5.3) [Computer software]. R Foundation for Statistical Computing. http://www.R-project.org/

Rahman, J. (2004). *It's a Serious Business: The Linguistic Construction of Middle-Class White Characters by African American Narrative Comedians.* [Ph.D. Dissertation]. Stanford University.

Rakić, T., Steffens, M. C., & Mummendey, A. (2011). Blinded by the accent! The minor role of looks in ethnic categorization. *Journal of Personality and Social Psychology*, *100*(1), 16–29. https://doi.org/10.1037/a0021522

Redford, M. A., Kapatsinski, V., & Cornell-Fabiano, J. (2018). Lay Listener Classification and Evaluation of Typical and Atypical Children's Speech. *Language and Speech*, *61*(2), 277–302. https://doi.org/10.1177/0023830917717758

Revelle, W. (2026). *psych: Procedures for Psychological, Psychometric, and Personality Research* (Version 2.6.1, p. 2.6.1) [Dataset]. https://doi.org/10.32614/CRAN.package.psych

Rickford, J. R. (1985). Standard and Non-Standard Language: Attitudes in a Creole Continuum. In *Language of Inequality* (pp. 145–160).

Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education*, *33*(4), 511–531. https://doi.org/10.1007/BF00973770

Rubin, D. L., & Smith, K. A. (1990). Effects of accent, ethnicity, and lecture topic on undergraduates' perceptions of nonnative English-speaking teaching assistants. *International Journal of Intercultural Relations*, *14*(3), 337–353. https://doi.org/10.1016/0147-1767(90)90019-S

Ryan, E. B., & Giles, H. (Eds.). (1982). *Attitudes towards language variation: Social and applied contexts*. E. Arnold.

Saito, K., Trofimovich, P., & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics*, *37*(2), 217–240. https://doi.org/10.1017/S0142716414000502

Sereno, J., Lammers, L., & Jongman, A. (2016). The relative contribution of segments and intonation to the perception of foreign-accented speech. *Applied Psycholinguistics*, *37*(2), 303–322. https://doi.org/10.1017/S0142716414000575

Shamoa-Nir, L., Razpurker-Apfeld, I., Dautel, J. B., & Taylor, L. K. (2021). Out-group prosocial giving during childhood: The role of in-group preference and out-group attitudes in a divided society. *International Journal of Behavioral Development*, *45*(4), 337–344. https://doi.org/10.1177/0165025420935619

Shimada, M. (2024). *Japanese Prosody and Its Role in Comprehensibility: Exploration and Training*. Graduate Studies. https://doi.org/10.11575/PRISM/43402

Shuck, G. (2006). Racializing the Nonnative English Speaker. *Journal of Language, Identity & Education*, *5*(4), 259–276. https://doi.org/10.1207/s15327701jlie0504_1

Silva, D. J. (2006). *Acoustic evidence for the emergence of tonal contrast in contemporary Korean*. *Phonology*, *23*(2), 287–308. https://doi.org/10.1017/S0952675706000911

Sim, S., & Kim, S.-A. (2025). Explicit or Implicit Instruction? Hangul Reading Instruction for Beginning KFL Learners. In H. K. Pae, H. Winskel, & S. Y. Kim (Eds.), *Handbook on the Korean Language and Literacy: Insights into Hangul and Text Processing* (pp. 435–455). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-88087-2_22

Solé, M.-J., & Ohala, J. J. (2010). What is and what is not under the control of the speaker: Intrinsic vowel duration. In C. Fougeron, B. Kühnert, M. D'Imperio, & N. Vallée (Eds.),

Laboratory Phonology 10 (pp. 607–656). DE GRUYTER MOUTON. https://doi.org/10.1515/9783110224917.5.607

Song, E. E. (2023). COVID-19, Anti-Chinese Sentiment, and Foreign Policy Attitudes in South Korea. *Asian Survey*, *63*(5), 823–850. https://doi.org/10.1525/as.2023.2008558

Song, J., Kim, B., Kim, M., & Iverson, P. (2023). The Korean Speech Recognition Sentences: A Large Corpus for Evaluating Semantic Context and Language Experience in Speech Perception. *Journal of Speech, Language, and Hearing Research*, *66*(9), 3399–3412. https://doi.org/10.1044/2023_JSLHR-23-00137

Strand, E. A. (1999). Uncovering the Role of Gender Stereotypes in Speech Perception. *Journal of Language and Social Psychology*, *18*(1), 86–100. https://doi.org/10.1177/0261927X99018001006

Strand, E. A. (with Johnson, K.). (2000). *Gender stereotype effects in speech processing*. Ohio State University.

Strand, E. A., & Johnson, K. (1996). Gradient and Visual Speaker Normalization in the Perception of Fricatives. *KONVENS*, 14–26.

Subtirelu, N. C. (2015). "She does have an accent but…": Race and language ideology in students' evaluations of mathematics instructors on RateMyProfessors.com. *Language in Society*, *44*(1), 35–62. https://doi.org/10.1017/S0047404514000736

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In *Psychology of intergroup relations* (pp. 7–24). Nelson-Hall.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Thomas, E. R., & Kendall, T. (2007). *NORM: The vowel normalization and plotting suite*.

Thomas, E. R., & Wassink, A. B. (2010). Variation and Identity in African-American English. In *Language and Identities* (pp. 157–165).

Trofimovich, P., & Baker, W. (2006). LEARNING SECOND LANGUAGE SUPRASEGMENTALS: Effect of L2 Experience on Prosody and Fluency Characteristics of L2 Speech. *Studies in Second Language Acquisition*, *28*(01). https://doi.org/10.1017/S0272263106060013

Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, *15*(4), 905–916. https://doi.org/10.1017/S1366728912000168

Trofimovich, P., Isaacs, T., Kennedy, S., & Tsunemoto, A. (2022). Speech Comprehensibility. In *The Routledge handbook of second language acquisition and speaking* (pp. 174–187). Routledge.

Van Engen, K. J., & Peelle, J. E. (2014). Listening effort and accented speech. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00577

Varonis, E. M., & Gass, S. (1982). The comprehensibility of non-native speech. *Studies in Second Language Acquisition*, *4*(2), 114–136. https://doi.org/10.1017/S027226310000437X

Vaughn, C. R. (2019). Expectations about the source of a speaker's accent affect accent adaptation. *The Journal of the Acoustical Society of America*, *145*(5), 3218–3232. https://doi.org/10.1121/1.5108831

Volaitis, L. E., & Miller, J. L. (1992). Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories. *The Journal of the Acoustical Society of America*, *92*(2), 723–735. https://doi.org/10.1121/1.403997

Wade, L. (2025). In-Group/Out-Group Dynamics, Contrast, and the Listening Subject in Sociolinguistic Perception. *Journal of Sociolinguistics*, *29*(5), 371–375. https://doi.org/10.1111/josl.12724

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Wayland, R. (1997). Non-native Production of Thai: Acoustic Measurements and Accentedness Ratings. *Applied Linguistics*, *18*(3), 345–373. https://doi.org/10.1093/applin/18.3.345

Winters, S., & O'Brien, M. G. (2013). Perceived accentedness and intelligibility: The relative contributions of F0 and duration. *Speech Communication*, *55*(3), 486–507. https://doi.org/10.1016/j.specom.2012.12.006

Wood, S. N. (2025). Generalized Additive Models. *Annual Review of Statistics and Its Application*, *12*(1), 497–526. https://doi.org/10.1146/annurev-statistics-112723-034249

World Bank. (2024). *GDP (constant 2025 US$)—United States, China, Korea, Rep.* https://data.worldbank.org/indicator/NY.GDP.MKTP.KD?locations=US-CN-KR

Wright, S. C., Aron, A., McLaughlin-Volpe, T., & Ropp, S. A. (1997). The extended contact effect: Knowledge of cross-group friendships and prejudice. *Journal of Personality and Social Psychology*, *73*(1), 73–90. https://doi.org/10.1037/0022-3514.73.1.73

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. https://doi.org/10.1006/jpho.1996.0013

Yi, H.-G., Phelps, J. E. B., Smiljanic, R., & Chandrasekaran, B. (2013). Reduced efficiency of audiovisual integration for nonnative speech. *The Journal of the Acoustical Society of America*, *134*(5), EL387–EL393. https://doi.org/10.1121/1.4822320

Zebrowitz, L. A., Bronstad, P. M., & Lee, H. K. (2007). The Contribution of Face Familiarity to Ingroup Favoritism and Stereotyping. *Social Cognition*, *25*(2), 306–338. https://doi.org/10.1521/soco.2007.25.2.306