AN EXPLORATORY STUDY OF FOREIGN ACCENT AND PHONOLOGICAL AWARENESS IN KOREAN LEARNERS OF ENGLISH

by

Mi Sun Park

Dissertation Committee:

Professor ZhaoHong Han, Sponsor
Professor Erika Shield Levy

Approved by the Committee on the Degree of Doctor of Education

Date 22 May 2019

Submitted in partial fulfillment of the
requirements for the Degree of Doctor of Education in
Teachers College, Columbia University

2019

ABSTRACT

AN EXPLORATORY STUDY OF FOREIGN ACCENT AND PHONOLOGICAL AWARENESS IN KOREAN LEARNERS OF ENGLISH

Mi Sun Park

Communication in a second or multiple languages has become essential in the globalized world. However, acquiring a second language (L2) after a critical period is universally acknowledged to be challenging (Lenneberg, 1967). Late learners hardly reach a nativelike level in L2, particularly in its pronunciation, and their incomplete phonological acquisition is manifested by a *foreign accent*—a common and persistent feature of otherwise fluent L2 speech. Although foreign-accented speech is widespread, it has been a target of social constraints in L2-speaking communities, causing many learners and instructors to seek out ways to reduce foreign accents. Accordingly, research in L2 speech has unceasingly examined various learner-external and learner-internal factors of the occurrence of foreign accents as well as nonnative speech characteristics underlying the judgment of the degree of foreign accents. The current study aimed to

expand the understanding of the characteristics and judgments of foreign accents by investigating *phonological awareness*, a construct pertinent to learners' phonological knowledge, which has received little attention in research on foreign accents.

The current study was exploratory and non-experimental research that targeted 40 adults with Korean-accented English living in the United States. The study first examined how 23 raters speaking American English as their native language detect, perceive, describe, and rate Korean-accented English. Through qualitative and quantitative analyses of the accent perception data, the study identified various phonological and phonetic deviations from the nativelike sounds, which largely result from the influence of first language (Korean) on L2 (English). The study then probed the relationship between foreign accents and learners' awareness of the phonological system of L2, which was measured using production, perception, and verbalization tasks that tapped into the knowledge of L2 phonology. The study found a significant inverse relationship between the degree of a foreign accent and phonological awareness, particularly implicit knowledge of L2 segmentals. Further in-depth analyses revealed that explicit knowledge of L2 phonology alone was not sufficient for targetlike pronunciation. Findings suggest that L2 speakers experience varying degrees of difficulty in perceiving and producing different L2 segmentals, possibly resulting in foreign-accented speech.

ACKNOWLEDGMENTS

I am indebted to many individuals who have supported my doctoral studies. First and foremost, I would like to thank my advisor and dissertation sponsor, Dr. ZhaoHong Han, for her consistent support, guidance, and encouragement. Without her help, I could not have completed, or even started, my doctoral studies. I was inspired by her passion for second language research, and every day with her has been a learning opportunity. I also thank Dr. Erika Levy for her warm support and valuable comments on my dissertation proposal and the final dissertation. A special thank you goes to Dr. Howard Williams, the chair of the dissertation committee. I have learned so many things from him, about linguistics, languages, life, and the world, and I will greatly miss our conversations in his office. I would also like to thank Drs. Harold Abeles and Hoa Nguyen for carefully reading my dissertation and providing insightful comments.

My sincere thanks go to the members of my doctoral seminar, who have helped me tremendously throughout my doctoral studies and provided thoughtful comments on the dissertation: Farah Akbar, Kaylee Fernandez, Rosette Finneran, Ji-Yung Jung, Peter Kim, Adrienne Lew, Shaoyan Qi, Natalia Sáez, Haimei Sun, Yayun Anny Sun, and Rainie Zhang. I am also thankful to Shafinaz Ahmed, Alice Chen, EunYoung Kang, and Sarah Sok. Our teamwork on varying projects was just wonderful, and I will miss all the moments of laughter and tears that we had together.

This dissertation, particularly its data collection, would not have been possible without the help of many participants and my colleagues at the Community Language Program at Teachers College. I am also grateful to Haena Kim at State University of New

York and EunYoung Kang at Lehman College of the City University of New York for coordinating meetings with participants. I would also like to thank Yun-Hee Kim for helping me with creating the instruments for the study, Fred Tsutagawa for piloting the instruments from a tester's perspective, and Jason Bennett for sharing a non-linguist's insights. I also express my gratitude for the Dean's Grant for Student Research and the Provost's Grant for Professional Development, both from Teachers College, Columbia University.

Two important men passed away while I was doing my doctoral studies: my grandfather gave me so much love and support even until his last day, and the other person, whom I leave unnamed, devoted his life to spreading the love of God. I dearly miss them and will always remember their strength and faith.

My deepest gratitude goes to my mom and dad, who have patiently and unconditionally supported my study for many years. No words can express my gratitude and love for them. This dissertation is dedicated to them. My younger brother Jae Moon has been a cheerful and thoughtful friend all the time, and I wish him all the best luck in the future. I also thank my grandmother, parents-in-law, and my aunts, uncles, and cousins, for their endless love and support. Last but not least, I thank my husband and bestest friend Jungho Kim for his love, patience, and understanding. I can't wait to open the next chapter of my life with you (and Yaong)!

MSP

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I – INTRODUCTION

Globalization and multilingualism have highlighted the role of language, which is indispensable not only in linking nations, but also in empowering individuals toward greater mobility. Worldwide, many people learn a language in addition to their mother tongue for purposes such as education, business, travel, and even personal joy. Second language acquisition (SLA) has thus become a widespread phenomenon, and an increasing volume of research has delved into its processes, factors, assessments, and pedagogical implications.

However, it is often a challenge for an individual to learn another language as a *late* learner—that is, to master a range of linguistic features including grammar, semantics, phonology, and pragmatics in a *new* language after early childhood. While only few late L2 learners reach native-level proficiency, cessation of learning, or fossilization, is a more common phenomenon in L2 acquisition (Han, 2004; Long, 2003; Selinker, 1972). The challenge is particularly greater for acquisition of phonology (Flege, 1981; Major, 2001) because most second language (L2) learners experience difficulty in targetlike L2 speech production and thus "never seem capable of ridding themselves entirely of a foreign accent" (Scovel, 1969, p. 245). Even for advanced, fluent L2 speakers, a foreign accent is easily detected in their speech. The special difficulty of L2 phonological acquisition is often exemplified by the writer Joseph Conrad, whose written works have become classics of English literature, while his English speech is reported to

have retained the heavy accent of his first language, Polish (Bongaerts, Planken, & Schils, 1995; Yule, 2014).

The widely observed cases of learners who retain a foreign accent often discourage many language teachers and late learners, even though cases of rare, exceptional learners (that is, adult learners who have achieved high L2 pronunciation proficiency without any trace of foreign accent) have been reported in some studies (e.g., Bongaerts, 1999; Moyer, 1999, 2004). The elimination of a foreign accent has become a goal that is neither realistic nor desirable in pronunciation instruction (Abercrombie, 1949; Derwing & Munro, 2005, 2009). Moreover, communicative language teaching, the dominant pedagogical approach nowadays, emphasizes phonological fluency rather than discrete-point accuracy, which was the core of language curricula in the days of audiolingualism (Pennington & Richards, 1986). The recent focus of phonological instruction in communicatively-oriented language learning settings is to promote the learner's ability to "negotiate for meaning" (Dalton & Seidlhofer, 1994). This approach can help the learner appropriately convey the meaning to complete a specific speech act, rather than holding him or her to a nativelike standard.

In the case of the English language, its current state as an international language or a lingua franca (i.e., spoken by nonnative speakers as much as (or even more than) by native speakers) has led researchers to argue that pronunciation instruction should aim for mutual intelligibility or comprehensibility, rather than paying attention to nonnativelike phonological forms that do not cause serious communication breakdowns (e.g., Jenkins, 2002). Consequently, L2 or foreign language (FL) pronunciation teaching has evolved to focus on the "aspects of the learner's speech that appear to interfere with listener's

understanding" (Munro & Derwing, 1995, p. 93) instead of achieving native-level accuracy (see Davies, 2003, for the "myth" of the native speaker).

Nevertheless, foreign accents remain a core topic in L2 speech research, also evident in a large number of accent reduction programs available in the market. There are several reasons for this trend. First, the salience of accent is immediate and real, signaling to the listener that the speaker is nonnative (Moyer, 2015). Listeners in general consistently point out that a foreign accent is a primary aspect of their comprehension of nonnative speech and communication with nonnative speakers (NNSs) regardless of the speakers' actual proficiency (Hayes-Harb & Hacking, 2015). Therefore, accentedness, defined as the perceived degree of an accent, is a frequently used measure in the current L2 speech literature (Munro & Derwing, 2011).

Furthermore, nativelike attainment still remains a goal (although it is often argued to be unrealistic) for a large population of highly dedicated and motivated L2/FL learners and instructors. The motivation of individuals to attain nativelike performance may vary. However, for some learners, nativelike-ness may be a means of overcoming constraints imposed on them by some external force such as persistent prejudice, intolerance, or discrimination against foreign accents, despite unceasing efforts on the part of sociolinguists and educators. A foreign accent is often seen as a kind of deficit or disorder suitable for treatment or intervention (Munro & Derwing, 1999). Studies have also shown that "sounding foreign" might reduce the credibility of nonnative speakers (Lev-Ari & Keysar, 2010), or might affect the attitudes of listeners toward nonnative speech/speakers (Hayes-Harb & Hacking, 2015; Munro, 2003) or toward even nonnative L2/FL teachers (Choi, 2007). These challenges have the potential to constrain interaction that plays a key

role in L2 development (Long, 1996). A strong foreign accent can be a barrier to advancement if it discourages learners from seeking out a connection to the target language community, whereas learners with a more authentic accent tend to have greater confidence in their L2 abilities and to pursue more contact with native speakers, thereby increasing L2 use and fluency (Moyer, 2014, 2015). Therefore, it remains a crucial goal for many L2 learners to accommodate their listeners and optimize desirable social distance from them by achieving norms for pronunciation. Accordingly, the psychological and social significance of accent has driven L2 speech researchers and instructors to elucidate foreign accent phenomena and investigate a complex interplay of internal and external factors. This trend will hardly fade away as society becomes more multilingual and multicultural.

Successful acquisition of L2 phonology and phonetics depends on how L2 learners process L2 input and whether they establish an L2 phonological system that is new to (or "different from") their first language (L1) system (e.g., Flege, 1995; Lado, 1957; Munro, Flege, & MacKay, 1996). Phonological processing occurs in two parts (Loewen, 2014): (a) perception, which involves auditory and cognitive processing to connect sound to meaning, and (b) production, which involves neuro-muscular control of the articulatory organs. Much research indicates that perception is the basis for establishing mental representations of L2 sounds, and further on, producing those sounds (e.g., Flege, 1995). Late L2 learners do not perceive (and likely do not produce) L2 sounds in a nativelike way, and thus have difficulties acquiring the full range of phonological, phonetic, and prosodic features of an L2. As a result, they produce nontargetlike pronunciation that contributes to accentedness (Major, 2001). Foreign

accent research has examined accentedness judgments in relation to a wide range of phonological and phonetic features of nonnative speech manifested in the major domains of pronunciation: (a) segmentals, involving consonants and vowels, and (b) prosody or suprasegmentals, which involves stress and intonation patterns as well as timing and rhythm (e.g., Anderson-Hsieh, Johnson, & Koehler, 1992; Trofimovich & Baker, 2006). However, it remains uncertain which phonological features of nonnative speech are particularly crucial for accent phenomena (Saito, 2011). Furthermore, much work is needed to analytically describe and characterize a language-specific accent (e.g., a Korean accent of English). Such information is essential for developing adequate contents and criteria upon which future L2 pronunciation assessments and curricula should be based.

Numerous studies have investigated factors that contribute to the occurrence and to the perceived degree of a foreign accent. Researchers have suggested that the age of learners at the onset of L2 learning or exposure (AOL) has a strong impact on L2 speech learning (Lenneberg, 1967; Scovel, 2000). A great deal of attention has also been given to variables related to the quality and quantity of cumulative language experience and input. These variables include the length of residence (LOR) in a community in which the L2 is the predominant language (Trofimovich & Baker, 2006); the extensive use of an L2 (Flege, Yeni-Komshian, & Liu, 1999); the inadequate L2 phonetic input (Flege, Takagi, & Mann, 1995); the continued and frequent use of an L1 (Piske, MacKay, & Flege, 2001); and L2 instruction, particularly L2 pronunciation instruction (Derwing, Munro, & Wiebe, 1998). However, a foreign accent also results from any number of factors inherent to individual learners, including socio-psychological factors such as motivation and

attitudes towards L2 (Moyer, 1999) and cognitive abilities such as aptitude and talent (Baker-Smemoe & Haslam, 2013). For instance, despite the reported and significant influence of the AOL, accentedness varies considerably across L2 speakers regardless of their age. Some late learners speak with no detectable foreign accent (e.g., Bongaerts, 1999), while other learners do not achieve native-level competence despite having started L2 acquisition well within the critical period (e.g., Flege, 1987).

A far smaller volume of foreign accent research has investigated individual speakers' awareness of the phonological system of the L2. The construct of awareness has been a core topic of discussions in L2 acquisition and instruction. A general position among researchers is that learning takes place when input is consciously registered through awareness, or *noticed* (Schmidt, 1990). When it comes to learning pronunciation of a language, a learner is required to monitor his or her own speech and notice the target L2 pronunciation features while making adjustments to productions in real time (Derwing, Thomson, Foote, & Munro, 2012). Studies have suggested that instructional environments designed to direct the learners' attention to L2 phonological input may positively affect L2 pronunciation (e.g., Kissling, 2013; Lambacher, Martens, Kakehi, Marasinghe, & Molholt, 2005) although opportunities for explicit pronunciation instruction and feedback are not available to most learners. It is important to note that the construct of awareness in these studies was often not clearly defined or measured, and the construct even included non-phonological domains such as paralinguistics. Studies have thus made distinctions among the domains of language awareness. One key aspect of language awareness is phonological awareness, which involves the knowledge of the phonological system of a language. Phonological awareness is reflected in the ability of

an individual to analyze and manipulate the phonological system, and this ability is exhibited through the explanation or use of phonology (e.g., Snow, Burns, & Griffin, 1998).

Previous studies on phonological awareness and L2 phonology have noted several important points. Most notably, the majority of studies employed young participants—that is, children whose L2 phonological system, or even L1 system, was not fully developed. Late L2 learners, on the other hand, have received relatively little attention. This suggests that the current understanding of phonological awareness may not be sufficient to account for the phonological acquisition of learners after childhood. This is also surprising, considering the volume of research that delves into L2 pronunciation instruction, which is largely based on the idea that raising learners' awareness of the L2 phonology is essential for L2 perception and production (e.g., Derwing et al., 1998; Lambacher et al., 2005). The levels of awareness examined in previous studies ranged from perception, which requires processing as simple as mere detection of the target sounds, to understanding, which involves the explicit explanation of the target sounds or sound systems. Yet, many studies focused on the highest level of phonological awareness only, requiring participants to express their conscious knowledge of L2 phonology, often with the use of metalanguage (most likely acquired through instruction), while neglecting implicit phonological knowledge that develops naturally through language contact and does not require instruction. Furthermore, the observed phonological system in these studies tended to be limited to the level of discrete phonemes—the smallest distinguishable auditory units in words. The tasks, therefore, mainly involved discrimination and manipulation of contrastive phonemes. Not until recently did

phonological awareness studies pay attention to prosodic features such as word stress (e.g., Venkatagiri & Levis, 2007).

Accordingly, foreign accent phenomena in L2 speech data have been evaluated mostly in terms of segmental accuracy, which is necessary, but not sufficient, for nativelike L2 pronunciation. Rarely have foreign accent features been examined in terms of how listeners receive or analyze them. In other words, the extent to which the L2 speech sounds foreign-accented to the listeners, beyond whether it is segmentally and prosodically accurate, has received little attention. The relationship between the learners' accent and phonological awareness thus remains to be elucidated. In light of these issues, a study is needed that conducts a more comprehensive, multi-level investigation of how L2 speech is received and perceived by listeners. Such a study would complement the limitations reported in the previous research.

1.1. Focus of the Study

The current dissertation study is motivated by the identified gap in foreign accent research: the lack of systematic research on the awareness of L2 phonology of late L2 learners and on the relationship between L2 phonological awareness and foreign accent. The first aim of this study is to advance the understanding of the concept of foreign accent as a characteristic of L2 speech by investigating the phonological and phonetic variables that underlie native listeners' perception of foreign accent. The second aim of this study is to explore the relationship between foreign accent and L2 phonological awareness. This study investigates whether L2 speakers exhibiting different degrees of foreign accent vary in their performance on L2 phonological awareness tasks by

administering tasks designed to tap into nonnative speakers' implicit and explicit knowledge of L2 phonology. L2 learners who have superior phonological awareness are hypothesized to be better at noticing the gap between nativelike and nonnativelike pronunciations and also at manipulating their pronunciation to eliminate the gap (i.e., the foreign accent). However, the relationship between the two needs to be investigated before drawing any conclusion about the influence of one on the other. Therefore, this study proposes the following research questions.

RQ1. What constitutes a foreign accent?

RQ2. Is there a relationship between foreign accent and L2 phonological awareness?

RQ3. If so, what is the nature of that relationship?

A non-experimental research design was used to carry out the study. The target population is a group of native speakers of Korean learning English as their L2. Because of the growing population of Koreans learning English, there has been an increase in the demand to identify and assess these learners' pronunciation problems and accent. Native English listeners assigned accentedness ratings for Korean-accented speech and identified phonological and acoustic-phonetic features of accents. This analysis is expected to provide an important source of information for Korean learners of English who seek to reduce their foreign accents and improve their comprehensibility and confidence in communicating with native speakers of English. The findings will also help both English language instructors and researchers understand and address L2 learners' pronunciation problems as well as pinpoint the psychological and social issues that these learners must deal with while carrying out communication in L2. Further on, the study explores the

relationship between accent and L2 phonological awareness, thereby contributing to theories on L2 phonological development. In-depth analyses of Korean-accented English and phonological awareness are expected to provide essential information for developing contents and criteria on which future L2 pronunciation curricula and assessments should be based.

1.2. Definitions of Key Terms

1.2.1. Phonology and phonetics

In this dissertation, the term *phonology* will be used more generally, to encompass both phonetics and phonology—a widespread practice in the L2 acquisition literature unless one is referring to particular learning phenomena, such as the acquisition of phonetic contrasts. Phonetics is strictly defined as the system of discrete speech sounds in a language, while phonology is understood more broadly as the rules governing the relationships between sounds (Moyer, 2015).

1.2.2. Accent, foreign accent, and accentedness

In human communication, accent is a medium through which we signal communicative intentions and control the flow of interaction. It serves as the main source of information about the speaker and includes features that constitute the speaker's personal and social identity, such as regional background, ethnic group, social class, and even trustworthiness in an immediate way. *Accent* refers to the phonology and phonetics of a given language variety—a set of dynamic segmental and suprasegmental habits encompassing the sounds, rhythms, and melodies of speech; they convey linguistic

meaning as well as social and situational affiliation (Moyer, 2015). Accents can be observed in both native and nonnative speakers.

On the other hand, the construct of *foreign accent* arises when the distinction between these two speaker groups—native speakers and nonnative speakers—is made. Although wording varies, foreign accent is more or less defined as "the deviations in pronunciation of nonnative speech compared to the norms of native speech" (Gut, 2009, p. 253). The deviation (nonnativelike-ness) is represented by a set of phonological characteristics that marks a speaker as nonnative. A term that is also frequently used in the literature is *accentedness*, which is the perceived degree of accent in a speaker's speech (e.g., Hayes-Harb & Hacking, 2015). The construct has been commonly operationalized in terms of the listeners' perceptions of speech and measured through the listeners' judgments on a scale (see Derwing & Munro, 2009).

1.2.3. Phonological awareness

In language acquisition research, *phonological awareness* typically refers to a speaker's awareness of the phonological system of a language (e.g., Cassady, Smith, & Huber, 2005). For L1 speakers, phonological awareness is generally understood as a conscious construct and explicit knowledge of their L1 phonology, which is strongly related to exposure to written language and literacy. On the other hand, L2 speech learning does not start from a blank-slate state because L2 learners are already equipped with the perceptual and articulatory systems of the L1. They may use the knowledge they possess about the L2 phonological system without providing analytic, explicit verbalization, regardless of whether they acquire the L2 phonological system naturalistically (i.e., without instruction) (R. Ellis, 2009; Mora, Rochdi, & Kivistö-de

Souza, 2014). Taking this point into consideration, this study defines *L2 phonological awareness* as both explicit and implicit knowledge of an L2 phonological system—segmentals and prosody.

1.3. Dissertation Outline

The introductory chapter is followed by Chapter II, which reviews the relevant literature on foreign accent and phonological awareness. Research on each of the constructs is reviewed examining the background, related concepts, and some of the major findings for each one.

Chapter III begins with a discussion of methodological concerns raised in previous foreign accent research and a pilot study of the current study. It then describes the method for the current study, divided into Phases 1 and 2.

Chapter IV presents the results obtained from Phases 1 and 2, and it presents findings from additional analyses that focused on the performance of highly- and little-foreign-accented L2 speaker participants. In-depth analyses of some of the rater and speaker participants will follow.

Chapter V discusses the key findings from Chapter IV, by addressing each of the research questions posed in this study.

Chapter VI summarizes the main findings of the study, then it outlines the limitations of the study and directions for future research.

II – REVIEW OF THE LITERATURE

This chapter provides an overview of the previous research findings with respect to the two main foci of the current study: (a) foreign accent and (b) phonological awareness. The first section of the chapter, which focuses on foreign accent, begins with some theories regarding the acquisition of L2 phonology and foreign accent phenomena. It further presents a review of the research findings on the phonological and phonetic manifestations of foreign accent, followed by a short review on the documented factors related to speakers' phonological ability. The second section of the chapter focuses on the construct of phonological awareness and discusses its measurement and impacts on L1 and L2 acquisition shown in previous research. The chapter concludes with a summary and addresses preliminary research questions for the current study.

2.1. Foreign Accent

2.1.1. Theories of L2 phonological acquisition and foreign accent

SLA researchers widely recognize that the acquisition of L2 phonology is often imperfect and affected by the existing phonological system of the learner's L1. Research on the role of L1 on L2 acquisition was first carried out within the context of the Contrastive Analysis (CA). The CA programme assumes that learners form a set of habits in acquiring an L2. A comparison and contrast of the structural systems of learners' L1 and L2 is crucial in predicting the areas of potential difficulty that L2 learners may face. For example, learners often take over (or *transfer*) their L1 habits while producing L2

utterance. Transfer in the domains of phonology is the process of carrying over certain features from the L1 phonological system to L2, which contribute to the emergence and degree of a foreign accent. The CA Hypothesis (CAH) has made better predictions about sound than about grammar, as shown by Whitman and Jackson (1972), who demonstrated the lack of transfer of such features as word order but definite transfer of phonetic non-equivalents of Japanese-accented English.

Different versions of the CAH have emerged over the past several decades. One of them, a strong or *a priori* version, was proposed by Lado (1957) as follows:

> The student who comes in contact with a foreign language will find some features of it quite easy and others extremely difficult. Those elements that are similar to his native language will be easy for him and those elements that are different will be difficult. (p. 2)

According to this version of the CAH, errors of a language learner can be predicted on the basis of a comparison of thorough descriptions of the L1 and L2, and L2 learning is a result of transferring L1 articulatory habits. In Lado (1957)'s view, the greatest phonological difficulty lies in the assignment of two or more allophones in learner's L1 to different phonemes in the L2. He took the sounds [d] and [ð] as examples: while these sounds are allophones of /d/ in Spanish, they are contrastive (i.e., separate phonemes) in English. Lado claimed that because of such a cross-linguistic difference, Spanish speakers find it difficult to assign the allophones [d] and [ð] to contrastive phonemes in English, so they are likely to make errors in using these sounds, being heard to produce Spanish-accented phones. Stockwell and Bowen (1965) expanded and refined the predictions of the CAH by constructing an eight-level hierarchy of difficulty. This hierarchy attempts to make predictions of difficulty based on whether any given sound is "optional" (a phoneme), "obligatory" (an allophone since it is environmentally

conditioned) or "null" (i.e., non-existent in the L1 nor the L2). According to this hierarchy, maximum phonological difficulty is predicted to arise from a learner acquiring an L2 allophone that is null in the L1.

Although Lado's (1957) and Stockwell and Bowen's (1965) claims were mainly based on anecdotal evidence without systematically gathered data, empirical data that revealed the influence of L1 on L2 phonology were provided by other researchers. For example, Moulton (1962) explained that German and English have distributional differences of phonemes. Whereas English has both voiceless and voiced obstruents in word-final positions, German allows only voiceless obstruents in that position. According to the CAH, this difference will cause German speakers to learn English sound distribution with difficulty. Moulton empirically confirmed that German learners of English had difficulty in pronouncing word-final voiced obstruents, while English speakers learning German had relatively less difficulty.

However, not all errors in L2 speech could be predicted by the strong version of the CAH, which was based on structural description and comparison of the sound systems of L1 and L2. In light of this limitation, Wardhaugh (1970) introduced a weak (*a posteriori*) version. Unlike the strong version which majorly makes predictions of learner errors and identifies causes for the errors, the weak version aims to explain recurring errors and analyze them to discover why the errors occur. This methodology, called error analysis (see Corder, 1971), assumes that errors represent learning difficulties and that the frequency of a particular error is evidence of the difficulty that learners experience in learning the particular form.

The strong and weak versions of the CAH faced problems, however. Research based on the CAH found that predictions of difficulty were not often borne out in actual learner data. Moreover, unpredicted errors were often more numerous than predicted errors, and not all errors could be attributed to transfer. Oller and Ziahosseiny (1970) thus proposed a moderate version of the CAH, which claims that similar structures in the L1 and L2 cause more confusion and difficulty than dissimilar structures: minimal differences are often unnoticed, resulting in non-learning. Their idea is based on the findings of their study where the similarities between the writing systems of the L1 and L2 caused mistakes. But such different versions of the CAH are not spared of the criticism of the claims that all learner errors are due to transfer. That is, the L1 is not the sole (and not necessarily primary) source of errors. In addressing the abovementioned limitations, Major (2008) proposed that the idea of prediction should be defined in the probabilistic sense instead of as an absolute occurrence or nonoccurrence of phenomena, which would rather falsify CA. Today, an interlanguage system is understood as a result of many factors, transfer being one of them (Selinker, 1972; Major, 2008). Non-transfer errors, which result from universals, developmental processes, or even ease of articulation, also exist (e.g., Eckman, 1977).

Meanwhile, the moderate version of CA, which claims that similar phenomena are harder to learn than dissimilar phenomena (Oller & Ziahosseiny, 1970), has inspired some researchers to reckon the impact of language similarity/dissimilarity in L2 phonological acquisition. According to them, similar sounds tend to be more difficult than dissimilar sounds, probably due to the fact that the larger the differences are, the more easily they are likely to be noticed. In contrast, minimal differences are less salient,

resulting in non-learning. Flege (e.g., 1987, 1995) has pioneered the research on similarity and what he called *equivalence classification*, which is fundamental to his Speech Learning Model (SLM). According to him, L2 sounds that are "equivalent" (or similar) to L1 sounds are difficult to acquire because the learner does not perceive or classify the L2 sounds as being different from the L1 sounds. As a result, the learner does not set up a new phonetic category for the L2 sounds but instead, establishes new phonetic categories for sounds that are perceived as different. The "new" (or different) L2 sounds are easier to acquire since there are salient differences.

The SLM was supported by a number of studies conducted by Flege and his colleagues (e.g., Bohn & Flege, 1992; Flege & Hillenbrand, 1984). They argued that adults can produce an L2 sound authentically if it is sufficiently dissimilar from any L1 sound and if the learners have had sufficient L2 experience. For example, Flege (1987) reported that experienced adult learners of French (L1 English) authentically produced French /y/, the "new" sound, whereas they produced French /u/ unauthentically. Flege attributed the difference between the two L2 vowels to their differing categorical status vis-à-vis English vowels: French /y/ has no easily identifiable counterpart in English. On the contrary, since French /u/ has a close counterpart in English, English speakers treated French /u/ as being equivalent to English /u/, and consequently, they produced English-accented /u/. Flege concluded that new vowels (like /y/ for L1 English speakers) evade equivalence classification and that sufficient input can enable adult learners to establish phonetic categories for new vowels.

The SLM focuses on surface phonetic categories, hypothesizing that learners perceptually relate positional allophones in the L2 to the closest positionally defined

allophones (or “sound”) in the L1. The phonetic elements comprising the L1 sound system and those comprising the L2 system exist in a common phonological space, and they mutually influence one another. The L1 and L2 subsystems interact resulting in category assimilation or dissimilation (Flege, 2008). The former occurs hypothetically when category formation is blocked; in this case, the L2 learner may initially produce the L2 speech sound just as if it were the corresponding L1 speech sound, without any modification. But modification is expected if the L2 sound differs audibly from the L1 sound with which it has been equated. Mackay, Flege, Piske and Schirru (2001) observed early and late bilinguals of Italian and English, focusing on the production and perception of English word-initial /b, d, g/, which are typically realized with short-lag Voice Onset Time (VOT) values that are similar to those for Italian /p, t, k/ but different from Italian /b, d, g/ that are realized as pre-voiced. They found that most of the late and early bilinguals pre-voiced English /b/ significantly more often than English monolinguals did. They also pre-voiced /b/ in English less often than is typical for Italian. In addition, they made errors identifying the English /b, d, g/ tokens, hearing them as /p, t, k/. The results indicated that the Italian-English bilinguals had not established separate phonemic categories for English /b, d, g/ but had been using a single phonetic category in which the properties of English and Italian /b, d, g/s had been merged. On the other hand, category dissimilation takes place when the learners establish a new category for a speech sound that is found in the L2 but not in the L1 (Flege, 2008). As the learners add new categories, their combined L1-L2 phonetic space becomes more crowded than that of monolingual speakers of either the L1 or the L2. In the case of bilinguals, they will strive to disperse phonic elements in order to maintain phonetic contrast both within and

between the phonetic categories comprising their L1-L2 phonetic subsystems. A newly established L2 category may even deflect away from the closest L1 category, and vice versa. The L1 category and the new L2 category will then be different from the categories possessed by monolinguals.

In short, according to the SLM, the degree of success with which L2 sounds can be learned is dependent on the perceived phonetic similarity between L1 and L2 sounds. However, despite its attempts to systematically account for all types of segmental deviations in L2 speech production as well as the cause and extent of foreign accents, the SLM is not without limitations because there are some exceptional speakers who perform better with the similar sound than with the new sound (e.g., Bohn & Flege, 1992). It is also problematic that definitions of similar and dissimilar are not always clear-cut. As pointed out by Major (2008), criteria for similar/dissimilar can include acoustic, articulatory, and perceptual factors, intuitions of native speakers and nonnative speakers, and even orthographic evidence.

2.1.2. Phonological parameters in the perception of accent

Theories such as the CAH and SLM have attempted to account for the pronunciation difficulties that L2 speakers may face when they establish an L2 phonological system in addition to the L1 system they already have. The difficulties have been predicted and observed to occur in all domains of phonology: segmentals and prosody (suprasegmentals) (e.g., Anderson-Hsieh et al., 1992). Researchers have examined foreign accents—the phonological/phonetic deviations from what is expected in nativelike speech—in relation to L1-L2 phonological and phonetic similarities and differences in those domains. Below is a review of the previous research findings on how

foreign accents are manifested phonologically and phonetically in each of the above areas.

In the literature of phonology, a segmental refers to a discrete, identifiable unit in the stream of speech, i.e., consonants and vowels in the phoneme inventory of a particular language. Learners using L2 segmentals that are "new" to them are likely to make deviations from nativelike speech, as predicted by the CAH and SLM as well, which result in foreign accents. Previous foreign accent studies on the segmental production of nonnative speakers took one of the two approaches. One was to focus on a few particular individual consonants and vowels in L2 that may attribute to accentedness, and the other was to use composite measures of segmental production, where a number of phonological and phonetic parameters are grouped as a whole. Most of the studies that took the former method conducted a contrastive analysis of the speaker participants' L1 and L2 sound systems and attended to particular segmentals that are predicted to be difficult. In their analyses, the studies also employed a scalar assessment method to examine the influence of the target segmental features on accentedness.

Vowel quality is one feature that has been examined extensively. Studies on nonnative vowel production have investigated the contribution of vowel quality (e.g., duration, formant frequency, or tenseness/laxness) to accentedness ratings, but they have had mixed findings. Magen (1998) assessed the contribution of various phonetic and phonological factors in the perception of Spanish accent in English utterance, one of them being vowel quality: (a) reduction/non-reduction of vowels in English unstressed syllables and (b) the tenseness/laxness of the English lax vowels /ɪ/ and /ʊ/ (these vowels are often pronounced close to the tense vowels, /i/ and /u/). Two native Spanish speakers

read English sentences containing phonological features that are characteristically difficult for Spanish speakers of English, and their productions were acoustically edited to sound more nativelike. For instance, in order to reduce the unstressed vowel as in 'lem[o]ns,' the duration ratio of the stressed to the unstressed syllable was adjusted. The original and edited versions were rated by monolingual speakers of English on a seven-point scale. Statistical analyses did not reveal any strong effects for vowels examined in this study. The listeners showed sensitivity to the vowel quality factor of tense-laxness for only one of the Spanish speakers, but vowel reduction did not affect the ratings.

Munro (1993) investigated the relationship between vowel quality and the perceived degree of foreign accent in her study of 21 Arabic learners of English who read lists of /bVt/ and /bVd/ words in carrier sentences. For each vowel, F1, F2, the range of F1 and F2 (i.e., F2 minus F1), movement in F1 and F2, and vowel duration were measured, and the measurements for Arabic speakers were compared with those of native English speakers. Arabic speakers did not consistently produce any of the examined vowels in a nativelike way. Moreover, the relationship between mean accentedness scores and the various vowel measures differed from vowel to vowel. The above findings indicate that the relationship between accentedness and vowel quality is not straightforward but rather quite complex. The relationship is different depending on the target vowel and even the context where the target vowel appears.

Among the extensively examined consonants are the liquid consonants in English: the lateral /l/ and the approximant /r/.[1] Riney and Flege's (1998) study and other subsequent liquid studies mainly targeted Japanese learners of English, as Japanese

[1] /ɹ/ in the original text. In the current study, the symbol /r/ is used for non-lateral approximant /ɹ/ as in *rose*.

speakers are often observed to substitute the Japanese apico-alveolar flap /ɾ/ for the English liquids /l/ and /r/, which are absent from the Japanese phoneme inventory. In Riney and Flege (1998), accentedness scores of Japanese learners of English, determined by five native English speakers judging the productions from the sentence reading task on a nine-point scale, were partially related to the accuracy of liquid pronunciation. Further analyses of some additional reading and spontaneous speech data carried out by Riney, Takada, and Ota (2000) revealed that substitution of the Japanese flap for the liquids in L2 English speech was positively associated with accentedness, but only in the beginning of the longitudinal study. Riney, Takagi, and Inutsuka (2005) ran a partial replication of Riney and Flege (1998) and also reported some contribution of segmental accuracy to accentedness ratings. The two groups of accentedness raters—L1 English speakers and L1 Japanese speakers (both untrained)—were found to base their respective perceptual accentedness judgments on different (or differently weighted) phonetic parameters: the American raters were more sensitive to the accuracy of the segmentals (especially /r/ and /l/) than were the Japanese listeners. However, these liquid studies used similar research methodologies, and because the relationship was not found in general subjects but in only a few of them, the studies could not firmly establish the relationship between liquid accuracy and foreign accentedness.

Voice Onset Time (VOT) is also known to affect the pronunciation of nonnative speakers whose native language and target language have considerably different VOT values for stop consonants. For example, Magen's (1998) study of Spanish-accented and acoustically-edited (in order to sound more native English) speech investigated the impact of stop voicing, by replacing all cases of the aspiration of voiceless stops

produced by Spanish speakers (VOT ranges 10–22 ms) with the aspiration produced by a native speaker of English (VOT ranges 46–82 ms). The listeners who participated in the rating experiment, however, did not show significant sensitivity to the voicing aspect of foreign accent. Other studies also reported weak or non-significant correlations between VOT and accentedness for nonnative speakers, suggesting that voicing and aspiration of stops, although they contrast much cross-linguistically, contribute in only a limited way to accentedness of nonnative speakers (cf. Riney & Takagi, 1999).

On the other hand, studies that used composite measures of segmental production assessed rates of segmental deviations (or "errors"), instead of attending to particular segmentals predicted to cause pronunciation difficulties. Segmental deviations were significantly related to accentedness in numerous studies. Anderson-Hsieh et al. (1992), for example, examined nonnative speakers' pronunciation errors (i.e., deviations from a native speaker norm) which included phonemic substitutions (e.g., /ʊ/ » /u/ in "put") and allophonic modifications (e.g., [r] » [ɾ] in "unnecessary"). The authors also found that deviation in segmentals had a significant impact on the pronunciation ratings impressionistically made by ESL (English as a second language) instructors.

In Munro and Derwing (2001), the authors transcribed the nonnative speakers' L1 (Mandarin) productions and assessed phonological errors by comparing each production with the production of a native speaker of English. Any occurrence of phonemic substitution (e.g., /i/ instead of /ɪ/), deletion (e.g., missing final consonant), or insertion (e.g., addition of a vowel or consonant) was counted as an error. For the accent ratings, 39% of the variance was explained by phonological errors. The study of Trofimovich and Isaacs (2012) on French learners of English also used holistic segmental errors as one of

their measures of nonnative speech. The segmental errors were significant for accentedness judgments given by novice and experienced raters: 27% of the raters' comments addressed speakers' pronunciation difficulties at the segmental level. Kashiwagi and Snyder (2010) also used composite measures of segmental production, while they counted the vowel and consonant error rates separately. The authors listened to the passage-reading recordings produced by Japanese learners of English and calculated consonant error rates and vowel error rates. Both consonant and vowel errors significantly affected accentedness ratings that were given by American English teachers on a seven-point scale. The other listener group in this study, native Japanese listeners who are teachers of English as a foreign language (EFL), was also sensitive to consonant errors, but not to vowel errors.

The above studies that used composite measures of segmental production have reported a relationship between segmental accuracy and accentedness. However, these composite measures are limited in that they implicitly assign equal weight to each of the segmental deviations. Most of these studies provided few examples of segmental features identified as nonnativelike. Kashiwagi and Snyder (2010), however, reported that L1 Japanese speakers' pronunciation of certain English vowels (r-colored vowels, followed in order by /æ/, /ɑ:/, /ʌ/, /oʊ/, and /ɒ:/) and consonants (/r/ topping the list, followed by /l/, /ð/, /θ/, /f/, and /v/) was responsible for listeners' understanding of the utterance.

Foreign accent is also manifested by nonnativelike use of L2 syllable structure, involving epenthesis (i.e., addition), deletion of a segmental or syllable, or metathesis (i.e., reordering) of segmentals in syllable. Vowel epenthesis and consonant deletion are most commonly reported types of syllable structure errors (e.g., Anderson-Hsieh et al.,

1992). In analyzing the speech produced by Spanish speakers of English, Magen (1998) considered two aspects of syllable structure problems: (a) the presence (or absence) of the initial epenthesis schwa in fricative plus stop clusters in English syllable onsets, and (b) the presence of non-initial epenthetic schwa (-*ed* ending). Magen synthetically repaired the epenthetic schwas by deleting the vocalic portions of the signal of initial/non-initial epenthetic schwas to make the stimuli more like English. The native English listeners who gave the accentedness judgments of the original and edited stimuli were particularly sensitive to these factors related to syllable structure. That is, deletion of epenthetic schwa led to significantly more nativelike ratings. A similar finding was reported in a recent study of French learners of English (Trofimovich & Isaacs, 2012) where the frequency of syllable structure errors (i.e., epenthesis and deletion of vowels/consonants) was significantly correlated with the accentedness ratings.

The other important aspect of speech is prosody (i.e. suprasegmentals), which has grammatical and discourse functions in speech comprehension and production (Trofimovich & Baker, 2006). The documented prosodic parameters that contribute to perceived accentedness include (a) stress, (b) pitch, and (c) speech fluency characteristics. Stress is realized in speech by combining higher pitch with increased vowel duration and intensity (Trofimovich & Baker, 2006). Stress types include the use of lexical (word) stress, prosodic stress, and timing (i.e., "variation in stress and in degree of vowel reduction from syllable to syllable" [Trofimovich & Baker, p. 12]). Based on these characteristics, native speakers can identify their own language and foreign accent (Moyer, 2015). Cross-linguistic differences in rhythm and timing have often been discussed in previous research, especially in the research of English: English is a stress-

timed language, which stretches out the duration of stressed syllables relative to unstressed syllables to preserve equal timing between stress groups. Thus, in English, most vowels in unstressed syllables are reduced to a neutral vowel [ə]. Other languages like Spanish or Korean preserve syllable timing—that is, they maintain relatively equal syllable durations (Jun, 1996; Trofimovich & Baker, 2006). Nonnative speakers of English whose native language is not stress-timed often find the English stress patterns and rhythm challenging (Major, 2001). Communication breakdowns may happen due to a combination of phonological errors related to stress patterns, such as "misplaced tonic (nuclear) stress along with a consonant substitution within the wrongly stressed word" (Jenkins, 2002, p. 89).

Lexical stress plays a central role in determining the identity of words and phrases, and previous studies have observed cases of L1 lexical stress transfer. For example, an English learner of L1 Spanish, French or Italian might say *probLEM*, instead of *PROblem*, as a result of transfer from his or her native language (Magen, 1998; Major, 2001; Missaglia, 1999). They may also show stress deafness (Trofimovich & Isaacs, 2012). Lexical stress errors have been reported to be partially responsible for pronunciation problems and undermine comprehensibility and intelligibility of L2 speech (Field, 2005). Prosodic (or phrasal) stress, on the other hand, refers to stress placed on the major stressed syllable within a major stressed word within a prosodic unit. Prosodic stress signals new and contrastive information in English discourse. In contrast, old or given information is expressed in unstressed elements (Bardovi-Harlig, 1986). Many nonnative speakers of low English proficiency experience difficulty mastering the stress system in English, exhibiting problems such as (a) misplacement of stress (e.g., stressing

given information, instead of new one; Hahn, 2004), and (b) use of stress on every lexical item in an utterance almost equally, "whether it be semantically important or a function word" (Juffs, 1990, p. 107). Hahn (2004) examined native English listeners' reactions to three nonnative speech samples where primary phrasal stress is correctly placed, incorrectly placed, or missing entirely, respectively. The listeners recalled significantly more content and responded more positively when presented with speech that had correctly placed primary phrasal stress.

While a large volume of research has identified nonnative speakers' stress patterns, only a small number of studies have investigated the effect of stress on accentedness. In one of these studies, Magen (1998) acoustically modified nontargetlike productions of lexical and phrasal stress produced by a Spanish speaker of English. Spanish speakers tend to stress the final or the penultimate syllable of English multisyllabic words, and they incorrectly stress function words. Comparison of the native listeners' accentedness judgments made for the original and F0-edited speech samples revealed that the stress factor consistently contributed to listeners' perceptions of accentedness.

Trofimovich and Baker (2006) analyzed the sentences produced by Korean speakers of English (collected via a repetition task) and computed the ratio of the stressed versus unstressed syllable durations. For example, timing used in a Korean-like way (i.e., unstressed and stressed syllables are roughly equal in length) would yield the ratio close to 1. According to the correlational analysis of the suprasegmental production and the accentedness ratings, stress timing was one of the predictors of accentedness ratings, but it was not as significant as fluency-related features such as duration of pauses and speech

rate. On the other hand, studies of Kang (2010) and her colleagues (Kang, Rubin, & Pickering, 2010) focused on the *pace* (number of prominent syllables per run) and *space* (proportion of prominent words) measures in lieu of the appropriacy of stress placement. These studies found that the above stress measures positively and significantly predicted listeners' accent judgments of nonnative speech. Furthermore, Kang (2010) reported that nonnative speakers of low proficiency placed stress on functional words or articles (e.g., *be, the, that,* and *this is*) and failed to use primary stress appropriately throughout the speech. As such, previous studies have confirmed the impact of stress on the perceived degree of foreign accent (e.g., Polyanskaya, Ordin, & Busa, 2017; Trofimovich & Isaacs, 2012), but not all measures of stress were significantly related to accentedness judged by native speakers. But the above studies operationalized and measured stress in different ways, either focusing on stress placement, the number of stressed syllables, or the ratio of stressed/unstressed syllables. Due to such variations, the findings only provide partial explanation of the stress-effect on accentedness.

Pitch measures comprise what is typically defined as the "linguistically meaningful use of pitch movement at the level of the spoken phrase or unit (as opposed to the lexical level)" (Pickering & Baker, 2014, p. 79). Research has shown that L2 speakers often use unexpected choices of tone, such as a replacement of rises with falls. Wennerstrom (1994) reported that Japanese, Chinese, and Thai speakers of English used low, falling tones at boundaries between related propositions, where native speakers of English would anticipate rising or mid-level tones. In his analysis of intonational foreign accent detected from English speakers of German, Jilka (2000) pointed out that some of

the causes of intonational foreign accent involve an incorrect choice or placement of tonal categories, as well as L1 transfer in the phonetic realization of tonal categories.

The phonetic realization of tone sequences is defined along two parameters: (a) scaling (i.e., the fundamental frequency—F0 value) and (b) alignment of the tones (Mennen, 2008). Regarding the former, research has found that pitch height (high, mid, or low; the overall or mean level), movement/contour (rising, falling, or level) and pitch range contribute to accentedness (e.g., Riney et al., 2005: use of falling intonation at the end of clauses). Trofimovich and Isaacs (2012) used the pitch contour as a measure of intonation accuracy, operationalizing it as the number of correct pitch patterns at the end of phrases (i.e., syntactic boundaries) over the total number of instances where pitch patterns are expected (as signaled by pre-boundary lengthening). The pitch contour measure was moderately correlated with the accentedness ratings for the speech of French learners of English.

Pitch range is typically calculated as the difference between the highest and lowest F0 values in a selection of speech. English is known to have a wider pitch range than German or Spanish; L2 English learners whose native language reportedly has a smaller pitch range than English tend to produce a smaller pitch range in their English discourse, and this is considered as a case of transfer (e.g., Mennen, 2008). Studies on pitch range and accentedness had mixed findings. Kang (2010) investigated the relative weights of temporal and prosodic features for listeners' judgments on accentedness of L2 English speech. She found that the nonnative English speech with less pitch variation was evaluated as more accented. Kang also suggested that the limited pitch range, making the speech sound monotonous, may affect listeners' feelings toward nonnative speech. This

was based on some raters' comments that a flat tone of voice made the speech frustrating, boring, and unattractive. The comments also support the generally accepted claim that the wider the pitch range, the more positively (or "lively") the speakers are characterized (Hayes-Harb, 2014). Mennen (2008) shared a similar idea that when English listeners hear German speakers of English with relatively flat and low (German-like) intonation, they may feel the speakers are uncompromising and self-opinionated. On the other hand, Trofimovich and Isaacs (2012) conducted correlational analyses of several speech-related variables (one of them was the pitch range) and accentedness ratings on the speech samples that were produced by French learners of English. In this study, no statistical association was detected between the pitch range and the accentedness ratings.

Peak alignment refers to the timing (or location) of the highest value (i.e., peak) of pitch (or its acoustic correlate, F0) with the syllable in speech. Peak alignment characterizes the rhythm of a prosodic unit of a particular language or dialect. Previous research has revealed some cross-linguistic differences in alignment. For instance, in English or German, pitch peak is used for signaling syllable prominence. It is usually aligned with the onset of the stressed syllable, and the following rapid fall takes place between the stressed and following unstressed syllable. In other languages, like Korean, pitch peak is used for marking a phrase boundary and is usually aligned with the offset of the syllable (Trofimovich & Baker, 2006). An adjustment of peak alignment may lead to improved intelligibility and less foreign-accented speech (Mennen, 2008). However, only a few studies have investigated the relative contribution of alignment patterning on the perception of foreign accent. Trofimovich and Baker (2006) conducted acoustic analyses of the speech produced by Korean speakers of English and computed peak alignment

(also called pitch-accent placement) by measuring the distance between the onset of the vowel in the stressed syllable and the peak in the pitch contour. The L2 speakers, regardless of their LOR, used peak alignment in a significantly different way than did the native English speakers. Nevertheless, peak alignment did not make a statistically significant contribution to the L2 speakers' accentedness.

While the abovementioned features (e.g., stress timing and pitch movement) characterize *speech melody*, there are still other features such as speech rate, pause frequency and duration that characterize *speech fluency* (Trofimovich & Baker, 2006). Fluency here is considered as the component of oral proficiency that is purely a performance phenomenon, defined as the capacity to use language in real time (Lennon, 2000). Rate of speech has been one of the most studied fluency-related suprasegmental features. A common observation is that learners often produce L2 speech at a slower rate than do native speakers, which, according to Munro and Derwing (2001), may have to do with "production problems due to incompletely developed syntactic and morphological knowledge, slower lexical access, and articulatory difficulties that arise in the production of segmentals and prosodic patterns that are less well established than native ones" (p. 453). Speech rate has been measured in several ways—most frequently as the number of "certain units of speech" per second (or minute), or divided by the total duration of speech. Another key temporal measure of fluency is pause. Learners tend to pause more in the L2 than in their L1, and research has also suggested that pauses may reflect processing or memory constraints that take place during the production of L2 speech (Schachter, Christenfeld, Ravina, & Bilous, 1991). The nature of pauses has been

investigated in terms of frequency, length, and proportion to the entire utterance (e.g., Kang et al., 2010).

Munro and Derwing (2001) extensively examined the role of speech rate in judgments of accentedness given to intermediate ESL learners of various L1s. The effect of speaking rate was significant, and the optimal rate for accent was estimated at 4.76 syllables per second. The authors found a curvilinear relationship between accentedness and speaking rate; in other words, speech produced at the optimal rate was rated as less accented than speech at faster or slower rate. In another experiment where L2 speech samples were manipulated (the speech was compressed and expanded to a rate of 10% faster or slower than the original stimuli), speaking rate still made a small yet statistically significant contribution to the ratings, independent of segmental accuracy. The authors suggested the rate effect may be related to error salience and processing difficulty. Fast speech may lead listeners to process the speech quickly, making them become less inclined to notice phonological errors and thus give better ratings. However, too great speech rate may rather place extra demands on the listener and make the speech rated as more accented than slightly accelerated speech. Slow speech, on the other hand, may also be difficult to process because listeners are required to maintain information in short-term memory. Moreover, the listeners may be more inclined to notice phonological errors and give poorer ratings to the speech. Similarly, Riney et al. (2005) commented that fast rate may affect how learners form L2 segmentals and the extent to which their L2 utterances are recognizable and intelligible to other listeners. A number of other studies similarly reported that L2 speakers who spoke slowly and had frequent, long, and inappropriate pauses were generally evaluated as more accented (Kang, 2010; Trofimovich & Baker,

2006; Trofimovich & Isaacs, 2012), while no effects of speech rate on accentedness were reported as well (e.g., Sereno, Lammers, & Jongman, 2016).

Foreign accent researchers have also attempted to determine the area of nonnative pronunciation that carries more weight in evaluations of accentedness than do the other areas of pronunciation. Answers to this empirical question have been searched in various studies that examined the effects of explicit segmental and/or prosodic instruction (e.g., Derwing & Rossiter, 2003; Missaglia, 1999). Furthermore, the technological development of speech analysis and manipulation has enabled researchers to overcome the drawbacks of impressionistic analysis. Separation of prosody from segmentals in speech (and vice versa) makes it possible to create sets of speech stimuli in which target features and their effects can be examined. Low-pass filtering, for instance, is a method of removing all energy components of speech signal above certain frequency (e.g., 450 Hz in Trofimovich & Baker, 2006) in order to preserve suprasegmentals (e.g., the acoustic correlates of stress, rhythm, and intonation) while removing (most of) the segmental content. Hence, low-pass filtering is used to confirm whether the listeners base their accent judgments on the segmental or the suprasegmental content of the speech. In addition to creating low-pass filtered stimuli, Jilka (2000) created and compared two versions of the same stimulus to describe the relative importance of tonal and segmental deviations (i.e., accents): one that is meant to show segmental but not intonational deviations (F0 generated and re-synthesized), and the other that has intonational but not segmental deviations (fully synthesized using diphones). The synthesized stimuli without segmental deviations received significantly better ratings, indicating that the contribution of segmental deviations is more important than intonational deviations.

Despite numerous attempts to determine the relative importance of segmental and prosodic features to accentedness, previous studies have had mixed findings. As generally admitted by researchers, segmental and prosodic aspects of speech are very closely associated. Thus, it is often difficult to draw a clear distinction between the two and avoid having confounding effects (e.g., Piske et al., 2001). Adaptation of newly developed technology has enabled researchers to tease apart and manipulate segmental and prosodic information of native and nonnative speech and has provided new insights, while methodological variations of the above studies have left several limitations which need to be taken into account in carrying out and interpreting foreign accent research.

2.1.3. Factors of foreign accent

The phonological and phonetic parameters of foreign accent described in the previous section are generally attributed to the influence of the existing knowledge of a learner (such as his or her L1). However, other factors of foreign accent (linguistic and non-linguistic) have also been investigated. The age of the learner at the onset of L2 learning or exposure (AOL) and the length of residence (LOR) in the L2 speaking environment have been suggested to have the greatest impacts on foreign accentedness (e.g., Flege & Fletcher, 1992; Piske et al., 2001). Foreign accent research has mostly focused on the conditions of language acquisition, whereas relatively little work has addressed the factors that are inherent to the learner. Some of these factors are associated with an individual's ability to carry out phonology-related tasks.

A core claim in language acquisition research is that there is a specific talent for L2/FL learning which exhibits variation across learners (Dörnyei & Skehan, 2003). Talented learners are seen to be unaffected by the critical period. Research has attempted

to answer questions as to whether such a talent is innate, distinct from more general abilities (e.g., intelligence), or related to variance in language learning success (Schneiderman & Desmarais, 1988).

Carroll devised a test battery called the Modern Language Aptitude Test (MLAT, Carroll & Sapon, 1959) that consists of four components of aptitude related to the ability to learn languages. Among the four components, phonemic coding ability, described as the ability to code unfamiliar sound, is considered to be significantly related to processing of phonological input. Carroll's approach to phonemic coding ability focuses on coding (analysis) of stretches of sound—beyond the level of simple sound discrimination. Through analysis, or often called "phonological processing" (Hummel, 2009), encoding and retrieval of sound information become more likely to occur. Phonemic coding ability is also considered to be a function of the phonological loop component of memory, which refers to the recall and rehearsal of auditory stimuli (e.g., new sound sequences) and is held to be critical in keeping a stimulus item in working memory, and further on, processing and storing it in long term memory (e.g., Baddeley, 1986). Another widely used language learning aptitude test is the Pimsleur Language Aptitude Battery (PLAB, Pimsleur, 1966). This test places greater emphasis on auditory factors (such as sound-symbol discrimination) but less on memory.

Thus far, only a small number of studies have attempted to link language learning aptitude with L2 speaking particularly. For example, Hummel (2009) examined the relationship between aptitude (as measured by the MLAT), phonological memory, and L2 oral proficiency of Francophone children and reported that oral proficiency was correlated with phonological memory and some components of the MLAT measures.

Baker-Smemoe and Haslam (2013) studied the effects of language learning aptitude (high versus low, as measured by the PLAB) and learning context of adult English learners' pronunciation strategies and achievement evaluated in terms of global foreign accent, fluency, comprehensibility, and accuracy. The participants' pronunciation scores in global foreign accent, fluency, and accuracy were positively correlated with auditory aptitude (aptitude in listening or phonological processing) and overall scores of aptitude, supporting the generally accepted claim that accurate perception leads to pronunciation improvement (e.g., Flege, 1995).

In other empirical studies that did not use commercially available language aptitude tests, phonological aptitude has been framed in terms of observable behaviors involving mechanisms like perception, articulatory flexibility, motor skills, or some combination of processing skills. They include (a) the ability to perceive new sound categories as precisely as native speakers do; (b) the ability to read aloud written L2 in a way comparable to native speaker controls, either with or without targeted training beforehand; and (c) the ability to mimic speech sounds. The first ability, perception, involves learners' ability to identify L2 segmental and prosodic components in the input and to discern contrasts between various L2 segmentals (like phonemes), as well as differences between L1 and L2 phonology (Loewen, 2014). The second ability has often been tested to measure holistic proficiency in general. The third one, namely, mimicry, has been considered in some studies as an explanation for the variation in foreign accents between nonnative speakers (e.g., Ioup, Boustagui, El Tigi, & Moselle, 1994; Thompson, 1991). In Purcell and Suter (1980), for example, the subjects' self-reported "innate aptitude for oral mimicry" was identified as the second best predictor of pronunciation

accuracy and accent scores. Flege et al. (1999) measured what they call "sound processing ability," which comprised subjects' self-estimated mimicry ability, musical ability, and ability to remember how English words are pronounced. This variable explained a small amount of the variation (2%) in the foreign accent ratings for Korean immigrants to the United States. In addition, Munro (2008) speculated that successful learners have some sort of "sensory memory" which encourages attention to fine phonetic detail, which learners store and use for perfect accent. However, researchers note that the emergence and nature of the above abilities and memory are still not fully understood (Piske et al., 2001). They also posit that accentedness cannot be determined by investigation of isolated factors, and exclusion of other confounding factors may lead to misinterpretation of results (Gut, 2009; Moyer, 2004).

For all learners, phonological processing is essential in acquiring "new" sounds. The processing skills include (but are not limited to) perceiving speech sounds, noticing the difference between the new L2 sounds and the sounds in learners' existing phoneme inventory, as well as between nativelike and nonnativelike pronunciations, and using phonological knowledge when producing utterance in the L2. However, the constructs related to phonology have been little studied in the field of foreign accent research. Assessment of phonological knowledge has been limited to testing abilities of perception or mimicry, and it has often been unsystematic and simply blended into the speaking and listening components of L2 proficiency tests. To address this, the current study aims to examine the phonological knowledge of individual learners, or phonological awareness, and explore whether this is associated with foreign accent phenomena. The next section of the chapter reviews the research background of phonological awareness.

2.2. Phonological Awareness

The second construct to be examined in the present study is phonological awareness. This section starts with a brief overview of cognitive notions related to phonological awareness, including attention, awareness, and further on, explicit and implicit knowledge in L2 learning. Then it describes how phonological awareness has been measured in previous studies, and reports on the documented impacts of phonological awareness on L1 and L2 acquisition. The section concludes with a short summary.

2.2.1. Attention, awareness, and knowledge

In the field of language acquisition research, language awareness is generally defined as an individual's explicit knowledge about language (according to the Association for Language Awareness) and conscious awareness of language, its learning/teaching, and use (according to the National Congress on Languages in Education of the United Kingdom, reported in Donmall, 1985). These definitions highlight key characteristics and concepts of the construct of awareness, such as consciousness, which is frequently discussed in association with intention, knowledge, and attention (Schmidt, 1990, 2010). Attention is the umbrella term among them. It refers to the process that encodes language input, maintains it active in working and short-term memory, and retrieves it from long-term memory (Robinson, 2005). To date, SLA researchers have proposed approaches to attention and its relationship to awareness and learning. One representative model is of Tomlin and Villa (1994), who proposed that attention incorporates the following three attentional functions: (a) alertness (overall

readiness to receive and deal with incoming stimuli or data); (b) orientation (allocation of attentional resources to a certain type of stimuli); and (c) detection (cognitive registration of the stimuli). Alertness and orientation contribute to the occurrence of detection, but this detection does not necessarily imply awareness. Tomlin and Villa thus argued for dissociation between awareness and learning, and they posit that attention without awareness can lead to learning. Posner and Peterson (1990) similarly argued that attention, in the form of detection, is a precondition for awareness. In their model, none of the components require awareness to operate.

On the other hand, the Noticing Hypothesis (Schmidt, 1990, 1994) acknowledges the role of awareness in language learning. Schmidt argued that a learner must come to *noticing*, which he defined as the "registration [detection] of the occurrence of a stimulus event in conscious awareness and subsequent storage in long term memory" (1994, p. 179). Schmidt (1990) recognized different levels of awareness (which he equated with consciousness) among which noticing is applied to the middle level.[2] The lowest level, awareness as *perception*, initiates creation of internal representations of external stimuli (such as L1-L2 differences of phonological properties of segmentals, rules, and processes) but is not necessarily conscious. This subliminal perception can be compared to Tomlin and Villa's *detection*. Awareness as *noticing* is operationally defined as "availability for verbal report, subject to certain conditions" (Schmidt, 1990, p. 132). Noticing is the necessary and sufficient condition for converting input into intake for learning. Input does not become intake for learning unless it is noticed, or consciously

[2] Truscott (1998) provides detailed criticism of this distinction and argues that "the hypothesis is not based on any rational theory of language" (p. 104). He posits that the noticing hypothesis should be limited to describing metalinguistic knowledge and not overall language competence.

registered through awareness. The highest level, awareness as *understanding*, involves further mental activities such as (a) analyzing and comparing what is noticed and has been noticed before, (b) problem-solving, and (c) making generalizations across instances, all within consciousness. Empirical studies have indicated that learners with higher levels of awareness of the target features perform better than those with lower levels of awareness (e.g., Leow, 1997, stem-change of Spanish verbs). However, understanding is not required but only facilitative for language learning (Schmidt, 2010). Schmidt further argued that noticing can be encouraged through instruction and that explicit instruction may also increase awareness at the level of understanding.

Robinson (1995) reconciled the models of Schmidt (1990, 1994) and Tomlin and Villa (1994) and proposed that the notion of noticing is "detection plus rehearsal in short-term memory, prior to encoding in long-term memory" (Robinson, 1995, p. 296). Rehearsal in Robinson's model is a result of resource allocation, and it can involve instance-based, data-driven processing (maintenance of stimuli (input) assembled in memory) or schema-based, conceptually-driven processing (integration of encoded stimuli within the context, activating previously attended information encoded into long-term memory). These rehearsal processes give rise to awareness, which is critical to noticing and distinguishes noticing from "simple" detection (i.e., detection without awareness).

Schmidt (2010) notes that although awareness and attention are closely linked ("what we are aware of is what we attend to," p. 31), awareness of certain abstract rules cannot be a prerequisite for learning. For example, native speakers have some intuitive understanding of subtle rules of a language that they cannot verbalize. Similarly, some

advanced naturalistic L2 learners may also have near-nativelike intuitive knowledge rather than instructed rules. Such cases are described as implicit knowledge, which is distinguished from explicit knowledge in SLA research. According to Rod Ellis (2009), explicit knowledge is conceptualized as "involving primarily 'analyzed knowledge' (i.e., structured knowledge of which learners are consciously aware)" and as "metalanguage." Explicit knowledge involves declarative, factual knowledge of abstract rules, fragments, and exemplars. On the other hand, implicit knowledge is characterized as "subsymbolic, procedural and unconscious" (p. 38) and therefore, highly automated. Implicit knowledge is largely systematic and structured. It is employed with some degree of certainty as to its correctness and is accessed by means of automatic processing during fluent performance, whereas explicit knowledge is often inconsistent and imprecise, involving controlled processing. The generative accounts of language acquisition regard linguistic competence as comprising implicit knowledge that enables learners cognize what is grammatical and ungrammatical (e.g., Chomsky, 1976). The cognitive accounts also consider implicit knowledge to be primary in language acquisition, but they have divided views of its relationship with explicit knowledge in L2 acquisition.

The non-interface position holds that implicit knowledge and explicit knowledge involve independent acquisitional mechanisms. It is represented by Krashen's (1982) distinction between subconsciously acquired knowledge (implicit knowledge) and consciously learned knowledge (explicit knowledge). The non-interface position views that explicit knowledge, obtained through learning, cannot transform into implicit knowledge, and it rejects the possibility of implicit knowledge transforming into explicit knowledge. Explicit knowledge is only available for use as a "monitor," which enables

the learners to monitor what they produce through the acquired system. In optimal conditions, explicit knowledge contributes to performance.

In contrast, the strong-interface position argues that transformation from implicit to explicit knowledge and the vice versa (i.e., explicit knowledge being converted into implicit knowledge) are possible through practice (DeKeyser, 1998). Through observation (and instruction), learners first obtain declarative knowledge (knowledge about facts and things), and begin to use that knowledge (then, it is knowledge about rules—how to perform various cognitive activities; i.e., procedural). Practice enables automatization of procedural knowledge: learners' performance at this stage becomes fluent, effortless, and automatic without deliberation, unlike during the initial stages of acquisition where they were consciously aware of rule applications. This automatized, or proceduralized, knowledge may appear in output as implicit knowledge that does not involve conscious attention (DeKeyser, 2003).

The weak-interface position regards that explicit knowledge (although its nature is still on a debate; R. Ellis, 2005, 2009) can be converted into implicit knowledge but limitedly. Rod Ellis (1993, 1994), for instance, posits that practice plays a role in transforming explicit knowledge into implicit one only when the learner is developmentally ready to acquire the target linguistic form. Nick Ellis (1994) argues that explicit knowledge and implicit knowledge are in a dissociable but cooperative relationship in which explicit knowledge indirectly contributes to the acquisition of implicit knowledge by, for example, enabling learners to notice relevant target features in the input (and to notice the gap between the input and their existing linguistic competence). The two weak-interface positions have a common understanding that

explicit knowledge may facilitate learning where implicit knowledge fails. In a recent paper of Han and Finneran (2014), the abovementioned theoretical debates were revisited along with their analyses of L2 learners' written constructions of different types of fossilization (i.e., under-passivization, over-passivization, and underused and overuse of articles and plurals), each suggesting a differing relationship between implicit and explicit knowledge. The authors claimed that while there is likely both implicit and explicit knowledge in the interlanguage of L2 learners, the three different types of relationships (i.e., interfaces) can co-exist within and across interlanguage subsystems, and within and across learners.

The notion of phonological awareness is closely associated with "knowledge." Phonological awareness is a conscious construct for L1 speakers since its measures are obtained through tasks that make the subjects consciously attend to, detect, and manipulate the phonological structure of spoken words (Snow et al., 1998). The L1 speakers' explicit knowledge of their L1, although the L1 itself is acquired implicitly, will be essential when one needs to analyze and manipulate L1 sounds. However, some recent studies include implicit knowledge to the construct of phonological awareness (e.g., Pullen & Justice, 2003). They argue that phonological awareness should be considered at more than one level: an implicit level where speakers rely on "epilinguistic" (spontaneous, intuitive, and automatic) ability to discriminate and manipulate phonological units, and a more explicit level where speakers have "metalinguistic" (intentional, reflective, and conscious) ability, on the basis of epilinguistic ability (e.g., Piasecka, 2011). In the case of L2 phonological acquisition, in addition to explicit metalinguistic knowledge of L2 phonology that they have acquired

(most likely through pronunciation-focused instruction), L2 learners may also have some implicit knowledge of the phonological system of the L2, such as the knowledge of the sound differences between the sound structure and properties of the L1 and those of the L2. Mora and others (Mora et al., 2014) describe the development of L2 phonological awareness as follows:

> [L2 phonological awareness] develops mostly implicitly, first through learners' ability to discern the differences between L2 and L1 sounds and their skill in extracting L2-specific phonetic and phonological regularities from L2 exposure and use; and then by using this phonological knowledge effectively to modify the nature of their L2 phonological representations. The L2 implicit phonological knowledge gained this way does not easily lend itself to explicit verbalization. (p. 58)

In this view, all learners, especially naturalistic learners who acquired L2 phonology implicitly without explicit instruction, can use what they are aware of the L2 phonological system without providing analytic, explicit verbalization (Ellis, 2009; Kivistö-de Souza, 2015; Venkatagiri & Levis, 2007). Phonological awareness is thus seen to involve both implicit and explicit knowledge of the phonological system of a language.

2.2.2. Measuring phonological awareness

Research has suggested several ways of measuring phonological awareness. In a chapter on testing of implicit and explicit knowledge in L2, Ellis (2009) summarized that a variety of tasks have been used in L2 phonology research in order to sample a range of speech styles: (a) casual speech; (b) careful speech (i.e., the speech found in interviews); (c) reading; (d) word lists; and (e) minimal pairs. These task styles are spread along a continuum according to the amount of attention the speakers pay to their own speech, the least attention being paid in casual speech and the most in minimal pairs. In other words,

a task designed to elicit spontaneous speech would require "low awareness, limited response time, a focus on meaning, and little opportunity to use metalanguage" (R. Ellis, 2009, p. 347), while a minimal pair task would require higher awareness and serve as a test of explicit knowledge of L2 phonology. However, a speaker may not certainly pay equal amount of attention while performing tasks of different speech styles. Ellis therefore suggested obtaining measures of the extent to which speakers monitor their pronunciation while performing tasks (e.g., speakers' self-reports of pronunciation monitoring).

Conventional measures of phonological awareness use tasks that require conscious processing of speech—attending to, detecting, and manipulating the phonological structure (e.g., Snow et al., 1998). A number of tests of phonological awareness have been created and used by language researchers and practitioners: e.g., Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, Rashotte, & Pearson, 2013) and Phonological Awareness Test (PAT; Robertson & Salter, 2007). Although differences exist among the tests, most of the tests consist of measures of phonological awareness that have been demonstrated to be valid and reliable. The commonly included measures are as described below.

Detection, recognition, discrimination, and matching assess whether a speaker participant has "an ear for the sounds" (Adams, 1990, p. 80) and whether he or she is able to detect, recognize, discriminate, and match certain phonological units (sounds, rhymes, or syllables) or patterns from speech stimuli. Tasks may involve judgment: e.g., Listen to the two words—minimal pairs—and answer if their beginning sounds are the same (S) or different (D). Oddity (or "odd-one-out") tasks assess whether participants can compare

sound units in the stimuli and identify "odd" ones: e.g., Listen to the words and choose the one that has an odd sound/rhyme/syllable (Cassady et al., 2005). Recognition and matching can be tested at a prosodic level. The DEEdee task is an example of the measure of prosodic sensitivity in recent studies (e.g., Goswami, Gerson, & Astruc, 2009; Whalley & Hansen, 2006). In this task, stimuli are created by replacing each syllable of an utterance with the nonsense syllable *dee*, while retaining the prosodic pattern of the utterance (e.g., *Cinderella = DEEdeeDEEdee*). On each trial, participants are asked to match a DEEdee stimulus to one of the response options.

Phonological manipulation is the most widely observed skill in phonological awareness tasks. It encompasses segmentation, blending, addition, deletion, substitution, and rearrangement of specified phonological units or patterns from the stimuli. Segmentation requires decomposing of a word into component phonemes and syllables, while blending requires combining phonemes into syllables and syllables into words. Deletion of a phoneme or syllable is one of the most frequently used phonological manipulation tasks (e.g., Chien, Kao, & Wei, 2008; Chung, McBride-Chang, Cheung, & Wong, 2013; Venkatagiri & Levis, 2007). The Spoonerism task in Venkatagiri and Levis (2007) is an example of the substitution and rearrangement task in which participants are asked to exchange the initial sounds of words in two-word phrases and create phrases with different words (e.g., *white fish* transformed into *fight wish*).

Repetition, rapid naming, and decoding are also used as measures of phonological awareness. Repetition often employs pseudowords and requires participants to code phonological information of the pseudoword for temporary storage in working or short-term memory. Rapid naming measures the ability to efficiently retrieve phonological

information from long-term memory and execute a sequence of operations rapidly. Rapid naming tasks may involve letters, digits, symbols, or non-symbolic objects. Decoding assesses general knowledge of letter-sound correspondence to blend sounds into words; however, it has been frequently used as a measure of oral word (or pseudoword) reading or word recognition skills rather than as a measure of phonological awareness (e.g., Schiff & Calif, 2007). The abilities to handle phonological-orthographic processing operations (e.g., naming, decoding) are similar to the skills implicated in the phonetic coding component of L2 aptitude (Dörnyei & Skehan, 2003).

The above measures, however, may not be adequate or sufficient to reflect upon the awareness of late L2 learners whose L1 phonological system is already in place and thus the L2 phonological system has been established in a more complicated manner—naturalistically or through instruction, or both. Moreover, many of the published phonological awareness test batteries target young children in early stages of L1 development. The CTOPP (Wagner et al., 2013) is one of the tests that are designed for older testees (ages 7-24); its "Alternate Phonological Awareness Composite Score" measures phonological awareness exclusively with pseudowords, memory, and the ability to execute a sequence of operations at high speed repeatedly. However, studies of adult or L2 speakers have mostly used the abovementioned tasks that were originally designed for L1-speaking children, or they have attempted to develop task materials by incorporating varied types of tasks to meet their own research purposes.

Although the number is not large, some recent L2 pronunciation studies have used qualitative reports or surveys as a means of assessing learners' awareness of phonology. Kennedy and Trofimovich (2010) and Kennedy, Blanchet, and Trofimovich (2014)

considered two types of awareness in their studies: qualitative awareness and quantitative awareness. The former is associated with awareness of meaning-based and context-based language usage, while the latter considers language as a set of items (e.g., sound system or word pattern) to be memorized. Both awareness types were observed from the L2 learner participants' journal entries, by applying Benson and Lor's (1999) analytical framework on dual conceptions of learning. In a similar vein, Ramírez Verdugo (2006) used questionnaires in order to gather information about L2 learners' degree of awareness, difficulties, sensitization, and feelings about L2 prosody and speech.

Stimulated recall protocols are emerging as a tool to obtain reliable insights into the cognitive processing of language learners in performing oral tasks, while partially eliminating memory effects. Wrembel (2011, 2015) investigated learners' self-perception of metaphonological awareness in spontaneous speech, by means of verbal protocols through which the learners consciously monitored their own pronunciation and identified strategies for improvement in pronunciation. More recently, Kopečková (2018) collected immediate and delayed recall protocols in which L1 German students verbalized any thoughts they had on their L3 Spanish production performance.

2.2.3. Research findings on L1 phonological awareness

Phonological awareness in childhood develops in a hierarchical manner, proceeding from awareness at the word level, to the syllabic level, to the onset-rime level, and finally to the phoneme level (Goswami & Bryant, 1990). As children acquire their L1, they build skills to detect (or isolate, recognize) or count the sound units within a word. As these isolation and detection skills are automatized, they progress to the level of manipulating the phonological units.

The effect of phonological awareness on children's L1 development has been attested in a number of studies. Children with high phonological awareness were superior in learning to read and write in their L1 orthographic system, whether the system was alphabetic or non-alphabetic (McBride-Chang, Bialystok, Chong, & Yanping, 2004; Piske, 2008) or whether it was opaque or transparent. Phonological awareness has also been found to play a significant role in the development of reading and vocabulary (Metsala, 1999). Studies have indicated that older monolingual speakers perform similarly. Williams and Wood (2012) investigated skilled readers' sensitivity to the lexical tone and amplitude patterns, and they reported a positive relationship between the sensitivity and standardized skills in reading.

2.2.4. Research findings on L2 phonological awareness

Previous studies on phonological awareness in L2 acquisition have varied in research focus, design, and measures. They thus have had mixed findings that do not provide straightforward explanations about the relationship between L2 phonological awareness and L2 development. A large number of the studies were non-experimental studies that ran correlational and/or regression analyses. Experimental (and quasi-experimental) studies have been conducted as well in order to examine the development of phonological awareness as a result of instruction. Studies have also observed and discussed transfer of L1 phonological awareness. The remainder of this section reviews the findings of the studies that have focused on the impacts of L2 phonological awareness on the development of L2—specifically, general proficiency, (oral) reading, and pronunciation.

Previous research has shown positive correlations between phonological awareness and general proficiency in L1 and L2. In both Verhoeven (2007) and Chien et al. (2008), young learners' L1 and L2 developmental patterns were related to the development of their phonological (phonotactic and phonemic, to be more specific) awareness in both L1 and L2. Similarly, Yeung, Siegel, and Chan (2013) found that explicit instruction on phonological awareness skills (such as syllable detection) had facilitated Chinese ESL kindergarteners' phonological awareness development and positively influenced their English word reading, spelling, and vocabulary acquisition.

Authors of studies on L2 reading commonly use the term *reading* to refer to either *reading comprehension*, which is primarily meaning-based cognitive work, or *oral reading*, which is often referred to as *decoding*. Koda (1998) made the distinction between the two in a study that investigated relationships between L2 phonological awareness (which was strictly phonemic in this study), L2 decoding and reading comprehension, and L2 speakers' alphabetic experience in their L1. The study involved two groups of adult ESL readers whose L1s were Chinese and Korean. The participants were tested on their phonemic awareness in English (via auditory discrimination, phonemic substitution, phonemic deletion, and phonemic insertion), L2 decoding (via pseudoword reading), and L2 reading (via the standard Test of English as a Foreign Language (TOEFL) and some cloze tests). Koda found that the two groups did not differ either in their phonemic awareness or in their decoding ability. However, strong interconnections were found between phonemic awareness, decoding, and reading comprehension among readers of Korean (a non-Roman alphabetic language) but not among readers of Chinese (a logographic language). Koda interpreted this as an

indication that while differential L1 orthographic experience is not directly associated with L2 phonemic awareness, variations in prior processing experience may engender the use of diverse phonological processing procedures, which may account for qualitative differences in L2 processing behaviors.

Oral reading, or decoding, is seen as an indicator of L2 literacy development and orthographic-phonological knowledge of learners, and it is thus relevant to L2 pronunciation (Schiff & Calif, 2007). Kang (2009) investigated the role of L1 and L2 phonological awareness in Korean elementary EFL learners by examining their L2 vocabulary, oral language comprehension (i.e., listening comprehension), and oral reading of L2 words and pseudowords. Overall, the learners' L1 phonological awareness and L2 phonological awareness were both found to be strong predictors of L2 reading. Another noteworthy finding from L2 decoding research is that phonological awareness, whether in L1s or L2s, may not strongly predict L2 decoding skills of children beyond a certain age (Swanson, Rosston, Gerber, & Solari, 2008).

Recent phonological awareness studies have attempted to include measures of prosodic awareness in their analyses, unlike earlier studies that mainly focused on phonological awareness at the segmental level. Chung et al. (2013) investigated the relationships between general auditory processing, speech perception, phonological awareness, and word reading in L1s and L2s in a study of L1 Cantonese-L2 English children. The auditory processing was measured through a task that required the subjects to listen and judge sequences of tone pairs: High-Low (HL), HH, LL, and LH. While this measure can be considered a measure of tone awareness (e.g., Yeung & Chan, 2012), Chung et al. treated it as separate from phonological awareness (which was measured in

their study through syllable deletion, onset deletion, and rhyme production in the L1 and through syllable deletion and initial/final phoneme deletion in English, the L2). The speech perception measure in their study used syllable discrimination tasks in Cantonese and English, which can be interpreted as measures of syllable awareness in many other phonological awareness studies. The authors reported that general audio processing, L1 and L2 phonological awareness, and L1 and L2 word reading were correlated at a significant level, although syllable discrimination remained independent from both L1 and L2 word reading.

While phonological awareness has been continuously researched in relation to L2 decoding skills, very little attention has been given to how L2 speakers actually use L2 sounds that they are aware of or how such performance is received by interlocutors. Venkatagiri and Levis (2007) conducted a study focusing on EFL adult learners' L2 speech comprehensibility—the perceived ease of understanding speech—and claimed that comprehensibility can be facilitated by phonological awareness which they measured using a set of sound detection and manipulation tasks. They interpreted their finding to lend some support for Schmidt's noticing hypothesis (1990, 1994): L2 speakers with superior phonological awareness were able to notice the gap between their own production of the target form (sound) and what the form actually requires, thus yielding more comprehensive pronunciation. However, the data were collected using the tasks that mainly required the learners' L2 metalinguistic knowledge, even though the authors acknowledged the implicit nature of L2 phonological awareness in the learners.

Some of the studies that examined learners' awareness and L2 speaking performance did not use conventional phonological measures such as detection and

manipulation, but rather used learners' verbal reflections. For example, Ramírez Verdugo (2006) argued that EFL learners' prosody greatly improved when they were provided explicit information about L2 prosodic form and meaning, whose effects were shown in learners' comments and reactions recorded during the prosodic training. For example, pitch graphs were presented to learners who were taught how to recognize similarities and differences between the prosody of native speakers and learners. Kennedy and Trofimovich (2010) determined the awareness of L2 learners from coded journal entries that learners made while taking an intensive pronunciation course that focused on the suprasegmental aspects of English. In this study, qualitative awareness was defined as an awareness of "how particular characteristics of English speech carry particular meaning" (p. 183). Learners with more qualitative awareness showed greater improvement in their L2 pronunciation, suggesting that learners' heightened awareness could be translated into their pronunciation. Kennedy et al. (2014) used journals of L2 French learners to elicit measures of pronunciation awareness, which they defined as "learners' conceptions of how pronunciation is acquired and how pronunciation patterns help speakers convey intended meanings in the L2" (p. 80). They reported that pronunciation awareness measures were associated with significant improvements in learners' segmental and intonational production as well as fluency. The studies discussed above did not strictly focus on the kind of phonological awareness commonly examined in most other studies and are therefore bound to have validity and reliability issues. For instance, Kennedy and Trofimovich admit that coding and analysis procedures for the qualitative reflection data (e.g., from journals) were never simple. Nevertheless, their findings provide insightful information about learners' awareness in L2 pronunciation development.

On the other hand, Wrembel's (2011) investigation of L2 learners' self-perception of metaphonological awareness found that her participants had a high degree of analyzed knowledge, as reflected in instances of self-repair and self-reflection on L2 pronunciation strategies; however, they showed a low level of control when monitoring their own L2 phonetic forms during their L2 speech performance. In her later study on the role of attention and noticing in third language (L3) input processing, Wrembel (2015) examined whether speakers of higher proficiency would demonstrate more enhanced metalinguistic awareness (e.g., more instances of self-corrections) by comparing speaker groups at two different proficiency levels. However, the performance of the two groups did not differ significantly. Kopečková (2018) adapted the stimulated recall protocols of Wrembel (2015) and revealed that even students as young as age 13 were able to notice their (L3 Spanish) pronunciation problems and point at aspects of both fluency and accuracy in their speech that they had performed three years prior to the study. These learners showed sensitivity mainly to the segmental features of their speech but minimally to intonation and speech rate. About 83% of their reflective comments involved noticing an auditory aspect of speech (but no further explication), while in the remaining comments, they were able to analyze and explain how they articulated target features. Because the focus of the above study was rather to examine the learners' phonological awareness, how these learners' phonological awareness was related to their pronunciation development remains unexplored.

2.2.5. Summary

The construct of phonological awareness has been frequently examined in SLA research along with related topics, such as attention and implicit and explicit knowledge.

Researchers in recent years have acknowledged that phonological awareness may not only reflect explicit understanding of L2 phonology but may also reflect implicit knowledge. It has also been pointed out that conventional phonological awareness tests may not be adequate in measuring L2 phonological awareness, which is much more complex than L1 phonological awareness. Previous research has also sought to find a relationship between L2 phonological awareness and diverse L2 skills such as general proficiency and oral reading (decoding). However, much less research has investigated whether (and how) phonological awareness is associated with L2 speech performance, although both rely on purely phonological and phonetic aspects—L2 sounds and sound systems. Moreover, previous studies were not without limitations. The studies had various foci, and their phonological awareness measures were not uniform, making it difficult to generalize their findings. As for the measures, most of the studies investigated explicit knowledge of L2 phonology only, although a few acknowledged the implicit nature of phonological awareness in L2s (e.g., Venkatagiri & Levis, 2007). It is also noteworthy that prior L2 phonological awareness studies have mostly focused on children (kindergarten age, at the youngest: e.g., Verhoeven, 2007), whereas adults or late L2 learners have been examined in only few studies (e.g., Koda, 1998). Therefore, much remains to be discovered regarding how L2 phonological awareness is developed in late learners and how it is related to L2 phonology and speech.

2.3. Summary of the Review

This chapter has reviewed key findings of the research of the two constructs, foreign accent and phonological awareness. The SLA literature is replete with theories

about the foreign accent phenomena and about how learners perceive and produce L2 sounds that are new to (and different from) pre-existing L1 sounds. A large number of phonological and phonetic parameters of foreign accents have been observed from speakers of a variety of L1s and L2s. However, many of the previous studies targeted only a limited number of L2 segmental and/or suprasegmental features that are predicted to be difficult to learners of particular L1s and L2s. Researchers have also argued that pronunciation difficulties, which likely yield foreign accents, cannot always be traced back to L1 systems. Therefore, more extensive and comprehensive investigation is necessary to identify what phonological and phonetic characteristics constitute the foreign accent and affect the degree of foreign accent perceived by listeners.

Previous research has noted that a variety of factors contribute to L2 speakers' accents, including variables pertaining to individual differences. The current study aims to concentrate on one of these learner-intrinsic variables, phonological awareness, defined as individual learners' implicit and explicit knowledge of their L2 phonological system. Phonological awareness is reflected in receptive and productive phonological abilities such as pattern identification/recognition, manipulation, coding/decoding, mimicry of phonological input, and verbalization of phonological rules. All of the above abilities are considered essential in speech processing, yet little is known about the relationship between phonological awareness and the foreign accent phenomena. A question for investigation is whether the variance in phonological awareness accounts for the variance in accentedness (i.e., nonnativelike-ness). Few studies have investigated L2 speakers' accentedness in relation to their abilities to use and explain the components of the L2 phonological system. While it has been suggested that higher phonological

awareness allows better pronunciation, it is still necessary to explore whether foreign accent and phonological awareness are indeed related and, further on, provide a detailed description of the relationship if it is found.

III – METHOD

This chapter aims to describe the methodology of the current proposed study. The chapter first reviews methodological concerns raised in the previous foreign accent research and then describes the related issues and findings observed in a pilot study. Then the chapter presents the research questions of the study and an overview of the study design. The rest of the chapter presents the method of the study, per phase, describing the participants, instruments, and procedures involved in data collection, coding, and analyses.

3.1. Methodological Issues in Foreign Accent Studies

Foreign accentedness is operationalized in terms of the listeners' perception of nonnative speech. Previous studies have thus addressed methodological issues experienced while collecting nonnative speech samples and assessing accentedness. This section discusses four principal points related to these issues: (a) speaker subject populations, (b) speech tokens, (c) accentedness ratings, and (d) listener/raters.

Speakers examined in previous foreign accent studies have varied in many important ways, such as number of participants (as many as 240 in Flege et al., 1999) and individual characteristics including participants' amount of L2 experience and degree of motivation to speak L2. Most of these studies have examined L2 pronunciation in learners of English, while a smaller number have examined learners of other languages. The subjects' native languages in these studies were diverse.

Some foreign accent studies have examined L2 pronunciation in groups of speakers of various L1s (e.g., Munro & Derwing, 2001; Yuan, Jiang, & Song, 2010), rather than homogeneous L1 speakers. Only a small number of these studies have investigated the effect of L1 background on L2 accentedness. For example, in Purcell and Suter (1980), native speakers of Arabic and Persian were found to have a better pronunciation of English than native speakers of Japanese and Thai. However, the relative effect of L1 background could not be identified because the subjects of this study were not matched for other variables such as onset age of learning L2. Anderson-Hsieh et al. (1992) conducted a sub-analysis of their subjects: the East Asian group consisting of native speakers of Korean and Chinese, and the Indian Subcontinent group consisting of native speakers of Assamese, Hindi, Tamil, and others. The authors reported that for each group, the strength of the correlations between pronunciation ratings and scores on segmentals and prosody varied.

Furthermore, not all studies have included a control group of native speakers in addition to nonnative subjects. While some have stressed the necessity of establishing native-speaker norms, others have presented only nonnative speech samples to raters (e.g., Kashiwagi & Snyder, 2010). Piske et al. (2001) noted that the decision not to include a control group did not impact the way that raters recruited to evaluate speech samples were able to distinguish native from nonnative speech. Trofimovich and Isaacs (2012) also wrote that explicit comparisons with a NS norm is not necessary because listeners in general have an intuition about foreign-accented speech. Moreover, the relative proportion of native-speaker stimuli as compared to nonnative-speaker stimuli may impact the ratings given to nonnative speakers. Flege and Fletcher's (1992) study

revealed that the scores accorded to nonnative speakers may be depressed if a greater number of native speakers are included. On the other hand, if fewer or no native speakers are included among the stimuli, nonnative speakers may be judged to be more nativelike than might otherwise be the case. This phenomenon is known as “range effect” (Hayes-Harb, 2014; Jesney, 2004).

Another methodological issue relates to the speech tokens by which accentedness judgments are made. First, the length of the speech tokens presented to raters have varied ranging from single segmentals excised from word productions (e.g., Snow & Hoefnagel-Höhle, 1977) to passages up to several minutes long (e.g., Anderson-Hsieh et al., 1992). Among them, sentence-length stimuli have been most prevalent (e.g., Bongaerts, van Summeren, Planken, & Schils, 1997; Flege & Fletcher, 1992; Munro & Derwing, 1999, 2001; Piske et al., 2001; Southwood & Flege, 1999). While it has been shown that foreign accentedness judgments can be made based on stimuli shorter than a sentence with good levels of inter-rater reliability (Jesney, 2004), Major (2001) argued that isolated words are not as unnatural as they have often been claimed, and that longer excerpts of speech are more informative for raters.

There has also been variation in the elicitation method used to collect the speech tokens presented to the raters. Reading and free (extemporaneous) speech have been the most common elicitation techniques. Speech elicited via reading tasks allows greater control over the stimuli, particularly in cases where certain specific features (like VOT) are of interest. Extemporaneous speech is considered to be more representative of communicative production and more appropriate for investigation of acoustic-phonetic features such as speech rate or overall prosody. Other elicitation techniques employed in

previous research include direct repetition based on native speaker norms (Snow & Hoefnagel-Höhle, 1977) and delayed repetition. The latter is often considered to be more efficient than the former, in that it prevents direct imitation from sensory memory (Flege, Birdsong, Bialystok, Mack, Sung, & Tsukada, 2006; Piske et al., 2001).

Some studies have used both reading and extemporaneous speech samples in order to determine the elicitation-type effects on L2 accentedness. Read speech was judged to have stronger foreign accents than extemporaneous speech samples in Thompson (1991), whereas Munro and Derwing (1994) reported no significant difference between the accentedness ratings made from the reading data and those from the spontaneous speech. Though findings vary, the advantages and disadvantages of each of the methods have been identified. Extemporaneous speech is efficient in preventing hyper-articulation or spelling pronunciation (Altenberg & Vago, 1983); however, it may contain morphosyntactic and lexical errors that can sometimes influence accentedness judgments (Calla McDermott, 1986). Read speech makes it possible to control for grammar and vocabulary, but it is not naturalistic and often lacks ecological validity (Levis & Moyer, 2014). Researchers have also noted that extemporaneous speech can allow subjects to use general communication strategies such as avoidance of words or expressions that contain challenging sounds. In addition, spontaneous speaking tasks are generally considered to be more cognitively demanding and complex than simple reading tasks (Robinson, 2001).

One trend in recent studies on L2 speech has been the use of technology—specifically, acoustically manipulating naturally-produced speech. Magen (1998), for example, compared accentedness ratings of computer-edited and unedited nonnative

speech samples to examine the contribution of the edited (i.e., corrected) features such as vowel quality to accentedness. Munro and Derwing (2001) used acoustically-edited speech data where L2 speech samples were mechanically manipulated (compressed or stretched) for speech rate. Similarly, the L2 prosody studies like Jilka (2000), Munro (1995), and Trofimovich and Baker (2006) used low-pass filtered speech in order to attend to all segmental cues, and other studies cross-transplanted prosodic features to original speech (e.g., Sereno et al., 2016; Winters & O'Brien, 2013). However, researchers acknowledge that naturalistic and synthesized data should not be treated in the same manner, as the latter type is not ecologically valid and thus may not be generalizable to other speech settings.

Methodological variation can also be determined using accentedness ratings. In some foreign accent studies, the raters were asked to make categorical judgments, classifying subjects in terms of native and nonnative categories (e.g., Alba-Salas, 2004). Most foreign accent studies have assessed accentedness using Likert-type scales, where raters select a point between two endpoints. These intermediary point scores are intended to represent the degree to which the raters agree with the labels at the extremities of the scale. The endpoints of the rating scale are reserved for extreme categories: one is usually marked as "no foreign accent" or "nativelike pronunciation," and the other, "heavy/strong foreign accent." The number of gradients on such interval scales has differed widely in resolution. In earlier studies, a five-point scale was commonly used (e.g., Oyama, 1976; Snow & Hoefnagel-Höhle, 1977; Thompson, 1991). A more common approach in recent studies has been the adoption of a seven- or nine-point scale (e.g., Munro & Derwing, 2001; Riney & Flege, 1998; Trofimovich & Baker, 2006). Other types of scales used in

previous studies include a continuous (or, sliding) scale where raters adjust a level or cursor along a continuum upon which only the endpoints were marked. Such a scale allows much finer distinctions to be recorded: e.g., up to 256 as in Flege and Fletcher (1992), and 1,000 as in Saito, Trofimovich, and Isaacs (2017). However, in contrast to the simple Likert scale, raters using a sliding scale are likely to be unaware of the individual gradients on the scale; in other words, raters may not be able to assign a clear accentedness value to each point on the scale (Jesney, 2004).

Brennan and Brennan (1981) proposed a method called Direct Magnitude Estimation. In this method, raters were asked to accord a numerical score to the first token presented and then rate subsequent speech tokens on the basis of the first token—for example, multiplying by two for a token deemed twice as accented as the first token. On the other hand, the study of Southwood and Flege (1999) aimed to determine whether foreign accent is a metathetic continuum (a continuum like pitch that can be partitioned into equal intervals) or a prothetic continuum (a continuum like amplitude that is not amenable to linear partitioning). The authors compared the accentedness of Italian speakers of English rated by native English listeners using direct magnitude estimation and a seven-point Likert scale whose endpoints were labeled as “least accented” and “most accented.” They found that the raters were able to partition accentedness into equal intervals. They interpreted this as an indication that foreign accentedness is a metathetic continuum and that interval scales can thus be effectively employed in accentedness rating. Southwood and Flege also reported ceiling effects in that Direct Magnitude Estimation enabled greater degrees of distinction among the tokens at the top end of the

scale than did the seven-point scale. They therefore suggest that use of a nine- or eleven-point scale may be ideal in exploiting listeners' full range of sensitivity.

In order to gain better understanding of how listeners actually arrive at accentedness judgments, recent studies have included qualitative analyses of raters' reports collected via post-rating surveys or interviews in addition to accentedness rating scores. Hayes-Harb and Hacking (2015), for example, carried out an extensive interview with ten native English listeners who rated the accentedness of five native Bosnian and five native English speakers. Despite high reliability obtained from the quantitative accentedness data (rated using a nine-point scale), listeners reported having activated a complex set of attitudes and expectations about speakers. Some listeners even tended to make inferences about the socio-cultural backgrounds of the speakers in addition to the phonological properties of the speech tokens. Such findings suggest that accentedness judgments likely result from a complex process that goes beyond a mere number assignment on a given scale.

Another important methodological question pertains to rater variability: variations in scores associated with rater characteristics, but not with examinees' actual performance or ability (Hsieh, 2011). Raters used in foreign accent studies have differed in several ways. The number of raters, for example, has ranged from merely one (Snow & Hoefnagel-Höhle, 1977) to more than 50: e.g., Kang (2010) used 58.

One of the rater-related variables is experience. The term *experienced* has been used to describe the listeners who were oral test raters (e.g., Anderson-Hsieh et al., 1992), language teachers (e.g., Hsieh, 2011; Kashiwagi & Snyder, 2010; Trofimovich & Isaacs, 2012), or linguistics majors (e.g., Hayes-Harb & Hacking, 2015). Experienced raters

were also expected to be familiar with the speech data presented to them and how that data may affect their accentedness judgments (Flege & Fletcher, 1992). On the other hand, inexperienced (or naïve) listeners have also been used in many studies (e.g., Calla McDermott, 1986), as it is assumed that L2 speakers likely have a greater chance to communicate with naïve listeners in real life than with experienced raters. Although Munro (1995) suggested that there is no need to provide raters with special training, raters commonly complete some practice trials before beginning the actual rating (Munro, Derwing, & Flege, 1999).

Several studies compared accentedness ratings given by raters with different amounts of experience. For example, Thompson (1991) compared experienced raters (those who "spoke at least one foreign language, had lived and studied abroad, had taken courses in linguistics, and had frequent contact with [L2 speakers' L1]" (p. 186)) with inexperienced raters (those who did not meet any of these criteria). Her results showed that although both groups of raters were able to accurately detect foreign-accented speech, the inexperienced raters generally perceived a higher degree of accent in nonnative speech than did the experienced raters. Additionally, a lesser degree of inter-rater reliability was obtained within the inexperienced group than within the experienced group. However, fairly good levels of inter-rater reliability were obtained in other studies, indicating that inexperienced raters may not necessarily be less reliable than experienced raters (e.g., Flege & Fletcher, 1992). On the other hand, many other studies reported no rater group differences (e.g., Bongaerts et al., 1997; Huang, 2013; Huang & Jun, 2015). Therefore, interpretation of comparative studies should be made with caution, since

raters' experience has represented traits such as rating experience, L2 experience, or knowledge in linguistics.

Hsieh (2011) investigated rater effects in the evaluations of nonnative speech samples made by ESL teachers (experienced raters) and American undergraduate students (naïve listeners). The undergraduate students were found to evaluate foreign accented speech more negatively—a factor which might pertain to the raters' amount of exposure to foreign-accented speech. Analysis of written comments produced by raters in addition to raters' assigned ratings showed that the ESL teacher raters tended to evaluate the speech more analytically than did the undergraduates by providing linguistically sophisticated descriptions of features (intonation or stress patterns) detected in the speech. Hayes-Harb and Hacking (2015) reported that their listener subjects—undergraduate native speaker students with varying amount of experience in linguistics—generally noticed similar properties of the nonnative speech samples while making accentedness judgments, and they only differed "somewhat in the level of sophistication with which they expressed themselves" (p. 61). Based on their data of accentedness ratings and comments made by novice and experienced raters, Trofimovich and Isaacs (2012) reported that it is difficult to describe phonological errors (especially prosodic errors, compared to segmental errors) without reference to specialized terminology.

Another common issue relates to whether or not listeners are native speakers of the target language. Earlier foreign accent studies relied on judgments of native speakers of the target language, who were the "models" of potential interlocutors in the setting where the L2 is the primary language for communication (Flege & Fletcher, 1992). More recently, however, accentedness ratings by nonnative listeners have received much more

attention. There has been a rise in the perspective that English is an international language (Jenkins, 2002), but according to Hayes-Harb (2014), this perspective "challenges even the very definition of accentedness as extent of deviance from a native speaker norm" (p. 47). To address this issue, some recent studies thus examined nonnative listeners' accentedness judgments in addition to native listeners' judgments. Although nonnative listeners may perceive lower degrees of foreign accent than native listeners, they are generally able to detect foreign accents (Major, 2007). For example, Yuan et al. (2010) compared English and Chinese listeners' assessments of foreign accent in spontaneous speech produced by speakers of eight different L1s. They found that the Chinese listeners perceived lower degrees of foreign accent than native listeners and were less sensitive to Chinese accents than to the other accents (e.g., French). The authors also carried out acoustic analyses which demonstrated that both native English and native Chinese listeners were sensitive to speaking rate and pauses in perception of foreign accent. On the other hand, phonetic deviation of segmentals and the F0 variation were important cues only in the native listeners' perception, but not in that of the nonnative listeners.

Other studies have noted that native and nonnative listeners perceive accents in fundamentally different ways. However, the results have been mixed. For example, Kashiwagi and Snyder (2010) found that in accentedness ratings given by Japanese and English listeners to Japanese speaker subjects, English listeners were sensitive to consonant and vowel errors, while Japanese listeners were more affected by pitch, stress, and rate of the speech tokens. On the other hand, Riney et al. (2005), who also had native Japanese and English listeners for rating, found that English listeners primarily used

segmental cues (e.g., /l, r/, vowels) in their perception of accent, whereas Japanese listeners primarily relied on non-segmental parameters such as pausing and intonation. Research in nonnative listeners' accentedness judgments is still in its infancy. Although it may be too early to make any generalizations, the topic of nonnative speakers' speech processing is expected to provide an insight to the understanding of accent and accentedness.

Recent studies in L2 speech assessments sought to explore whether raters' familiarity with a particular accent creates rater biases (Huang, 2013; Huang, Alegre, & Eisenberg, 2016; Winke & Gass, 2013). These studies reported that some of the raters—those who had cultural or ethnic ties to the language primarily spoken by the ratees—occasionally sympathized with the ratees and expressed concerns about the objectivity of their judgments. However, quantitative analyses of the data did not find any significant effect of raters' accent familiarity on their numerical ratings of accentedness.

In summary, previous foreign accent studies have employed various methods of collecting nonnative speech samples and assessing them in terms of accentedness. The studies have also discussed methodological concerns related to speaker populations, speech tokens, accentedness ratings, and listener/raters. Each of these factors needs to be taken into account in conducting and interpreting foreign accent research, and the current study sought to be a part of it. The following section reviews the design and the findings of a pilot study and discusses some lessons that provided insights for designing the current study.

3.2. Pilot Study

A pilot run of the current study was conducted examined speech performance of an intact group of adult ESL learners at advanced level. The pilot study had a single research question: what is the relationship between ESL learners' phonological awareness and their perceived degree of foreign accent? The study used the same definition of foreign accent as the one set in the current study. Phonological awareness, on the other hand, was understood as the explicit knowledge of L2 phonology, tied to consciousness. In this study, four sentences from each of the 12 ESL speakers' read speech were extracted for accentedness rating purposes. Four experienced ESL instructors judged the accentedness of each sentence on a nine-point scale: 1: very heavy foreign accent; 3: heavy foreign accent; 5: moderate foreign accent; 7: slight, negligible foreign accent; 9: no foreign accent. A mean score for each sentence of each L2 speaker was obtained. The ESL speakers also performed a set of L2 phonological awareness tasks: segmentation, blending, initial and final consonant deletion, word-initial and word-final cluster identification, and pseudoword reading. The task items were presented either visually or aurally, and the responses were given either orally or in writing.

Descriptive statistics of the accentedness ratings revealed that none of the speakers produced nativelike, accent-free English speech, despite the high proficiency level they had been assigned to according to their ESL course placement test results. Variance in rating among the speakers was also detected. A sub-group analysis (L1 Korean speakers versus speakers of other L1s) revealed that L1 Korean speakers were averagely rated lower (i.e., more accented) than speakers of other L1s for three of the four examined sentences. As for the phonological awareness tasks, the speakers exhibited

varying degrees of L2 phonological awareness, scoring in the middle ranges on average. Correlational analyses of the accentedness ratings and phonological awareness data provided an answer for the research question: overall, there was no clear pattern of association between accentedness and phonological awareness. Those who performed better on phonological awareness tasks (and thus were considered to be more sensitive to phonological structures) did not necessarily speak with a lighter foreign accent. In other words, the awareness did not seem to have translated into L2 production. But there was one exception: the accent ratings for one of the read sentences were significantly correlated with some of the phonological awareness tasks, such as blending and pseudoword reading. This particular sentence was found to contain a word which the ESL speakers were not familiar with and thus read dysfluently (e.g., pausing at the word, stuttering, or slowing down). This underscored the importance of considering potential influences of variables such as lexical familiarity and frequency on L2 speech performance.

Interestingly, the correlational analyses obtained different results for the L1 Korean speakers and for the speakers of other L1s. For the L1 Korean group, strong and significant correlations were found between some tasks and accentedness ratings and even between the aggregate means of phonological awareness and the means of the accentedness ratings, while such relationships were not found among other speakers. The difference between the accent ratings of the two groups suggests that speakers of a particular L1 (Korean in this pilot study) may face relatively more difficulty than speakers of other L1s in reading certain English sentences in a nativelike way.

The pilot study acknowledged several limitations. First, the scalar ratings of accentedness were based on native listeners' holistic judgments, not zeroing in on which feature(s) of the L2 speech—whether segmental, prosodic, or both (then, to what extent)—affected the perceived degree of foreign accent. The phonological awareness tasks, on the other hand, encompassed items that mainly focused on segmental features (e.g., consonants), with prosodic features largely unexamined. Moreover, as previously mentioned, phonological awareness in this pilot study did not include the implicit, non-verbalizable kind of L2 phonological knowledge, but only included explicit knowledge requiring use of metalanguage. Lastly, the study's exclusive use of reading data, which are not as naturalistic as spontaneous speech, creates a generalizability issue. These points show limitations of the pilot study in reflecting the constructs of accent and phonological awareness as well as their relationship. By expanding the pilot study and reflecting upon the above lessons learned from it, the present study aimed to gain deeper understanding of foreign accent, phonological awareness, and their relationship.

3.3. The Present Study: Research Questions and Overview

In light of the review of the literature and the concerns addressed in the previous sections, the present study proposed the following research questions:

RQ1. What constitutes a foreign accent?

RQ2. Is there a relationship between foreign accent and L2 phonological awareness?

RQ3. If so, what is the nature of that relationship?

This exploratory study was in a non-experimental design conducted in two phases. The first phase focused on the nature of foreign accent—how it is detected, perceived, and judged by native speaker (NS) listeners. It also identified phonological and phonetic factors of accentedness by analyzing accent perception data. More specifically, this study examined L2 English speech produced with a Korean accent, and it is based on North American English, which is the most widely accepted English accent in Korea. Although predictions of difficulties in English pronunciation have been made in previous research, there is little documentation of the actual surface characteristics of Korean-accented English speech (see Hong, Kim, & Chung, 2014). A close examination of a language-specific accent (here, a Korean-accent in English) was expected to provide an informative answer to the first research question. The second phase of the study sought to answer the second and third research questions. First, it probed phonological awareness of L2 speakers by administering tasks involving production and perception of L2 phonological units at segmental and prosodic levels. The remainder of the second phase explored the relationship between L2 phonological awareness and foreign accent, based on the data collected during the first phase.

3.4. L1 Korean Speakers' L2 English Speech

Previous research has reported that foreign-accented speech is largely attributed to L1-L2 cross-linguistic differences and transfer (e.g., Flege et al., 1999; Flege et al., 2006; Sereno et al., 2016). Since the current study delves into a foreign accent produced by L1 Korean speakers of L2 English, a brief introduction on the Korean phonology and characteristics of Korean learners' English would be necessary in interpreting the results

of the study. Vowels and consonants of Korean are shown in Tables 3.1 and 3.2, respectively. Vowels and consonants of (American) English are displayed in Tables 3.3 and 3.4, respectively.

Table 3.1

Vowels of Korean

Adapted from Lee (1996)

	Front	Central	Back
Close	i		ɯ **u**
Mid	e ɛ		**o**
Open		a	ʌ
Diphthongs	je jɛ ja jʌ jo ju ɥi we wɛ wa wʌ ɯi		

Note: Vowels in bold are rounded

Table 3.2

Consonants of Korean

Adapted from Kabak and Idsardi (2007)

		Bilabial	Alveolar	Post-alveolar	Velar	Glottal
Stop	Aspirated	p^h	t^h		k^h	
	Fortis	p’	t’		k’	
	Lenis	b̥	d̥		g̥	
Affricate	Aspirated			$tɕ^h$		
	Fortis			tɕ’		
	Lenis			d͡ʑ̥		
Fricative	Fortis		s’			
	Lenis		s			h
Nasal		m	n		ŋ	
Liquid			l			

Table 3.3

Vowels of (American) English

Adapted from Fromkin, Rodman, & Hyams (2003)

	Front	Central	Back
High	i		**u**
	ɪ		**ʊ**
Mid	e	ə	**o**
	ɛ	ʌ	**ɔ**
Low	æ		ɑ
Diphthongs		aj aw ɔj	

Note: Vowels in bold are rounded

Table 3.4

Consonants of (American) English

Adapted from Fromkin et al. (2003)

		Bilabial	Labio-dental	Inter-dental	Alveolar	Palatal	Velar	Glottal
Stop	Voiceless	p			t		k	
	Voiced	b			d		ɡ	
Affricate	Voiceless					ʧ		
	Voiced					ʤ		
Fricative	Voiceless		f	θ	s	ʃ		h
	Voiced		v	ð	z	ʒ		
Nasal		m			n		ŋ	
Liquid					l, r			
Glide		(w)				j	w	

The Korean monophthongs can be classified in terms of tongue height, backness, and lip rounding. Only the two back vowels /o, u/ are rounded. Korean lacks the vowel equivalent to /æ/ in English; accordingly, previous research showed that many Korean learners find it difficult to discriminate /æ/ and /ɛ/ even though Korean /ɛ/ is acoustically closer to English /ɛ/ than /æ/ (Tsukada, Birdsong, Bialystok, Mack, Sung, & Flege,

2005). In addition, the Korean vowels do not make the distinction between tense and lax vowels. The lax counterparts of /i/, /u/, and /o/ are not in the phonemic inventory as shown in Table 3.1. Studies on Korean learners' use of English vowels reported that Korean learners experience difficulty in both producing and perceiving tense-lax contrasts in English; for instance, English /ɪ/ was identified as four different Korean vowels /i, e, ɛ, ɯ/ (Tsukada et al., 2005), and English /ɔ/ was frequently substituted with /oʊ/ (Hong et al., 2014). Hong et al. also reported misidentification of English /ʌ/ as /ɑ/ or /æ/. Korean also has a range of diphthongs which are combinations of a glide approximant and a vowel, but not in the opposite order, as shown in Table 3.1.

Cross-linguistic differences in the consonantal system must be noted as well. While all of the Korean consonants shown in Table 3.2 (except for /ŋ/) can occur at the beginning of a word, only seven consonants /b̚, d̚, g̚, m, n, ŋ, l/ can occur in syllable-final position. Korean has typologically unusual three-way contrasts of stops and affricate: lenis (slightly aspirated), fortis (laryngealized and unaspirated), and aspirated (heavily aspirated). These consonants contrast in VOT, F0 at vowel onset, and close duration, unlike the English stop voicing contrast that is defined primarily by VOT (Schertz, Cho, Lotto, & Warner, 2015). The Korean stops and affricate are voiceless in word-initial position, while the lenis stops and affricate /d͡ʑ̚/ are voiced between voiced segments. Korean also has one glottal and two alveolar fricatives, which have a two-way contrast of lenis and fortis. The fricative /s/ is palatalized to [ʃ] before high front vowel /i/. There are three nasal consonants /m, n, ŋ/. The liquid /l/ is included as a phoneme but does not occur word-initially. Rather, its allophone flap [ɾ] appears in word-initial and intervocalic positions. Overall, the acoustic-phonetic detail of the aspirated stops and affricate, nasals,

and the fricatives /s, h/ is considered to be similar to the phonetic detail of corresponding English consonants (Schmidt, 1996). On the other hand, there are English consonants that are not in the Korean inventory (/f, v, θ, ð, z, ʒ, r/). They are "new" sounds that may cause pronunciation difficulty for Korean learners of English, as predicted by the CAH (Lado, 1957) and the SLM (Flege, 1995).

The syllable structure in the Korean language is much simpler than that of English, allowing only a maximum of one consonant in both the onset and coda positions of syllables. Korean learners of English tend to pronounce English consonant clusters with a vowel inserted between the consonants; for instance, 'please' /pliz/ is likely to be pronounced with an additional vowel between [p] and [l] (and possibly another after [z]) (Weinberger, 2015). Cross-linguistic differences in prosody also cause L2 pronunciation difficulties (e.g., Lee, Guion, & Harada, 2006). The rhythm structure of Korean is typically described as syllable-timed, unlike that of English, which is stress-timed. The pitch peak in Korean is typically aligned with the offset of the stressed syllable, whereas the pitch peak in English is aligned with the onset of the stressed syllable (Jun, 1996; Trofimovich & Baker, 2006). The study of Kim and Kim (2001) on Korean-accented English sentences reported that the pitch accent tended to fall on the last syllable of the phonological word in focus and that Korean speakers showed distinct tonal patterns for accentual phrases—such as the LHLH tone sequence, which sometimes gave the impression that they were asking a question rather than making a statement in English. According to the authors, these instances of Korean-to-English transfer seemed to have yielded intonation patterns in Korean-accented English that are clearly distinct from those of natively-produced English.

3.5. Phase 1

3.5.1. Participants

This study examined 40 native Korean speakers learning English as an L2. These L2 speaker participants were gathered from intermediate to highly advanced English proficiency level in light of the fact that foreign accent is a persistent (or fossilized) attribute of L2 speech that is detected even among advanced, fluent speakers (Yule, 2014). The L2 speaker participant group was composed of 28 females and 12 males, and their mean age was 27.8 years old, ranging from 20 to 46. Because the speakers' L2 use experience varied, they were subdivided into three groups according to the amount of experience in L2[1]: up to 2 years (Group 1, n = 14), 3 to 5 years (Group 2, n = 10), and 6 years and above (Group 3, n = 16). The L2 speakers in the third group were considered to be highly experienced in L2. None of the L2 speaker participants, even those in Group 3, had learned languages other than their L1 (Korean) before age 12. Table 3.5 presents a summary.

[1] In the current study, L2 experience is defined as meaningful L2 practice, characterized by consistent interaction with native speakers (Moyer, 2015): e.g., the length of residence (LOR) in the L2-speaking environment or intensive L2 instruction in immersion settings.

Table 3.5

Summary of L1 Korean-L2 English Speaker Participants' Backgrounds

Variables		Group 1	Group 2	Group 3
Gender	Female	11	5	12
	Male	3	5	4
Age	Mean	27.43	24.00	30.25
	SD	8.22	6.24	4.94
L2 experience	Mean years	1.57	3.70	11.56
	SD	.73	.64	3.64
Age of learning L2	Mean	25.85	20.30	18.69
	SD	8.50	6.41	6.91

Note: SD: standard deviation

Twenty-three native speakers of English served as judges for accentedness rating in Phase 1. The recruited raters, 17 females and six males, were raised and educated in a variety of regions in the United States: New York (6), New Jersey (4), Ohio (3), Illinois, Indiana, Maryland (2 each), Florida, Massachusetts, Michigan, and Utah (1 each). Although several studies found evidence that both expert and novice raters can give reliable ratings for accentedness (e.g., Huang & Jun, 2015; Yuan et al., 2010), the current study sought to employ NS listeners who are experienced in English teaching and have matalinguistic knowledge of English. Raters with these linguistic experiences were expected to aid in collecting extensive, analytical, and metaphonological reports to justify their accentedness judgments (cf. Hayes-Harb & Hacking, 2015; Saito et al., 2017; Trofimovich & Isaacs, 2012). The raters on average have taught English for 5.48 years in a second language setting (ranging from a half year to 27 years) and for 1.72 years in a foreign language setting (ranging from 0 to 10 years). The raters' experience in phonology and phonetics ranged from taking one course at a graduate level to teaching the subject or even earning a doctoral degree in phonology. Most of the raters reported

that they speak a language other than English (Spanish being the majority), although the reported proficiency levels varied. The raters also varied in familiarity with the Korean language or speakers of it. A summary is presented in Table 3.6. Lastly, all raters reported normal hearing and no history of hearing or speech problems.

Table 3.6

Summary of NS Rater Participants' Backgrounds

Rater	Gender[a]	Home state	Other languages (proficiency)[b]	Teaching (years)			Training in phonology[c]	Familiarity with Korean[d]
				ESL	EFL	Total		
E01	F	NJ	Korean (her), Chinese (int)	3	1	4	L	6
E02	M	UT	Spanish (low), Japanese (low)	17	3.5	20.5	LT	5
E03	F	NJ	Korean (her)	1	0	1	L	6
E04	F	OH	Spanish (flu), Portuguese (low)	1	1	2	L	2
E05	M	NJ	Italian (flu), Spanish (low)	1	0	1	L	1
E06	F	NY	Cantonese (low), German (low)	2	5	7	L	4
E07	F	NY	Chinese (her), Japanese (low)	3	0	3	LT	3
E08	F	IN	German (low)	1	0	1	L	5
E09	F	OH	None	1	0	1	L	1
E10	F	NY	Spanish (flu)	1	0	1	L	4
E11	F	NY	Spanish (flu)	.5	0	.5	L	4
E12	M	NY	Spanish (int)	.5	0	.5	L	4
E13	F	MD	Spanish (int)	4	5	9	LT	3
E14	F	MA	French (flu), Spanish (int)	9	0	9	LT	3
E15	F	MI	Korean (flu), Japanese (low)	4	5	9	LT	6
E16	F	IL	Bengali (her), Spanish (int)	12	0	12	LT	2
E17	F	FL	Spanish (flu)	2	4	6	L	5
E18	F	NY	French (int), Italian (int)	27	0	27	LT	5
E19	M	OH	French (int), Japanese (int)	18	10	28	LT	2
E20	F	IL	Arabic (low)	12	0	12	LT	2
E21	F	MD	Spanish (int)	4	5	9	LT	3
E22	M	IN	German (low)	1	0	1	L	5
E23	M	NJ	Italian (int), Spanish (int)	1	0	1	L	1

Notes: [a]Gender: F: female; M: male;
[b]Proficiency indicators: flu: fluent; int: intermediate; low: low; her: heritage
[c]Training in phonology: L: learned; LT: learned and taught
[d]Familiarity with Korean (based on the response to the question, "Have you been in contact with a native speaker of Korean (e.g., friend, student, or neighbor)? If yes, how regularly?" on the background questionnaire): 6: "Yes, almost every day;" 5: "Yes, 3-5 times a week;" 4: "Yes. 1-2 times a week;" 3: "Yes, 1-2 times a month;" 2: "Yes, but not as frequently as above;" 1: "No."

3.5.2. Instruments

Collection of L2 speech samples. The L2 speakers were asked to complete two speaking tasks: reading aloud and narration. The first task used the elicitation paragraph from *The Speech Accent Archive* (Weinberger, 2015), which is available for public use. The archive is composed of a large set of speech accents demonstrated in recordings of an English paragraph read by speakers from a variety of language and demographic backgrounds, allowing users to compare and analyze accent variables. The paragraph, consisting of 69 words as shown in Appendix A, contained various sounds and sound sequences which are absent in the Korean phonological system and others that are similar but not precisely (or acoustically) the same as the components in the Korean phonology. These elements are prone to cause pronunciation difficulties for Korean speakers of English, as predicted by the CAH (Lado, 1957) and the SLM (Flege, 1995). Eliciting speech data by having participants read texts aloud allowed for controlling for grammar and vocabulary so that speakers could pay less attention to meaning but more to phonological and phonetic forms when reading aloud. This was to minimize the possibility of the foreign accent ratings being influenced by grammatical errors or nontargetlike vocabulary use (Munro & Derwing, 2001).

Following the reading loud task, a spontaneous speaking task—narration, to be specific—was administered. Participants were asked to play the role of a tenant who has just moved into a new apartment and were given a picture of a room with several problems, including a broken window, a leaking faucet, and piled up trash (Image shown in Appendix A). The speakers needed to leave a voice message for the landlord complaining and describing the problems shown in the picture. Spontaneous speaking

tasks in general require less attention to form than do reading or metalinguistic tasks (Ellis, 2005, 2009), although they may not prevent non-phonological (morphosyntactic or lexical) errors that could influence accentedness judgments (Calla McDermott, 1986). Even so, inclusion of spontaneous speech data was intended to expand the examination of L2 speech carried out in the pilot study.

Collection of raters' judgments. For accentedness rating, the NS raters first received general rating instruction via email and were provided detailed instruction presented on the online rating forms. The instruction included two reading speech samples: one produced by a native English speaker and the other by a native Korean speaker, both of which were obtained from the *Speech Accent Archive*. The samples were provided as a baseline norm of the data. After norming, the raters were asked to judge the degree of foreign accent of each L2 speech sample on a set of rating scales. Based on evidence from the accent scaling literature, at least nine levels were necessary to capture the magnitude of accent ratings from widely variable L2 speech samples (Southwood & Flege, 1999). The individual raters were thus presented with a set of nine-point Likert scales (see Figure 3.1), adapted from Kang (2010):

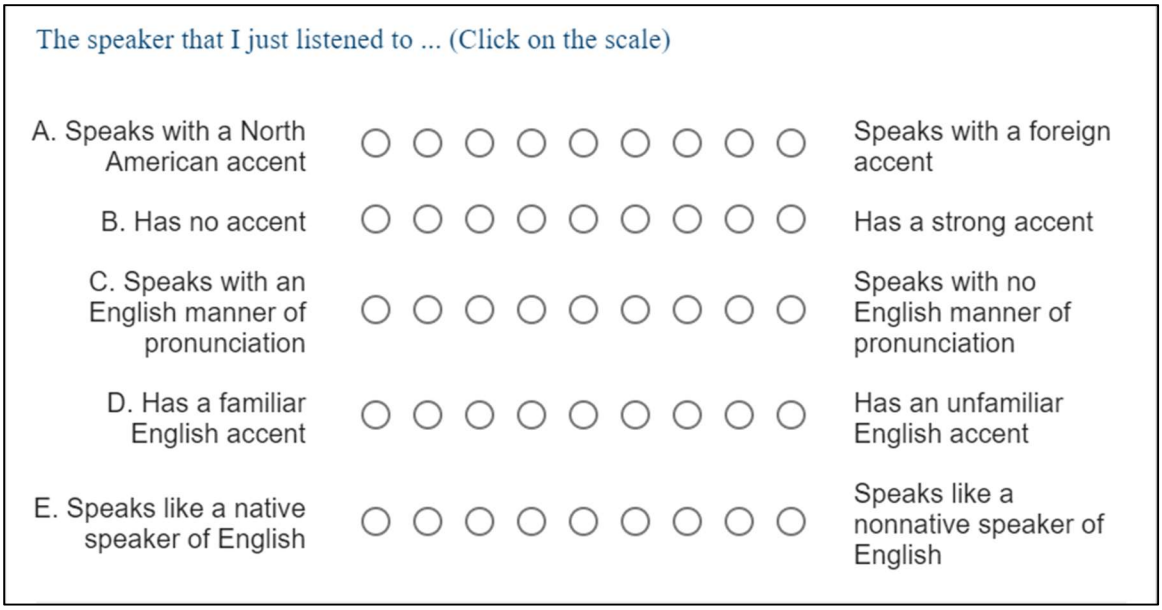


Figure 3.1 Accentedness Rating Scales

The accentedness rating scales consisted of five items which were intended to help raters approach the notion of accentedness in a dynamic way. It is also argued in Kang (2010) that an extended scale of this sort can increase item reliability, compared to a single item scale like those commonly used in L2 accent studies, which merely measures the accentedness (e.g., Munro & Derwing, 1995).

In addition to the rating scales above, the rating form also included open commentary slots where the raters can provide written comments about phonological/acoustic-phonetic parameters that they noticed during accentedness rating. The commentary slots were provided with a guideline displaying the phonological and acoustic-phonetic parameters in the perception of foreign accent that have been documented in previous research (see Table 3.7). A transcript of the speech sample was also provided. This was to circumvent the possibility in which the raters are distracted by

intelligibility or comprehensibility of the speech sample. The image used for the narration task also appeared on the rating form for narration samples.

Lastly, a questionnaire was provided to the raters for the purpose of collecting demographic and linguistic background information such as the number of years they have been teaching ESL/EFL. The questionnaire data were used as supportive evidence to elaborate and help account for the quantitative data results.

Table 3.7

Phonological/phonetic Parameters of Deviant Forms

(Provided to the NS Raters as a Guideline for Accentedness Rating)

Deviations	Descriptors	Examples
Segmental errors	Phonemic substitutions	*"thing"* spoken as *"ting"*
Syllable structure errors	Vowel/consonant insertion and deletion errors	*"spoon"* spoken as *"supoon"*
Lexical stress	Incorrect stress placement in polysyllabic words	*"PLAstic"* spoken as *"plaSTIC"*
Phrasal stress, or rhythm	Failure in vowel reduction in both polysyllabic words and function words	*"ASK her to BRING"* spoken as *"ASK HER to BRING"*
Intonation, or pitch contour	Inaccuracy in intonation at the end of phrases (i.e. syntactic boundaries)	*"Please call Stella* [falling tone]" spoken as "*Please call Stella* [level tone]"

3.5.3. Procedure

Collection of L2 speech samples. The L2 speakers were tested individually in a quiet room equipped with a recording device—a digital voice recorder with a microphone. Brief instructions were provided to the participant before beginning each

speaking task. The reading and spontaneous speech data were audio-recorded, and the intensity of all speech files was normalized (75 dB) using PRAAT (Boersma & Weenink, 2017, Version 6.0.25). While the reading samples were about 30 seconds long on average, the lengths of the spontaneous speech samples varied, ranging from 40 seconds to two minutes long. The long speech samples were thus trimmed to 45 seconds long in order to minimize the possibility in which raters are affected by the overall length of the discourse while giving accentedness judgments. Transcription of the spontaneous speech data was also created. After audio-recording, the L2 speakers were asked to fill out a questionnaire that was designed to collect background information such as their age, language learning experience, and desire for pronunciation accuracy.

Collection of raters' judgments. For accentedness rating, the NS raters first received general rating instruction via email and were provided detailed instruction presented on the online rating forms. The raters were asked to use the full range of the scale. Because online rating forms were used, the raters assessed samples at their own, individual pace, and were allowed to replay and listen to the sample as many times as they needed. The raters used their own headsets while listening to the speech samples. In addition to the accentedness judgments given on rating scales, the raters were asked to justify their ratings, by describing the aspects of nonnative speech that they noticed while rating for accentedness. The raters were asked to refer to the guideline presenting categories of accent features (e.g., consonants, vowels, syllable structure, lexical and phrasal stress, and pitch; see Table 3.7). This analytical, qualitative assessment was expected to provide insight into how listeners actually arrive at accentedness judgments and thus strengthen the internal validity of the study (e.g., Hayes-Harb & Hacking, 2015;

Trofimovich & Isaacs, 2012). However, the analytical assessment comments were not mandatory, thus there was a variation in the numbers of comments made by the raters. The L2 speech samples were presented in a randomized order, and only one sample was presented per rating page. The raters gave scores to all of the L2 speech samples: 40 in reading and 40 in narration, 80 in total.

3.5.4. Coding and analysis

The accentedness rating data consisted of the analytical comments and the rating scores. The raters' qualitative assessment of accentedness was compiled and categorized into comments related to particular parameter categories: segmentals and prosody. The codes were inductively extracted from the data. The purpose of collecting the qualitative assessment data was not to assign scores to individual L2 speaker participants, but rather to examine the phonological factors that the raters attended to while judging the degree of Korean accent in L2 English speech. The rating scores were submitted to quantitative analyses: individual L2 speakers' scores were examined for internal consistency and were computed into a composite measure for analyses for Phase 2.

3.6. Phase 2

Phase 2 of the study measured phonological awareness of the L1 Korean-L2 English speakers from Phase 1 and analyzed the data with the accent judgment data obtained in Phase 1.

3.6.1. Operationalization of L2 phonological awareness

As described in the earlier chapters, L2 phonological awareness in this study was conceptualized as explicit and implicit knowledge of L2 phonological systems. Operationalization of this construct therefore required consideration of the construct of *knowledge* and its scope pertaining to a particular linguistic domain, *phonology*.

In previous studies, explicit knowledge has been operationalized as "learners' explanation of specific linguistic features" (Ellis, 2009, p. 34). Researchers have elicited this type of knowledge typically by asking learners to verbalize specific grammatical rules, often involving use of metalanguage. Implicit knowledge has been determined by examining learners' use of these features in oral or written language, and spontaneous production tasks have been most commonly used for elicitation. In light of this, the current study operationalizes *explicit knowledge* as the ability to *explain* specific linguistic features (verbally, possibly with metalanguage) and *implicit knowledge* as the ability to *use* the target features. The use of L2 phonological features involves core speech processing skills such as identification and manipulation of the components of the L2 phonological system (Cassady et al., 2005; Snow et al., 1998). In short, L2 phonological awareness in this study is operationalized as the learners' ability to (a) provide explanation and (b) demonstrate use of L2 phonology as they identify and manipulate specific L2 phonological features.

3.6.2. Participants

The same Korean learners of English who participated in Phase 1 participated in Phase 2 (N = 40).

3.6.3. Instruments

For assessment of the L2 speaker participants' phonological awareness, the study used five tasks designed to touch upon the knowledge of the L2 phonological system at gradient levels of implicitness/explicitness. This was also to avoid using commercially available phonological awareness test batteries which mainly rely on test takers' metalinguistic knowledge of L2 segmentals.

Task 1 was repetition of pseudowords, which involved both perception and production of L2 sounds and aimed to tap into the implicit phonological knowledge of the L2 speakers. Repetition requires speakers to first perceive—that is, to detect and identify the foreign sound(s)—then, reproduce the mental representations of it/them to the same sound as the input. Repetition thus involves awareness at a level higher than mere perception but lower than the level of understanding, since the participants are not required to demonstrate metalinguistic knowledge of the sounds they heard (Mora et al., 2014; Schmidt, 1990). Ellis (2005, 2009) also noted that oral imitation tasks, especially when timed and intended to focus the participants' attention on meaning, can be an effective measure of implicit knowledge of a language. In order to prevent direct imitation from sensory memory (Flege et al., 2006), the participants were directed to give a response after a short delay.

Task 1 used twenty pseudowords—sequences of phonemes (or segmentals) that conform to the phonotactic constraints of the language, without holding any semantic content. This was to control any possible effects of the participants' L2 lexical knowledge on performance. The pseudowords were collected from Gathercole, Willis, Baddeley, and Emslie (1994) and Elgort (2011). The pseudowords, all multisyllabic (having three to five

syllables), were presented aurally. The audio files of the stimuli were recorded by a female American speaker (age 27), born and raised in New York. The length of the stimulus containing a pseudoword did not exceed two seconds in total, following Baddeley, Thomson, and Buchanan's (1975) finding that an item rehearsable within two seconds can be retained in the phonological loop of the working memory.

For each item in Task 1, the participant was presented with a random image on the screen and heard an audio file that began with an introductory phrase *'This is…'* followed by a pseudoword (e.g., *bannifer*), and after a short pause, asked the participant to repeat the name of the introduced item—i.e., the pseudoword. Then, after a beep sound, the participant imitated the pseudoword as closely to the stimulus as possible, in terms of both segmentals and prosody. The participants were asked to focus on the pseudowords only, and their oral responses (i.e., repetition of pseudowords) were audio-recorded. One practice item was provided.

Task 2 was designed to measure the participants' awareness of L2 segmentals. It required the participants to perceive L2 sounds, twenty in total, and attend to whether a specified L2 phonological unit (phoneme) is present or absent in the stimuli. Similarly to Task 1, Task 2 also used pseudowords to minimize the influence of the participants' existing L2 lexical knowledge. The pseudowords, all monosyllabic, were collected from the ARC Nonword Database (Rastle, Harrington, & Coltheart, 2002). Each stimulus consisted of (a) an existing English word containing the target segment (presented aurally and visually; e.g., /p/ as in *pop,* while the IPA symbol was not displayed) and (b) four pseudowords (presented aurally only). The audio files of the pseudowords and example words were recorded by the same American speaker from Task 1.

For each item in Task 2, the participant was presented with a target phoneme and its example both aurally and visually. Then, the participant heard four pseudowords that either had or did not have the target phoneme. While listening, the participant was required to determine whether each of the pseudowords had the target phoneme and circled the one(s) that had the target phoneme. The sound files were played up to two times, although most participants did not need a replay. The participants provided answers using an answer sheet, in which the spellings of the pseudowords were not shown. One practice item was provided.

Task 3 required the awareness at the levels of perception and of noticing of L2 prosody—particularly, word level stress, which is generally realized by combining higher pitch with increased intensity and vowel length. Some basic metalanguage, such as stress and syllable, was also used. The task used twenty multisyllabic pseudowords collected from Gathercole et al. (1994) and Elgort (2011). The pseudowords in Task 3 were recorded by the same speaker from Tasks 1 and 2 and were presented to the participants aurally.

For each item in Task 3, the participant was presented with a pseudoword whose syllables were divided in slots, and was asked to circle the most, primarily-stressed syllable after hearing the pronunciation of each pseudoword. The responses were given on an answer sheet, on which syllables of the pseudowords were divided in slots. One practice item was provided before the actual task. The participant was allowed to listen to the item up to two times, but most participants did not need a replay.

Task 4 assessed the awareness at the levels of perception and noticing of L2 intonation encompassing pitch contour and rhythm. The task used 15 written sentences

and questions that do not have any punctuation and thus remain ambiguous in meaning. Identification of the intended meaning of the text was only possible by hearing the intonation of the text. For instance, a text written as “What is he eating” can be realized as a wh- question a falling boundary tone (e.g., “What is he eating (HL%)?” or as a yes/no question with a rising tone (e.g., “What, is he eating (LH%)?”). The former case would be followed by a response like “Dinner,” while the latter case is expected to have an answer in a yes/no form. Another classical example of intonational ambiguity is “Would you like some coffee tea or milk?” from Goldsmith (1978). If asked, “Would you like some coffee (LH), tea (LH), or milk (HL%),” the respondent is expected to make a choice out of the three items; on the other hand, “Would you like some coffee, tea, or milk (LH%)” intends to have a yes/no response. The participant’s performance on this task would be an indicator that he or she has phonological awareness of English intonation. The intonation of the text stimuli was produced and recorded by the American speaker from the previous tasks. The recordings were low-pass filtered (removing all sounds above 400Hz) so that they maintained the prosodic characteristics such as stress, rhythm, pitch, and intensity which the participants needed to focus on. An answer sheet was used to collect the responses—i.e., identification of the meaning of the text.

In Task 4, the participant was first presented with written text without any punctuation. Then, they heard the intonation of the text and were asked to choose the appropriate response or meaning shown on the answer sheet. There were a couple of practice trials before the actual task, and replay of the speech was allowed up to two times.

Lastly, Task 5—explicit verbalization of L2 phonological knowledge—assessed learner participants' ability to notice the difference between the native and nonnative pronunciation and collected their verbal reports. The participants were aurally presented with four pairs of speech samples, each of which consisted of a phrase produced by a native speaker of English (reading English words and phrases) and a phrase produced by a Korean speaker of English (reading the same words and phrases). The speech samples were from Weinberger's (2015) *The Speech Accent Archive*. Similarly to Task 1, the length of each item in this task did not exceed two seconds. After hearing the audio stimuli, the participants were asked to detect and identify the gap(s) between the two (i.e., NS-NNS samples), and then verbally explain how the items of each pair (i.e., one native and the other Korean-accented) are distinct from one another. The participants were asked to address all phonological/phonetic characteristics they could identify and analyze from the stimuli. In other words, the Korean speakers had to explicitly describe the phonological and phonetic differences between the native and Korean-accented English productions they noticed. This task thus involved awareness perception, noticing, and understanding of L2 phonemes and prosody. The task also required use of metalanguage to a certain extent because the participants had to use terms such as consonants, vowels, stress, or intonation. The verbal responses were expected to reflect the nonnative speakers' explicit knowledge of the L2 phonological system across levels of segmentals and prosody.

The participants were given freedom to replay as many times as they needed, and were permitted to take notes if they would like so that they could spare cognitive load to both phonological knowledge and memory. The participants were allowed and

encouraged to respond in their native language, Korean, in order to give as much verbal explanation as possible. The free oral responses were audio-recorded and transcribed.

3.6.4. Procedures

The phonological awareness tasks were administered to each participant individually in a room equipped with a monitor screen that displayed visual task items as well as audio-recording devices and a headset. The five phonological awareness tasks were presented to the participant one by one, with 1- or 2-minute break in between. Before each task, the participant received written and aural instructions primarily in both English and Korean so that he or she could best understand the requirements of each task. For each task, one or two examples of the task items were also provided to help the participant's comprehension of the task requirements.

With the exception of Task 1, which required the speakers to imitate stimuli promptly, the phonological awareness tasks in this study were not time-pressured. Responses from Tasks 2, 3, and 4 were collected using answer sheets, while those from Tasks 1 and 5 were audio recorded and digitized for coding and analyses. All task items underwent a pilot run with two native English speakers and two L1 Korean-L2 English speakers prior to data collection.

3.6.5. Coding and analysis

The L2 speaker participants' performance on phonological awareness tasks was coded for analyses. The audio-recorded responses made for Task 1 were evaluated for repetition accuracy, based on phonetic analyses of the responses. Two phonetically-trained judges listened to each production sample and compared it to the original prompt.

For each of the repeated pseudowords, the presence of segmental errors (such as phoneme substitution) and the presence of prosodic errors (such as incorrect stress placement) were determined dichotomously. The original stimuli functioned as the baseline data upon which errors were judged. An accurately repeated, error free word was scored 1 point for segmental accuracy and 1 point for prosodic accuracy. A word with an error (or more than one error) was scored 0. The reliability between the two judges was high: $\alpha = .88$. The coding scheme and examples are presented in Table 3.8.

Table 3.8

Coding Scheme and Examples for Task 1

Pseudoword	Segmental accuracy score	Prosodic accuracy score
For each:	1: accurate repetition 0: presence of segmental error(s), e.g., phoneme addition, deletion, or substitution, detected audibly or acoustically	1: accurate repetition 0: presence of prosodic errors, e.g., incorrect lexical stress, detected audibly or acoustically
e.g., *bannifer* If pronounced as [ˈbænɪfər], then scored as:	1 (accurate segmentals) *Note: Use of neutral vowels (e.g., /ə/) / for /ɪ/ would be accepted.	1 (accurate stress/rhythm)
If pronounced as [bæniˈpər], then scored as:	0 (substitution of /ɪ/ with /i/, which can be shown from formants; substitutions of /f/ with /p/, which can be shown from spectrograms)	0 (inaccurate stress/rhythm, which can be shown from pitch contour and waveform)

Performance on Task 2 was assessed in terms of sensitivity to the presence and absence of a signal (cf. Signal Detection Theory). The discrimination index, *d'* scores, were computed in order to reduce possible effects of response bias. In doing this, each of

the participants' responses made on the answer sheet were first coded as 'Hit (presence of signal is accurately perceived),' 'False Alarm (signal is perceived to be present while being absent),' 'Miss (signal is perceived to be absent while being present),' or 'Correct Rejection (absence of signal is accurately perceived).' Then, *d'* of the individual participant was computed using the z-score (standard deviation) of the signal distribution and rates of Hits and False Alarms: $d' = z(\text{Hit rate}) - z(\text{FA rate})$. An example of coding is illustrated in Table 3.9.

Table 3.9

Coding Scheme and Examples for Task 2

e.g., ***pop*** Displayed on the answer sheet	**A**. *(speem)*	**B**. *(graff)*	**C**. *(purf)*	**D**. *(jamped)*	
If responded as:	**A**.	**B**.	**C**.	**D**.	
Then scored as:	Miss (/p/ is present in *speem*)	False Alarm (/p/ is absent in *graff*)	Hit (/p/ is present in *purf*)	Hit (/p/ is present in *jamped*)	Hit rate: .05 FA rate: .03 *d'* = .29

The data obtained from Tasks 3 and 4 were coded for accuracy: the responses made on the answer sheet were assessed for accurate perception of the primary stress of each pseudoword (Task 3) and for accurate perception of the intonation of each stimulus (Task 4). A correct response was given 1 point, while an incorrect response was given 0. The scores were aggregated in each task.

Performance on Task 5 was assessed by tallying the number of phonological and phonetic differences (consonants, vowels, and prosodic features) correctly explained by

the participant. The verbal responses were first transcribed and segmented into information units, yielding 861 entries in total. During the initial round of coding, each entry was labeled with a keyword and type of the noticed forms as follows: segmentals, prosody, and unspecified at the word level. The entries were then coded for levels of complexity based on the coding system below (Table 3.10), which is partially adapted from Wrembel (2015).

Table 3.10

Coding Scheme for Task 5

For each L2 speaker, obtain the number of entries in which noticing of NS-NNS differences takes place, and code for:	
A. Accuracy of the explained nonnative features (accurate vs. inaccurate)	
- Accurate	
- Inaccurate	
B. Pronunciation domain of the explained NS-NNS difference	
Code	Description
- S: Segmentals	Vowels and consonants
- P: Prosody	Lexical stress, phrasal stress and rhythm, intonation, fluency (speech rate, pause)
- U: Unspecified at word level	The domain of the difference is not specified
C. Levels of complexity	
Code	Description
- Level 0: Perception	Perceiving differences
- Level 1: Noticing	Noticing and attentional focus on relevant forms
- Level 2: Understanding	Description or explanation of relevant forms
- Level 3: Understanding with metalanguage	Description or explanation of relevant forms using metalanguage

Statistical analyses followed the coding procedure. Descriptive statistics of the data were expected to provide an overview of the phonological awareness scores, encompassing central tendency, score ranges, and normality of distributions. These

analyses were also expected to reveal whether the L2 speaker participants performed significantly better or worse on particular phonological awareness tasks. Correlational analyses, stepwise multiple regression analyses, and additional qualitative analyses were conducted to investigate the relationship between phonological awareness and accent—its perceived degrees as well as phonological characteristics.

IV – RESULTS

This chapter aims to present and describe results obtained from Phases 1 and 2 that are described in the previous chapter and answer the research questions proposed in the current study.

4.1. Phase 1

The main goal of the first phase of the study was to identify phonological characteristics associated with accentedness. The study examined comments that NS raters provided upon making accentedness judgments. This was to find an analytical answer to the first research question of the study: *What constitutes a foreign accent?*

This section first reports the findings from the analyses of the written comments in which the NS raters described the phonological features of foreign-accented L2 speech. It then reports the findings from the quantitative analyses of the accentedness rating scores that the raters assigned to each L2 speaker participant. Lastly, the section closely inspects the judgments of selected rater participants to better understand their accentedness rating behaviors.

4.1.1. Raters' analytical comments

During the accentedness rating, the raters focused their commentaries on the L2 phonological and phonetic features that contributed to the L1 Korean-L2 English speakers' accentedness. The comments were analyzed both qualitatively and

quantitatively to characterize a Korean accent of English. During the first cycle of coding, each of the written comments was labeled according to the keyword, main idea, and the nontargetlike phonological feature type specified by the comment providers. The comments were then classified into main categories, based on their contents: segmental issues (substitution, insertion, and deletion of consonants or vowels) and prosodic issues (problems associated with lexical stress, phrasal stress, intonation, and speech fluency, such as speech rate and pauses). Examples of the comments coded for each coding category are presented below to illustrate the specific aspects of accented speech commented by the raters. Examples 1 through 8 were coded as "segmental issues" and were further classified into "substitution" (Examples 1–4), "insertion" (Examples 5–6), and "deletion" (Examples 7–8). Examples 9 through 13 were coded as "prosodic issues" and were classified into "lexical stress" (Example 9), "phrasal stress" (Example 10), "intonation" (Example 11), and "fluency" (Examples 12–13).

Example 1: "brother" as "broder" [E17_K35_R][1]
Example 2: "floor" pronounced as "ploor" [E20_K32_N]
Example 3: Line 4 "cheese": vowel not high enough [E14_K01_R]
Example 4: "snek" for "snack" [E01_K35_R]
Example 5: "luggage" as "lurgage" [E17_K03_N]
Example 6: Line 3: "six" » "six-eu" [E03_K10_R]
Example 7: "outlet" was said as "owlet" in line 2 [E05_K05_N]
Example 8: "situation" spoken as "sit-ation" [E16_K08_N]
Example 9: "winDOW" in line 3 [E02_K02_N]
Example 10: "She can SCOOP" » "She CAN scoop" [E09_K03_R]
Example 11: Rising intonation on "station" [E20_K31_R]
Example 12: A fast taker. She reads through the sentences very quickly which might have affected her performance. [E16_K19_R]
Example 13: Too long pauses between each word when saying "from the store" [E17_K07_R]

[1] The label of a commentary example consists of the following information: Rater, L2 speaker, and sample type (Reading or Narration); e.g., [E01_K01_R] = Rater E01's comment on Speaker K01's Reading

The raters' description of accent features showed some patterns. One of the widely used ways was to pinpoint the foreign-accented word and transcribe the problematic sound(s) as in most of the examples above (Examples 1–2, 4–8). The transcription, however, did not use conventional phonetic alphabets because the responses were typed in on the online rating form. In the comments on prosodic issues, the stressed syllable or word was typed in capital letters (Examples 9–10) or marked with quotation marks. Several raters provided even more analytical comments using some metalanguage (e.g., Examples 3 and 11).

There were other "free" comments, such as holistic impression of the nonnative speech (Examples 14 and 15).

Example 14: This person more or less sounds like a native speaker. [E22_K02_R]
Example 15: Speaker has a lisp. [E23_K31_R]

It was also revealed during the first coding cycle that some comments were made on simple reading mistakes; for instance, while reading the given text, some L2 speakers read *snake* as *steak* (substituting /n/ with /t/) or *we* as *she* (substituting /w/ with /ʃ/), or mixed the words *snake* and *snack* that occurred in adjacent lines, which seemed to be a reading mistake rather than a pronunciation issue. There were also comments concerning lexical and grammatical errors that emerged in the narration samples (e.g., plural markers added to noncount nouns such as *garbage* and *trash*). Such comments that are not relevant to the pronunciation were classified as "irrelevant." Table 4.1 presents the frequency counts of the four coded main categories commented by the raters. Figure 4.1 graphically illustrates the proportions of the written comments coded for each main category.

Table 4.1

Frequency Counts and Proportions of Written Comments

Main Category	Reading	Narration	Total
Segmental issues	2316 (31.64%)	1905 (26.03%)	4221 (57.67%)
Prosodic issues	1650 (22.54%)	1039 (14.20%)	2689 (36.74%)
Holistic impression	48 (.65%)	62 (.85%)	110 (1.50%)
Irrelevant to pronunciation	210 (2.87%)	89 (1.22%)	299 (4.09%)
Grand total	4224 (57.70%)	3095 (42.30%)	7319 (100%)

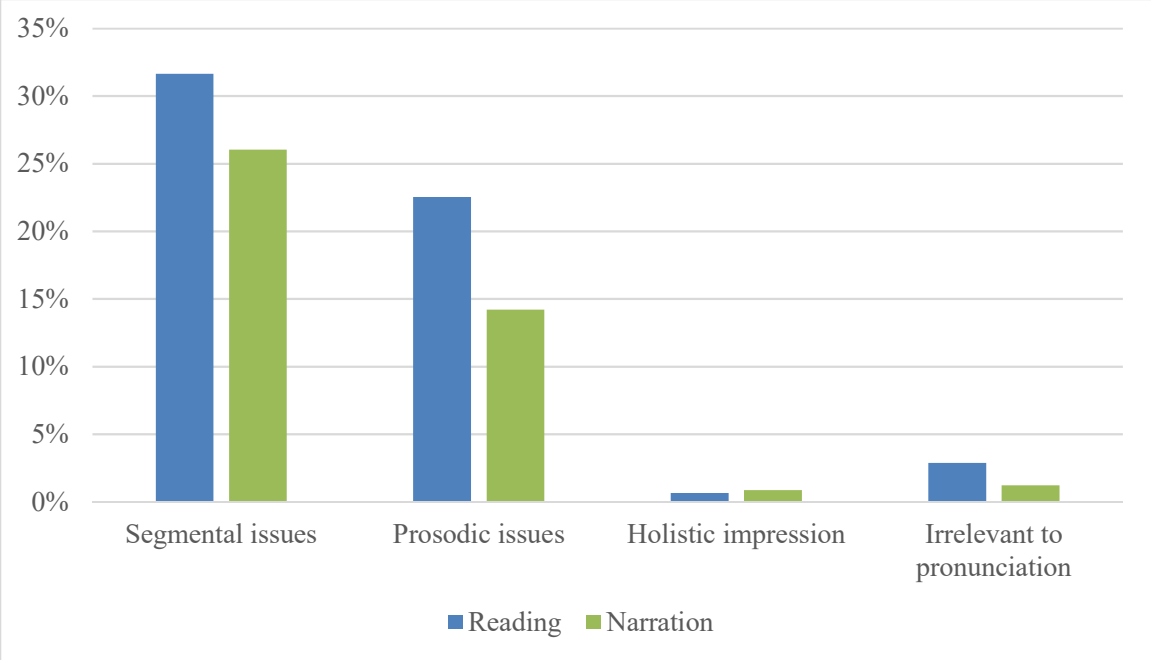


Figure 4.1 Proportion Distribution of Comments Coded for the Main Categories

A Korean-English bilingual speaker who is experienced in English phonology as well as EFL teaching served as a second coder. He coded 20% of the comments (n = 1,488, randomly selected). Because qualitative data were coded into nominal variables, a

Cohen's Kappa was run to compute the inter-coder reliability for the selected comments. The Kappa coefficient showed a high agreement, $\kappa = .87$, $p < .01$ (95% CI .82 to .93).

The remainder of this section focuses on the comments in the first two main categories (i.e., segmental and prosodic issues) in characterizing a Korean accent of English. It is also important to note in interpreting the current data that the coded comments were the individual raters' subjective judgments exhibiting inevitable variations; for instance, none of the accent features in the coded categories were unanimously agreed upon by the raters in this study (see Tables 4.3–4.11). Thus, the quantitative analyses in this section include not just the tallies of the coded categories but also the counts (and proportions) of the raters who commented on each of the coded categories.

The first main category, segmental issues, comprises three subcategories: substitution, insertion, and deletion. Table 4.2 presents the frequency counts and proportions of these three subcategories.

Table 4.2

Frequency Counts and Proportions of Commented Items: Segmental Issues

Subcategory	Reading	Narration	Total
Substitution	1682 (39.85%)	1480 (35.06%)	3162 (74.91%)
Insertion	507 (12.01%)	270 (6.40%)	777 (18.41%)
Deletion	127 (3.01%)	155 (3.67%)	282 (6.68%)
Total	2316 (54.87%)	1905 (45.13%)	4221 (100%)

In both reading and narration datasets, the raters made a substantial number of comments pertaining to substitution of a consonant or vowel (3,162 comments in total), indicating that they were highly sensitive to nontargetlike use of segmentals in L2 speech.

The comments in this subcategory are summarized in Table 4.3. These comments were further classified into substitution of consonants (see Examples 1 and 2 above) and substitution of vowels (see Examples 3 and 4). In addition, a relatively smaller number of comments were coded as "unspecified." These comments included those in which the raters did not specify the problematic segmental feature while typing in their comments in the slots made for segmental issues (see Examples 16, 17, and 18).

Example 16: "blue" is mispronounced in line 4 [E02_K21_R]
Example 17: phonemic substitute in "these" in line 2 [E10_K04_R]
Example 18: pronunciation of "apartment" is off [E17_K01_N]

These unspecified comments were excluded in narrow analyses of the reported substitution of segmentals because it was not clear from these comments regarding which segmentals the raters were referring to and, further on, which substitution case(s) affected the L2 speakers' accentedness. However, it is important to note that approximately 74% of the raters in this study left such unspecified comments at least once (338 comments in total; see Table 4.3). This result suggests that the raters, despite their backgrounds in English language teaching and formal training in linguistics, did not always perform as expected but ended up providing limited information when describing their realization of L2 accent features.

Table 4.3

Summary of Commented Items (Segmental Substitution)

	Reading	Narration	Total		Example(s)
	Counts of comments	Counts of comments	Counts of comments	Counts (%) of commenting raters	
Consonants					
Fricatives					
ð » d	165	230	395	17 (73.91%)	1, 19
ð » t	4	1	5	3 (13.04%)	
θ » t	125	14	139	18 (78.26%)	20
θ » d	33	12	45	7 (30.43%)	
θ » s	5	5	10	6 (26.09%)	21
s » θ	1	8	9	6 (26.09%)	
f » p	30	25	55	15 (65.22%)	2
v » b	2	10	12	9 (39.13%)	
f » h	8	0	8	5 (21.74%)	
z » ʤ	1	8	9	3 (13.04%)	24
z » s	19	19	38	10 (43.48%)	22, 23
s » ʃ	3	6	9	7 (30.43%)	
Liquids					
l » r	70	123	193	20 (86.96%)	25, 26, 27, 99, 102
r » l	86	51	137	19 (82.61%)	28
r » w	11	11	22	12 (52.17%)	29
ɾ » r	0	28	28	11 (47.83%)	30
ɾ » t	0	5	5	4 (17.39%)	31
Obstruents					
b » p	57	8	65	12 (52.17%)	32
d » t	13	17	30	14 (60.87%)	
g » k	21	5	26	8 (34.78%)	33
p » b	0	3	3	3 (13.04%)	
t » d	2	3	5	5 (21.74%)	
k » g	2	0	2	2 (8.70%)	
ʤ » ʧ	0	5	5	3 (13.04%)	
p, t, k » p^h, t^h, k^h	3	21	24	7 (30.43%)	34
p » f	10	0	10	7 (30.43%)	
b » f	4	3	7	5 (21.74%)	
b » v	5	9	14	8 (34.78%)	35
Others (e.g., w » h)	33	29	62	12 (52.17%)	
Consonants total	713	659	1372	23 (100%)	
Vowels					
i » ɪ	204	72	276	17 (73.91%)	3, 38
i » ɛ	13	1	14	11 (47.83%)	
ɪ » i	119	136	255	17 (73.91%)	39
ɪ » æ	5	3	8	7 (30.43%)	
e » ɛ	11	1	12	6 (26.09%)	

ɛ » æ	8	23	31	11 (47.83%)	46
æ » ɛ	67	62	129	14 (60.87%)	4, 45
u » ʊ	49	3	52	12 (52.17%)	40
u » o	14	1	15	9 (39.13%)	
ʊ » u	0	7	7	7 (30.43%)	41
o » ɔ	6	11	17	10 (43.48%)	37
ɔ » o	165	41	206	21 (91.30%)	36
ɔ » ʌ	6	0	6	6 (26.09%)	
ɑ » a	29	4	33	9 (39.13%)	48
ɑ » o	2	27	29	10 (43.48%)	47
ʌ » o	0	25	25	11 (47.83%)	
ʌ » ɑ	0	13	13	8 (34.78%)	
ə » ɯ	6	29	35	7 (30.43%)	43
ə » ʌ	94	64	158	18 (78.26%)	42
ə » ɑ	10	23	33	9 (39.13%)	44
ə » o	16	61	77	12 (52.17%)	
ə » ɛ	0	7	7	6 (26.09%)	
Others	5	9	14	6 (26.09%)	
Vowels total	829	623	1452	22 (95.65%)	
Unspecified	140	198	338	17 (73.91%)	16, 17, 18
Grand total	713	659	1372	23 (100%)	

Subsequent, narrow analyses of the comments on substitution showed deviation patterns concerning phonemic and phonetic categories that were reported as main sources of pronunciation difficulties experienced by Korean speakers of L2 English (e.g., Hong et al., 2014). Regarding consonantal issues, a large number of raters pointed out the substitution of interdental fricatives (/ð/ and /θ/) with dental stops of similar place of articulation /d/ and /t/, and marginally, /s/, as shown in Table 4.3 and the examples below.

Example 19: “the” as “da” [E17_K01_N]
Example 20: “thick” sounds like “tick” [E04_K05_R]
Example 21: “thick” » “sick” [E09_K08_R]

According to 17 (73.91%) of the raters, the substitution of /ð/ with /d/ was detected mostly in word-initial position (331 comments), as compared to word-medial or final positions. Only a small number of comments reported the substitution of /ð/ with /t/ (i.e., five comments provided by three (13.04%) of the raters). While this figure may suggest that the L2 speakers maintained the [+ voice] feature of /ð/ during substitution, a

different substitution pattern was found with /θ/, the [- voice] counterpart: voiceless /θ/, particularly in word-initial position, was substituted not only with /t/ (perceived by 18 (78.26%) raters; Example 20) but also with /s/ (perceived by six raters; Example 21). It is also interesting that there is little difference between the number of raters who perceived the substitution of /ð/ with /d/ (73.91% of the raters) and that of those who perceived the substitution of /θ/ with /t/ (78.26% of the raters), even though a significantly larger number of comments referred to the former type of substitution (359 comments vs. 139 comments). This result indicates that the number of comments alone is not sufficient in examining the degree of noticeability, or significance, of the pronunciation issue.

It is also important to note that the text used for the task may have affected the current data. For instance, the number of comments on /θ/ was substantially larger in the reading data than in the narration data, and this difference may have to do with the words *things, thick,* and *three,* all of which appeared in the reading text. The larger number of comments on voiced /ð/ than on voiceless /θ/ can be explained by relatively high occurrence/use rates of *the, these/those,* and *this/that* in both reading and narration.

Other English fricatives were substituted with a sound chosen, probably unconsciously, by the L2 speakers. Labiodental /f/ and /v/ were often substituted with labial stops of the same voicing quality, /p/ and /b/, respectively, according to at least 39% of raters in this study. On the other hand, voiced alveolar /z/ was pronounced as its voiceless counterpart /s/, particularly when it was in a word-final position (Example 22) or used as a plural marker [z] (Example 23), as commented by 43.48% of the raters. Substitution of /z/ with a Korean-sounding affricate (which is closest to English /ʤ/) was also reported by 13.04% of the raters in nine comments (Example 24).

Example 22: Doesn't voice /s/ in "please" [E17_K07_R]
Example 23: "peas" pronounced with a final /s/ sound instead of /z/ [E22_K10_R]
Example 24: Line 8: "disaster" » "dijester" [E03_K09_N]

Mixed use of the liquid consonants /l/ and /r/ also caused the L2 speakers to sound foreign-accented, as perceived and reported by more than 80% of raters in this study. The issues of /l/-/r/ contrast are classic examples of Korean learners' pronunciation difficulties because of the absence of /r/ in Korean (e.g., Iverson & Sohn, 1994). Substitution of /l/ with /r/ occurred word-initially (Example 25), word-finally (Example 26) or in consonant clusters (e.g., *bl, fl, pl,* and *sl*; Example 27). This type of substitution (i.e., /l/ » /r/) was reported substantially less in the reading data, perhaps due to the low rate of /l/, especially in word-initial position, in the reading text. Substitution of /r/ with /l/ was detected in all positions across the reading and narration data (Example 28). In some samples, word-medial /r/ was perceived as /w/, according to 52.17% of the raters (Example 29). Pronunciation problems related to flap [ɾ] as in the word *water* in North American English were also reported. Some L2 speakers pronounced it as [r] (according to 47.83% of the raters in 28 comments; Example 30), while others pronounced it as [t] (according to 17.39% of the raters in five comments; Example 31).

Example 25: "lamp" with /r/ [E15_K18_N]
Example 26: Approximates an /r/ in the word "call" [E22_K30_R]
Example 27: "please" was said with an /r/ in line 1 [E23_K13_R]
Example 28: "red" as "led" [E17_K05_R]
Example 29: /r/ in "broken" sounds almost like /w/ [E08_K35_N]
Example 30: "ter" in "water" » "rer" [E04_K34_N]
Example 31: Doesn't 'flap' the /t/ in "water," aspirates it instead (she does say "later" correctly) [E08_K33_N]

Approximately half of the raters also pointed out that the L2 speakers tended to devoice word-final voiced obstruents /b, d, g, ʤ/ (Example 32). When a plural marker [z]

was added, the speakers often devoiced the entire coda consonant cluster (Example 33), but this issue was examined only in the reading data. Other issues with obstruents include excessive aspiration of word-final voiceless stops /p, t, k/ (Example 34) and frication of labial stops /p, b/ (Example 35). Overall, all of the 23 raters in this study appear to have attended to substitution of consonants while assessing L2 speech samples.

Example 32: "Bob" sounds like "bap" [E13_K18_R]
Example 33: Line 8: "bags" » "baeks" [E03_K21_R]
Example 34: Overly emphasizes each /t/ in "apartment" (aspiration) [E08_K04_N]
Example 35: "blue" sounds like "vlue" [E04_K05_R]

On the other hand, almost all of the raters in this study commented on the pronunciation issues that involve substitution of English vowels (see Table 4.3). Among the comments related to vowels, the majority of the noticed problems in both reading and narration pertained to mixed usage of English tense and lax vowels of the same height and backness, supporting previous findings on L1 Korean-L2 English speakers' use of vowels (e.g., Tsukada et al., 2005). That is, many Korean speakers of English show difficulty in discriminating tense and lax vowels (whereas NSs are sensitive to the difference between the two). More than 90% of the raters in this study commented on the mixed use of mid back vowels—particularly with respect to substitution of lax /ɔ/ with tense /o/ (206 comments; Example 36). The substitution of /o/ with /ɔ/, on the other hand, were pointed out by 43.48% of the raters (17 comments; Example 37). The largest number of comments concerned high front vowels (/i/ and /ɪ/): according to 17 (73.91%) of the raters, the L2 speakers in this study tended to substitute tense /i/ with lax /ɪ/ (276 comments; Examples 3 and 38) and vice versa (i.e., /ɪ/ » /i/); 255 comments; Example

39). Meanwhile, instances of substitution of tense /u/ with lax /ʊ/ and the opposite case (i.e., /ʊ/ » /u/) were also reported by a smaller number of raters (Examples 40 and 41).

Example 36: "call" sound like "col" [E11_K06_R]
Example 37: "hall" for "hole" in line 2 [E15_K40_N]
Example 38: "meet" as "mit" [E17_K08_R]
Example 39: Pronounces first vowel in "severe" as /i/ [E22_K16_N]
Example 40: using a slightly more lax vowel than /u/ in "scoop" [E08_K02_R]
Example 41: "today" » "too-day" [E03_K28_N]

Another prominent vowel issue was associated with the nontargetlike use of the neutral vowel [ə]. As pointed out by many raters (see Table 4.3 for detail), [ə] was substituted with a Korean-colored, mid-open back vowel [ʌ] (reported by 78.26% of the raters in 158 comments; Example 42), or for a closed back vowel [ɯ], which is neutral in the Korean phonology (reported by 30.43% of the raters in seven comments; Example 43). The raters also commented on the incidents in which [ə] was mispronounced as other vowels including [o], [ɑ], and [ɛ] (Example 44). Meanwhile, more than half of the raters reported that the low front vowel /æ/ was also problematic in the L2 samples. The L2 speakers tended to produce /ɛ/ for /æ/ (reported by 60.87% of the raters in 129 comments; Examples 4 and 45) and vice versa (reported by 47.83% of the raters in 31 comments; Example 46), indicating that the Korean speakers likely experience difficulty in producing a perceptually effective contrast between English /æ/ and /ɛ/.

Example 42: Line 2, "her" as "hur" (back) [E07_K03_R]
Example 43: "brokun" (for "broken") [E01_K04_N]
Example 44: Uses low front vowel instead of schwa at the end of "Stella" [E08_K25_R]
Example 45: "trash" » "tresh" [E09_K29_N]
Example 46: "red" sounds like "rad" [E22_K22_R]

Lastly, approximately 40% of the raters attended to the L2 speakers' use of the vowel /ɑ/ (see Table 4.3). They pointed out in 29 comments that the vowel /ɑ/ as in

bottle, problem, and *shocked* was substituted with /o/ (Example 47), an issue which may be associated with spelling pronunciation phenomena (Altenberg & Vago, 1983). In some words, particularly in the reading data, /ɑ/ was pronounced in a manner that is perceived to be somewhat Korean—slightly in the front (like /a/) (Example 48). The reported vowel substitution patterns indicate that some English vowels were not realized with acoustic accuracy in the English sound inventory of the L2 speakers in this study, and nontargetlike use of these vowels has influenced, to varying degrees, the raters' judgments.

Example 47: First vowel in "problem" is /o/ [E08_K32_N]
Example 48: Line 5: "Bob" » "bab" [E03_K10_R]

The two other subcategories of the segmental issues—insertion and deletion—are associated with English phonotactics and syllable structure. Insertion or deletion of a vowel, for instance, results in a change in the total number of syllables within a word, and possibly, re-syllabification (i.e., consonants attached to syllables other than those from which they originally come from). Re-syllabification also occurs because of insertion or deletion of a consonant (e.g., Labov, 1995). During accentedness rating in this study, all NS raters commented on a problem pertaining to insertion of a vowel—as a result, addition of a syllable—at least once. Table 4.4 presents a summary of the comments concerning insertion of a segmental.

Table 4.4

Summary of Commented Items (Segmental Insertion)

	Reading	Narration		Total	Example(s)
	Counts of comments	Counts of comments	Counts of comments	Counts (%) of commenting raters	
Consonants					
r insertion	55	54	109	12 (52.17%)	5, 55–57, 94–98, 100, 101
d insertion	0	6	6	6 (26.09%)	
Others	4	10	14	8 (34.78%)	
Consonants total	59	70	129	14 (60.87%)	
Vowels					
VC_$	260	116	376	22 (95.65%)	49
VCC_$	95	20	115	22 (95.65%)	6, 51
VC_C$	0	8	8	7 (30.43%)	50
V_C$	9	12	21	9 (39.13%)	
$C_CV	63	16	79	17 (73.91%)	52
$_CCV	0	3	3	2 (8.70%)	
Diphthongization	17	11	28	11 (47.83%)	53
Others	4	14	18	7 (30.43%)	54
Vowels total	448	200	648	23 (100%)	
Grand total	507	270	777	23 (100%)	

Deeper analyses of the written comments revealed that the L2 speakers tended to add a vowel after a syllable-final consonant (i.e., VC_$[2]; Example 49), even though one coda consonant is allowed in the Korean phonotactics. Almost all of the raters in this study (95.65% of them) touched upon this issue in a total of 376 comments throughout the dataset. Vowel epenthesis also occurred in English consonant clusters, which Korean phonotactics does not allow, but interestingly, the raters observed variations in the L2 samples. In the case of coda consonant clusters, an epenthetic vowel was used between the consonants (i.e., VC_C$; Example 50) or after the second consonant (i.e., VCC_$;

[2] $: syllable boundary (cf. #: word boundary); _ : position for epenthesized vowel

Example 51). The latter case was more noticeable than the former case, which was observed only in the narration data: reported by 95.65% of the raters in 115 comments vs. 30.43% of the raters in less than ten comments. On the other hand, in onset consonant clusters, the epenthetic vowel was used only between the two consonants (i.e., $C_CV; Example 52).

Example 49: "ovu" (Korean-like) for "of" [E15_K32_N]
Example 50: Line 1: "just" » "ju-sut" [E03_K06_N]
Example 51: vowel insertion after "ask" [E08_K13_R]
Example 52: "three" was said "thuree" [E23_K09_R]

The outcome of vowel epenthesis in the above contexts was manifested in two ways. On the one hand, the L2 speakers produced English vowels as suggested by the orthography, either using the orthographic conventions of English or of the transliteration systems of their L1. On the other hand, the inserted vowels in the current data were often in the form of the vowel [ɯ], which is neutral in Korean but does not exist in the phonemic inventory of English. The raters familiar with Korean (having learned it as a heritage or foreign language) were able to described this vowel as the neutral vowel in Korean. But other raters (including those who were not highly familiar with Korean) transcribed this vowel as <U> or <UH> in their comments or noted that the inserted vowel did not sound like English. As such, frequent use of such an unfamiliar vowel may have affected the raters' judgments to some extent.

Other instances of vowel insertion include separation of a diphthong into two separate syllables (Example 53), an issue which was reported by 48% of the raters in 28 comments, and insertion of an additional vowel before a coda /r/ (Example 54), a problem which was reported by 39.13% of the raters in 21 comments.

Example 53: “toy” » “to-ee” [E03_K22_R]
Example 54: “floor” sounds like extra syllable “er” at the end [E01_K09_N]

The raters also commented on insertion of a consonant (see Table 4.4). Approximately half of them described that some L2 speakers tended to insert [r] in some contexts—particularly before /l/ (Examples 55–56), or after a word-final vowel, thus changing the syllable structure with a new coda /r/ (Example 57).

Example 55: “also” was said as “urlso” [E05_K30_R]
Example 56: “call” was said as “curl” in line 1 [E23_K39_R]
Example 57: Says “stellar” instead of “Stella” [E22_K21_R]

Deletion of a segmental was also pointed out by the raters in assessing accentedness of the L2 speech samples (see Table 4.5). According to many of the raters, one of the noticeable patterns of consonant deletion was deletion of word final /z/ which was present as a plural marker (45 comments; Example 58) or as a word part (11 comments; Example 59). Other patterns of consonant deletion include: (a) deletion of /r/ that served as a single coda consonant or was in coda consonant clusters (reported by 65.22% of the raters in 54 comments; Example 60), (b) deletion of /t/ as a single coda consonant or in coda consonant clusters (reported by 43.48% of the raters in 65 comments), and (c) deletion of /d/ in coda consonant clusters (reported by 60.87% of the raters in 51 comments; Example 61); in some contexts, /d/ was a past tense marker (Example 62). Deletion of a stop consonant after a nasal coda (Examples 63–64) was reported as well. Finally, deletion of a vowel was noted by 34.78% of the raters in 11 comments, most of which pointed at deletion of the second vowel in the word *situation* (i.e., /ʊ/; Example 8).

Table 4.5

Summary of Commented Items (Segmental Deletion)

	Reading	Narration		Total	Example(s)
	Counts of comments	Counts of comments	Counts of comments	Counts (%) of commenting raters	
Consonants					
t deletion	4	61	65	10 (43.48%)	7, 63
z deletion	56	0	56	16 (69.57%)	58, 59
r deletion	25	29	54	15 (65.22%)	60
d deletion	27	24	51	16 (60/87%)	61, 62
l deletion	6	2	8	4 (17.39%)	
p deletion	0	8	8	7 (30/43%)	64
k deletion	2	4	6	4 (17.39%)	
n deletion	5	3	8	3 (13.04%)	
Others	2	13	15	9 (39.13%)	
Consonants total	127	144	271	20 (86.96%)	
Vowels					
ʊ deletion	0	11	11	8 (34.78%)	8
Vowels total	0	11	11	8 (34.78%)	
Grand total	127	155	282		

Example 58: line 7, “kids” – missing s/z [E14_K01_R]
Example 59: the “se” in “these” was cut off in line 2 [E23_K04_R]
Example 60: Deletes /r/ in “first” [E08_K01_N]
Example 61: Deletes final /d/ in “around” [E08_K11_N]
Example 62: “moved” is pronounced as “move” in line 1 [E02_K27_N]
Example 63: Deletes /t/ in “can’t,” which makes it indistinguishable from “can” [E08_K36_N]
Example 64: “lamps” » “lams” [E09_K20_N]

The second main category, prosodic issues, comprises four subcategories: lexical stress, phrasal stress, intonation, and speech fluency. Table 4.6 presents the frequency counts and proportions of these four subcategories.

Table 4.6

Frequency Counts and Proportions of Commented Items: Prosodic Issues

Subcategory	Reading	Narration	Total
Lexical stress	124 (4.61%)	304 (11.31%)	428 (15.92%)
Phrasal stress	758 (28.19%)	190 (7.07%)	948 (35.25%)
Intonation	701 (26.07%)	527 (19.64%)	1228 (45.67%)
Speech fluency	67 (2.49%)	18 (.67%)	85 (3.16%)
Total	1650 (61.36%)	1039 (38.64%)	2689 (100%)

The number of comments regarding lexical stress issues was relatively small (428 in total), compared to the other subcategories of prosodic issues. These comments were provided by 21 (91.30%) of the raters, indicating that most of the raters in this study attended to deviant use of English lexical stress while assessing L2 speech samples. The relatively small number may be due to a low number of multi-syllabic words used in the reading and narration datasets. In the reading text, for instance, there were only eight bi-syllabic words in which the L2 speakers could make lexical stress errors: *also, brother, into, maybe, plastic, station, Stella,* and *Wednesday*. All of the above words had a lexical stress issue in the L2 samples, while there were a few that received most attention from the NS raters in this study. About 65% of the raters commented on the incorrect stress placement in *Stella* (reported in 60 comments). Lexical stress issues in *plastic* (reported by 43.47% of the raters in 16 comments), *Wednesday* (reported by 30.43% of the raters in 13 comments), and *station* (reported by 26.08% of the raters in seven comments) were noticeable as well (see Examples 65–68).

Example 65: line 1 – large stress on /la/ in “Stella” [E07_K03_R]
Example 66: “PLAstic” pronounced as “plaSTIC” [E16_K29_R]
Example 67: “WednesDAY” in line 9 [E02_K32_R]
Example 68: “station” is pronounced “STATION” instead of “STAtion” [E04_K04_R]

On the other hand, the narration samples, collected using a spontaneous speaking task, had a greater variety of multi-syllabic words. Because of the context of the speaking prompt (i.e., describing problems in an apartment), there were certain words that had to be used to fulfill the task. The words that appeared highly frequently and at the same time, were found to be most problematic in the narration data include the following: *broken* (reported by 26.08% of the raters in 58 comments), *apartment* (reported by 21.73% of the raters in 27 comments), *cabinet* (reported by 34.78% of the raters in 17 comments), and *complain* (reported by 34.78% of the raters in 11 comments). Below are some examples of relevant comments, followed by Table 4.7, which presents an overview of selected words from the reading and narration samples:

Example 69: Too much stress on the second syllable of "broken" [E08_K31_N]
Example 70: "apartMENT" [E02_K27_N]
Example 71: "cabinet" – second and third syllables not reduced [E14_K05_N]
Example 72: "complain" was said with the stress on "COM" [E23_K09_N]

Table 4.7

Selected Words that Reportedly Had a Lexical Stress Issue

	Counts of comments	Counts (%) of commenting raters	Example(s)
Reading			
- *also*	2	2 (8.70%)	
- *brother*	5	3 (13.04%)	
- *plastic*	16	10 (43.47%)	66
- *station*	7	6 (26.08%)	68
- *Stella*	60	15 (65.21%)	65
- *Wednesday*	13	7 (30.43%)	67
Narration			
- *apartment*	27	5 (21.73%)	70, 73
- *broken*	58	9 (26.08%)	69
- *cabinet*	17	8 (34.78%)	71
- *complain*	11	8 (34.78%)	72
- *problem(s)*	11	4 (17.39%)	
- *window(s)*	12	2 (8.70%)	

In the current analysis (see Table 4.8), lexical stress problems were characterized as: (i) misplacement of primary stress (reported by 91.30% of the raters in 352 comments; Examples 65–67, 69–70, and 72) and (ii) no reduction of unstressed vowels—i.e., placement of equal stress on all syllables within the words (reported by 65.22% of the raters in 67 comments; Examples 68 and 71). In a smaller number of comments, the raters only pointed to problematic words but did not provide any description of the problems (Example 73) or they did not specify the locus of lexical stress errors (Example 74). These comments were thus coded as "unspecified" as shown in Table 4.8.

Example 73: lexical stress in "apartment" is off [E17_K04_N]
Example 74: wrong syllable stress in line 3 [E12_K03_R]

The raters provided other comments that fell into the subcategory "phrasal stress," or rhythm, one of the suprasegmental characteristics of a foreign accent documented in the research literature (e.g., Moyer, 2015). The raters in this study overall provided 948

comments on the issues related to phrasal stress: 758 in the reading data and 190 in the narration data (see Table 4.9 for a summary). The difference between the two datasets seems to suggest that the raters detected fewer phrasal stress issues in the spontaneous speech samples.

Table 4.8

Summary of Commented Items (Lexical Stress)

	Reading	Narration	Total		Example(s)
	Counts of comments	Counts of comments	Counts of comments	Counts (%) of commenting raters	
Misplacement of primary stress	107	245	352	21 (91.30%)	65, 66, 67, 69, 70, 72
No reduction	10	57	67	15 (65.22%)	68, 71
Unspecified	7	2	9	4 (17.39%)	73, 74
Total	124	304	428	21 (91.30%)	

Table 4.9

Summary of Commented Items (Phrasal Stress)

	Reading	Narration	Total		Example(s)
	Counts of comments	Counts of comments	Counts of comments	Counts (%) of commenting raters	
Locus of nontargetlike rhythm:					
- Pronouns	177	37	214	16 (69.57%)	75
- Prepositions	41	18	59	11 (47.83%)	76
- Modal verbs	82	3	85	17 (73.91%)	10, 77
- Articles	30	1	31	11 (47.83%)	78
- Verb *is*	0	6	6	4 (17.39%)	
- Noun phrases or compound nouns	273	49	322	19 (82.61%)	79, 80, 81
Others:					
- Linking	31	0	31	12 (52.17%)	82
- No rhythm	13	2	15	8 (34.78%)	83
Unspecified	111	74	185	17 (73.91%)	84
Total	758	190	948	23 (100%)	

The raters commented on phrasal stress placed on the words that should remain unstressed, including pronouns (reported by 69.57% of the raters in 214 comments; Example 75), prepositions (reported by 47.83% of the raters in 59 comments; Example 76), modal verbs such as *will* and *can* (reported by 73.91% of the raters in 85 comments; Examples 10 and 77), and articles such as *the* and *a* (reported by 47.83% of the raters in 31 comments; Example 78). Although phrasal stress was placed on content words such as nouns and verbs, there were a number of phrases where the rhythm was perceived to be unnatural, according to more than 80% of the raters (Example 79). Performance on compound nouns was also problematic, according to the comments (Examples 80–81), and revealed the L2 speakers' lack of knowledge of targetlike compound prosody. In other comments, the raters pointed out problems in linking (Example 82) and absence of rhythm (i.e., every word was equally stressed throughout the sentence; Example 83). In addition, there were 185 comments in which the raters (73.91%) did not specify the locus of the rhythm problem (Example 84).

Example 75: stress on "her" in line 2 [E02_K09_R]
Example 76: "FOR the kids" [E15_K22_R]
Example 77: "we WILL go," wrong phrasal stress [E03_K03_R]
Example 78: "from THE store" [E20_K36_R]
Example 79: "these THINGS" rather than "THESE things" [E16_K38_R]
Example 80: "snow PEAS" instead of "SNOW peas" [E22_K25_R]
Example 81: "train STAtion," rather than "TRAIN station" [E13_K01_R]
Example 82: "meet her," don't blend [E21_K06_R]
Example 83: There was no rhythm at all [E11_K13_R]
Example 84: wrong phrasal stress in line 8 [E04_K40_N]

The third subcategory of prosodic issues is related to intonation (1,228 comments in total). According to the rater comments, the L2 speakers often produced a High (H-

H%) or rising (LH%) boundary tone at the end of declarative and imperative sentences (reported by 82.61% of the raters in 335 comments; Examples 85–86), while these tones are normally in the intonation of questions or utterances of uncertainty, approval, or continuation in English (Huang & Jun, 2011). Some speakers sounded flat and monotonal throughout the utterance (reported by 56.52% of the raters in 283 comments; Example 87) or they had one or more than one nontargetlike pitch peaks (e.g., steep rising LH sequence) within a clause/sentence (reported by 52.17% of the raters in 94 comments; Example 88). Some speakers used falling intonation at the end of yes-no questions (reported by 8.70% of the raters in three comments; Example 89). Also, there were more than 500 comments in which the raters did not specify the type of the intonation problem but only left simple notes (Examples 90–91). About 65% of the raters provided such "unspecified" comments at least once. A summary is presented in Table 4.10.

Example 85: "Please call Stella" rising at end [E20_K33_R]
Example 86: Rising intonation in "from the store," like she is asking a question [E04_K01_R]
Example 87: Very flat intonation [E19_K35_N]
Example 88: Intonation was rising and falling and then rising again in "bad" at the end of line 2 [E23_K07_N]
Example 89: Line 10, should have rising intonation with question form, but is falling [E01_K04_N]
Example 90: unnatural intonation in line 2 [E04_K16_N]
Example 91: inaccuracy in intonation, line 4 [E10_K04_R]

Table 4.10

Summary of Commented Items (Intonation)

	Reading	Narration	Total		Example(s)
	Counts of comments	Counts of comments	Counts of comments	Counts (%) of commenting raters	
Statement ends with a rising tone	196	139	335	19 (82.61%)	85, 86
Flat overall	164	119	283	13 (56.52%)	87
Unnatural peak(s)	85	9	94	12 (52.17%)	88
Question ends with a falling tone	0	3	3	2 (8.70%)	89
Unspecified	256	257	513	15 (65.22%)	90, 91
Total	701	527	1228	23 (100%)	

The last subcategory of prosodic issues to be examined in this study, fluency issues, was not included in the rating guideline provided to the raters at the time of data collection (see Table 3.7). However, it was revealed during the data coding cycles that some of the raters had commented on the fluency features (e.g., speed and pauses) of the L2 samples; these comments were left in the slots for free notes included in the rating forms. Because previous studies (e.g., Munro & Derwing, 2001; Pickering & Baker, 2014; Sereno et al., 2016; Trofimovich & Baker, 2006) have considered these features to be essential aspects of speech prosody in making accentedness judgments, the current category, speech fluency, was included in the analyses. Two aspects of speech fluency were reported: speech rate and pause (see Table 4.11).

Table 4.11

Summary of Commented Items (Speech Fluency)

	Reading	Narration	Total		Example(s)
	Counts of comments	Counts of comments	Counts of comments	Counts (%) of commenting raters	
Pause(s)	44	8	52	13 (56.62%)	13, 92
Speech rate (fast)	23	10	33	9 (39.13%)	12, 93
Total	67	18	85	16 (69.57%)	

As shown in Table 4.11, almost 70% of the raters provided a total of 85 comments to address fluency-related issues by which (they thought) their accentedness judgments were affected. Specifically, nearly 57% of the raters pointed out that there were unnatural pauses that broke the overall flow of the speech (Examples 13 and 92). But it was not clear from the current data whether these unnatural pauses were filled (e.g., "uh") or silent. Approximately 39% of the raters were sensitive to speech rate. These raters reported that some L2 speakers spoke too fast to allow the listeners to identify the overall flow of the speech (Examples 12 and 93).

Example 92: "snack …… for her brother" [E15_K35_R]
Example 93: Generally speaks too fast and many words blur together [E22_K34_N]

Results shown thus far can be summarized as follows: examination of the comments revealed that the raters attended to diverse pronunciation segmental and prosodic deviations while assessing Korean-accented English speech samples, and two major types (seven subcategories in total) were described in detail. Most of the raters left comments (more than half of the total comments collected in this study) that addressed issues regarding segmental deviations. A variety of segmental substitution, insertion, and deletion patterns were considered to be relevant in characterizing a foreign accent.

Meanwhile, many raters noted that problems associated with prosody—lexical stress, phrasal stress, intonation, and fluency of the utterance—also influenced their assessment. Among the prosodic issues, intonation was pointed out the most in the raters' comments.

4.1.2. Accentedness rating scores

Analyses of the scalar assessments of accentedness followed. First, in order to examine inter-rater reliability and determine whether it is statistically appropriate to combine the raters' accentedness ratings, the internal consistency of the scale items was computed. As shown in Table 4.12, the raters' consistency (Cronbach's alpha) for each of the five accentedness measure items was acceptable, indicating that the 23 raters showed consistency across all five items in the rating scales for both reading and narration. Furthermore, correlations among the five items were also computed, and the coefficients were overall consistent and large, with Pearson *r* values of .64 and above, all at the $p < .01$ level in both reading and narration ratings (see Table 4.13). Accordingly, following the methods in previous work in this area (e.g., Huang & Jun, 2015; Kang, 2010), ratings on the five accent rating items were summed and then averaged into a single scale measure, which was used as a composite measure for subsequent analyses.

Table 4.12

Inter-rater Agreement (Cronbach's α) for Accentedness Rating Items

Rating Item	Reading	Narration
A. Speaks with a North American accent … Speaks with a foreign accent	.98	.98
B. Has no accent … Has a strong accent	.98	.98
C. Speaks with an English manner of pronunciation … Speaks with no English manner of pronunciation	.97	.98
D. Has a familiar English accent … Has an unfamiliar English accent	.97	.98
E. Speaks like a native speaker of English … Speaks like a nonnative speaker of English	.98	.98

Notes: In all reliability statistics, number of cases: 40; number of items: 23; $p < .05$

Table 4.13

Inter-item Correlation Matrix

		Reading					Narration				
		A	B	C	D	E	A	B	C	D	E
R	A	-	.93	.83	.84	.88	.71	.73	.67	.70	.70
	B		-	.86	.84	.91	.71	.74	.68	.69	.72
	C			-	.88	.82	.66	.67	.71	.70	.66
	D				-	.81	.68	.68	.68	.74	.65
	E					-	.67	.70	.65	.64	.73
N	A						-	.92	.85	.87	.87
	B							-	.87	.88	.90
	C								-	.89	.87
	D									-	.84
	E										-

Note: All at $p < .01$

As summarized in Table 4.14, for each of the L2 speakers, the average value of the scores given by the 23 raters was calculated. The possible score range was from 5 (no foreign accent; almost nativelike) to 45 (heavy foreign accent). From the reading task, the L2 speakers scored 25.52 on average (SD = 8.61), and the scores ranged from 7.48 to

38.87. The mean accentedness score from the narration data was 25.33 (*SD* = 8.72), and the scores ranged from 7.43 to 37.00. This indicates that the L2 speakers in the current study exhibited little to substantial degrees of foreign accent in their speech. The skewness statistics revealed that the accentedness scores from both tasks were not in normal distribution ($p = .038$ for reading and $p = .004$ for narration, as assessed by a Shapiro-Wilk test of normality). Rather, the score distributions were somewhat negatively skewed, as presented in Figures 4.2 and 4.3. Therefore, a Wilcoxon Signed Rank test, a nonparametric statistical test, was conducted to compare the accentedness scores from the two tasks. The difference between the reading and narration scores was not statistically significant ($Z = -.571$, $p > .05$). This result suggests that there was no effect of the task type (or the speech elicitation method) on accentedness, at least for the L2 speakers and the raters in the current study.

Table 4.14

Descriptive Statistics of Individual L2 Speakers' Accentedness Ratings

	Reading					Narration				
Speaker	Mean	SD	Min	Max	Mdn	Mean	SD	Min	Max	Mdn
K01	36.52	5.09	26	45	37	36.91	5.18	29	45	38
K02	11.91	3.53	5	20	12	15.09	6.68	5	30	15
K03	32.70	5.35	15	40	33	33.04	5.08	24	41	35
K04	31.96	6.56	16	43	32	32.04	5.89	15	43	33
K05	33.26	5.57	23	44	32	34.57	5.26	25	43	35
K06	25.57	7.58	13	38	25	33.09	8.04	17	44	34
K07	17.17	7.77	7	38	15	18.04	6.49	9	27	15
K08	25.83	6.39	15	35	27	20.87	6.83	10	35	21
K09	27.83	7.57	15	40	29	28.09	6.80	15	38	30
K10	35.39	5.38	23	45	36	31.17	5.43	20	41	32
K11	16.35	7.52	7	32	13	20.17	7.36	8	35	20
K12	7.48	4.66	5	26	6	7.74	3.97	5	20	5
K13	38.87	4.32	32	45	38	37.00	7.04	20	45	40
K14	17.78	6.35	5	29	17	16.09	7.00	8	30	13
K15	17.87	7.71	5	35	15	19.78	8.19	10	38	18
K16	24.39	7.80	5	37	26	31.09	5.20	20	40	30
K17	27.87	5.39	18	39	28	29.78	6.24	15	40	29
K18	27.13	6.74	16	38	29	31.39	7.72	15	42	34
K19	28.52	7.58	11	43	27	23.48	7.80	10	40	24
K20	24.39	7.20	10	40	24	21.39	7.45	10	40	22
K21	35.35	4.65	24	43	35	36.48	4.53	29	42	37
K22	25.57	7.10	12	35	28	28.30	7.05	15	40	25
K23	9.26	4.99	5	21	7	7.43	3.04	5	15	6
K24	8.87	5.40	5	24	6	7.96	3.94	5	20	6
K25	14.35	7.16	5	30	11	11.04	6.75	5	30	8
K26	14.00	7.80	5	26	12	9.74	4.97	5	20	7
K27	16.96	7.62	6	32	16	17.78	7.59	8	32	16
K28	19.74	7.50	9	32	18	18.17	7.38	6	35	17
K29	34.43	5.76	18	45	35	30.00	7.62	15	40	32
K30	34.52	5.23	22	43	35	32.57	7.35	15	42	34
K31	32.39	6.30	20	43	33	29.96	6.35	19	40	29
K32	35.35	6.44	23	45	37	32.48	7.19	17	43	32
K33	31.09	8.09	16	44	30	32.78	5.67	15	40	33
K34	31.17	8.31	15	42	33	31.61	6.57	20	45	30
K35	30.43	7.55	17	43	31	26.96	8.51	13	42	27
K36	21.52	7.15	10	37	20	25.04	7.11	11	35	24
K37	34.96	6.15	25	45	35	34.00	5.85	15	41	36
K38	23.13	7.65	15	40	22	24.48	8.60	10	40	25
K39	30.39	6.47	17	42	31	31.30	8.02	10	41	32
K40	28.35	8.79	14	40	29	24.26	7.51	10	36	25
Total	25.52	8.61	5	45	28	25.33	8.72	5	45	28

Notes: SD: standard deviation; Min: minimum; Max: maximum; Mdn: median

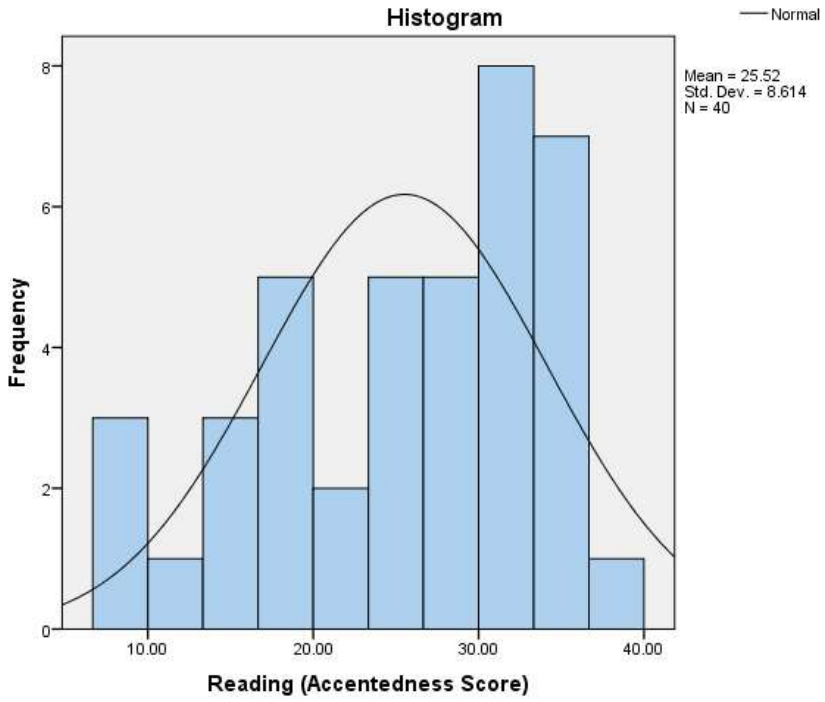


Figure 4.2 Distribution of Accentedness Scores on Reading

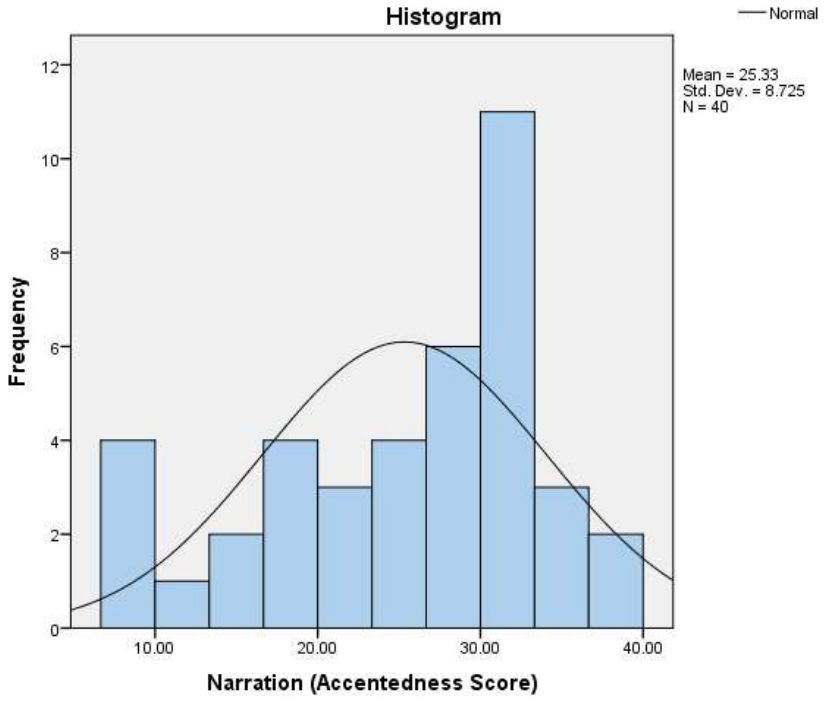


Figure 4.3 Distribution of Accentedness Scores on Narration

A Kruskal-Wallis H test was carried out in order to examine the effect of the amount of L2 experience on rated accentedness (see Table 4.15). There was a statistically significant difference in reading scores among the different L2 speaker groups: $\chi^2(2) = 18.72$, $p < .001$, with a mean rank readings core of 27.46 for Group 1, 26.40 for Group 2, and 10.72 for Group 3. Follow-up tests were run to evaluate pairwise differences among the three groups, controlling for Type I error across tests by using the Bonferroni approach. Group 3 was significantly different from Group 1 (Adjusted $p < .001$) and Group 2 (Adjusted $p < .01$), but the difference between Groups 1 and 2 was not significant. Similar results were obtained from the narration data: a significant effect of L2 experience on the narration scores was found ($\chi^2(2) = 16.97$, $p < .001$). A post-hoc test with Bonferroni correction showed significant differences between Groups 1 and 3 (Adjusted $p < .01$) and between Groups 2 and 3 (Adjusted $p < .01$). It can be thus interpreted that significant improvement in accentedness does not manifest itself until some amount of L2 experience is accumulated.

Table 4.15

Descriptive Statistics of Accentedness Ratings across Groups

Accentedness Ratings		Group 1	Group 2	Group 3
	N	14	10	16
Reading	Mean	30.75	30.39	17.89
	SD	5.29	4.08	7.43
	Min	21.52	23.13	7.48
	Max	38.87	36.52	33.26
	Mean Rank	27.46	26.40	10.72
Narration	Mean	29.91	30.77	17.92
	SD	5.15	3.54	8.34
	Min	20.87	24.48	7.43
	Max	37.00	36.91	34.57
	Mean Rank	26.29	27.30	11.19

Notes: SD: standard deviation, Min: minimum, Max: maximum

Results shown thus far in this chapter provide a general picture of the characterization and perception of a foreign accent. The 23 raters in this study showed high consistency in assigning accentedness scores to the L2 speech samples. In addition to assigning numeric accentedness scores, the raters provided their own description of the segmental and prosodic issues detected in Korean-accented English speech. However, as previously noted in this chapter, the raters in this study also showed varied degrees of sensitivity to each of the phonological deviations that affected their accent judgment. It was therefore necessary to conduct a deeper investigation of the performance of individual speaker and rater participants to examine more closely what underlies the accentedness scores and judgments. The findings of these "focused" analyses are presented in Sections 4.1.3 and 4.1.4.

4.1.3. Focused analyses 1: Two raters' assessment of an L2 speaker

This subsection reports qualitative analyses of how two individual raters assessed the accentedness of an L2 speaker. Two raters and one L2 speaker were selected: E02, E05, and K16. The two raters were selected purposefully because their background characteristics lie at two extreme ends, as presented in Table 4.16. According to his responses in the background questionnaire, Rater E02 grew up in Salt Lake City, Utah, and currently lives there working as an English instructor. He was highly experienced in teaching ESL/EFL (approximately 20 years in total) and also English pronunciation/speaking skills. In the questionnaire, he responded that he had been in contact with a Korean speaker regularly—three to five times a week. In contrast, Rater E05 was a resident of central New Jersey his entire life. His experiences in teaching and

familiarity with Korean were somewhat limited: he has taught ESL for one year and taken one ESL teaching methodology course in which some contents of phonology, phonetics, and phonics were included. He responded in the questionnaire that he had not been in contact with a Korean speaker. Therefore, these two raters were considered to represent the two extreme ends of the spectra of experience with respect to teaching, formal training in phonology, and familiarity with Korean. Both raters completed rating of the reading samples first, then assessed the narration samples.

Table 4.16

Comparison of the Selected NS Raters (E02 & E05)

Background characteristic	E02	E05
Gender	Male	Male
Hometown	Salt Lake City, Utah	Central New Jersey
Teaching ESL	17 years	1 year
Teaching EFL	3.5 years	None
Training in phonology	Learned and taught	Learned
Contact with Korean	3-5 times a week	Never
Other languages (proficiency)	Spanish (low), Japanese (low)	Italian (fluent), Spanish (low)

The following is a brief introduction of the selected L2 speaker (see Table 4.17 for summary): Speaker K16, age 31, had been in the United States for approximately one year at the time of data collection. He had been preparing for applications for graduate school admission, but was not taking any ESL classes. He reported that he had scored 900 out of 990 on the Test of English for International Communications (TOEIC)[3], a very common English proficiency test for learners in Korea. His first exposure to the English

[3] The TOEIC Test here refers to the one that consists of the listening comprehension test and the reading comprehension test. The test is developed by Educational Testing Service.

language was in his middle school years back in Korea. This speaker was selected for the current analyses because his total accentedness score ranked 20th, the median of the forty L1 Korean-L2 speakers in this study. He was thus assumed to be the average, in terms of accentedness scores, among the speaker participants. The following reports on how this speaker's performance was perceived and assessed by the two selected raters.

Table 4.17

Summary of Backgrounds and Performance of the Selected L2 Speaker (K16)

Speaker	K16
Gender	Male
Age	31
L2 experience in years	1
L2 experience group	1
Accentedness Ratings:	
Reading	
Mean (SD)	24.39 (7.80)
Min-Max	5-37
Mdn	26
Narration	
Mean (SD)	31.09 (5.20)
Min-Max	20-40
Mdn	30
Total accentedness score rank (out of 40)	20

Notes: SD: standard deviation; Min: minimum; Max: maximum; Mdn: median

Ratings of E02. Tables 4.18 and 4.19 present the commentary that Rater E02 provided while assessing Speaker K16's reading and narration, respectively. In his assessment of reading, E02 commented on the segmental issues by (i) transcribing the problematic sound(s) as in line 3 (Table 4.18), (ii) providing his own interpretation of what the speaker may have intended to say (as in line 4, Table 4.18), or (iii) simply writing that a particular word was mispronounced, without specifying the mispronounced

segmental(s) (e.g., line 8, Table 4.18). In addition, Rater E02 was very sensitive to the prosodic issues in K16's reading. He wrote that the intonation was overall flat (lines 1, 2, 3, 4, and 6), but no other intonational issues were pointed out. He provided quite analytical comments about awkward phrasal stress. According to his comments, K16 placed excessive stress to the pronoun *her* throughout the speech, modal verbs *can* and *will*, and final nouns in noun phrases or compounds (e.g., *snow PEAS* in line 3, *toy FROG* in line 7, and *train STATION* in line 9). The rater did not comment on any fluency issues in this speech. To this speech, Rater E02 assigned accentedness scores around the middle point of the scale, as presented in Figure 4.4.

Table 4.18

E02's Commentary on K16's Reading

Line	Transcription	Segmentals	Prosody
1	Please call Stella.		- flat intonation
2	Ask her to bring these things with her from the store.		- flat intonation - "her" is stressed, twice
3	Six spoons of fresh snow peas,	- "freSHA"	- flat intonation - "peas" is stressed
4	five thick slabs of blue cheese,	- "of" is mispronounced or the reader tried to add a vowel after it	- flat intonation
5	and maybe a snack for her brother Bob.		- "her" is stressed
6	We also need a small plastic snake		- flat intonation
7	and a big toy frog for the kids.		- "frog" is oddly stressed
8	She can scoop these things into three red bags,	- "red" is mispronounced	- "can" is stressed
9	and we will go meet her Wednesday at the train station.		- "will" is stressed - "her" is stressed - "station" is stressed

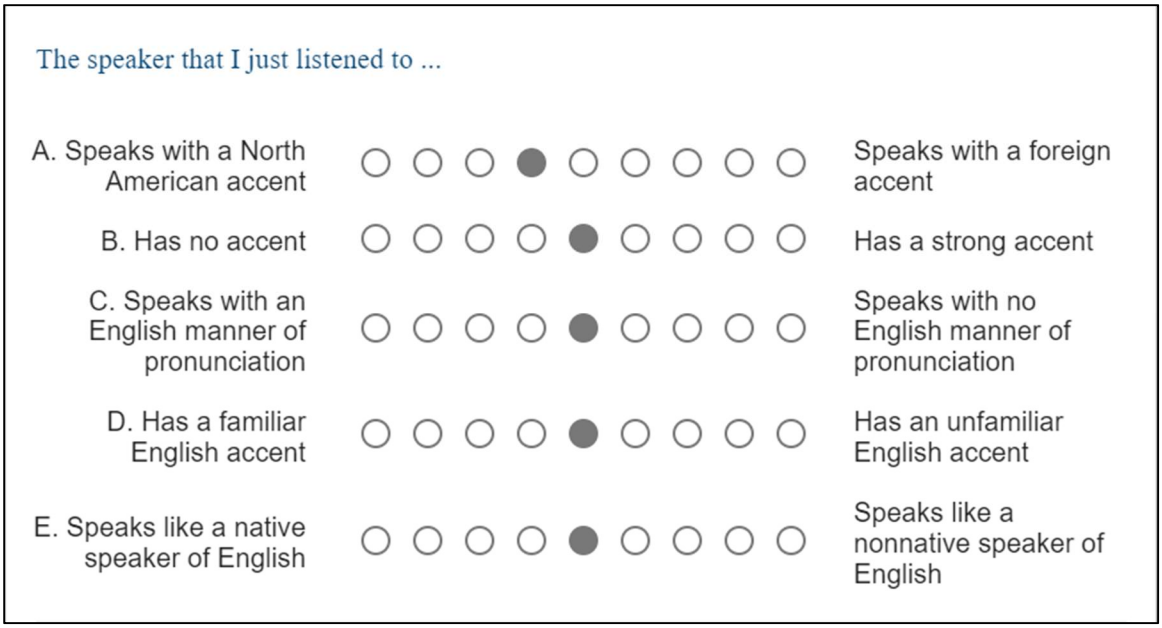


Figure 4.4 E02's Accentedness Ratings of K16's Reading

Similar commenting behaviors were observed in E02's assessment of K16's narration (see Table 4.19). Rater E02 left quite several comments concerning the rather flat intonation of the speaker. It should be noted, however, that "flat intonation" was not the only description that E02 provided to the L2 speakers in this study. In his commentary, other speakers' intonation was described as, for instance, "flat to rising intonation at the end of line 3" [E02_K11_R], a comment which indicates that he was attending to the overall pitch movement in each speech sample. Speaker K16 appears to have sounded monotonous to a few other raters as well (e.g., "flat overall" [E15_K16_N], "neutral intonation" [E23_K16_R]). Meanwhile, E02 also thought that K16 produced a nontargetlike rhythm by stressing unfocused words—preposition *in* (line 2) and pronoun *me* (line 6). These findings, along with similar comments provided in the reading sample, suggest that K16's use of L2 English intonation and rhythm was nontargetlike regardless

of the speech type and that Rater E02's perception of the speaker's prosodic issues was consistent across the samples.

Table 4.19

E02's Commentary on K16's Narration

Line	Transcription	Segmentals	Prosody
1	I'm tenant living in the apartment that you rented yesterday.	- "apartment" was mispronounced - "rented" is mispronounced	- flat intonation
2	Unfortunately, there are so many problems in the apartment.		- "in" is stressed
3	The most important, I mean, the most severe thing is that there is a rat.		- flat intonation
4	Please handle this problem.	- "please" is mispronounced	- flat intonation
5	And then the second thing is, um, the window is broken.	- "thing" is mispronounced	- flat intonation
6	It's very hard to handle for me.		- "me" is stressed - flat intonation
7	So please handle these problems.	- "these" is mispronounced	
8	And furthermore, the sink is leaking.	- "furthermore" is mispronounced - "leaking" is mispronounced	- flat intonation

On the other hand, E02's comments on the segmental issues were not as specific as in the assessment of reading (see Table 4.18 for comment examples). He merely described that certain words were "mispronounced," without providing phonological/phonetic details. But such a comment was not commonly found in his assessment of other speakers (e.g., "peas" is pronounced as "peace" in line 3 [E02_K13_R]; the /k/ sound is not aspirated fully in line 3 [E02_K04_R]). It seems that E02 chose to type in such simple, later coded as "unspecified," comments in some of the

samples, and K16's narration happened to be one of them. Lastly, E02 rated that K16's narration was slightly more foreign-accented than his reading, as shown in Figure 4.5.

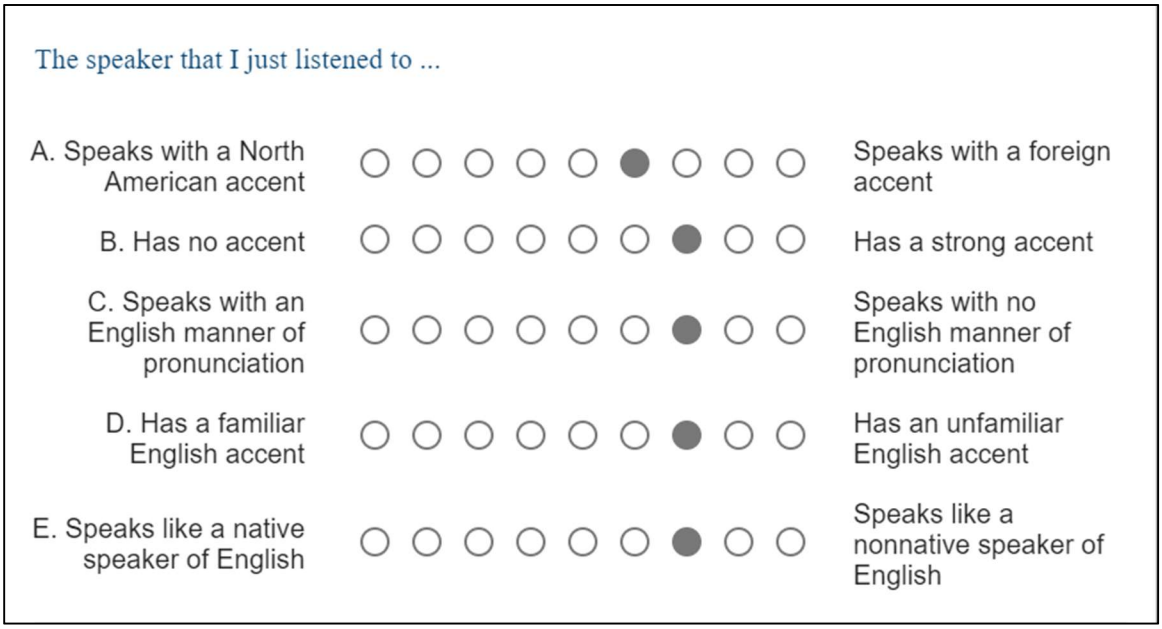


Figure 4.5 E02's Accentedness Ratings of K16's Narration

Ratings of E05. While assessing Speaker K16's speech for accentedness, Rater E05 attended to several segmental and prosodic features and provided relevant comments (see Tables 4.20 and 4.21). Very little metalanguage appeared in his commentary. In describing segmental issues in the reading sample, he either transcribed pronunciation of the mispronounced word (e.g., line 1: "call" was said as "col") or provided his own analysis of the issue (e.g., line 8: "these" was said with a /d/). It is notable that just as E02 did, E05 commented on the same segmental issues in the words *fresh* in line 3 and *of* in line 4. Although E05 did not comment on the word *red* in line 8, which E02 attended to during rating, he showed sensitivity to a few other segmental issues in K16's reading, as

presented in Table 4.20. On the other hand, E05 left fewer comments concerning prosodic issues in K16's reading. He wrote that one of the lines (line 1) ended with a "neutral" tone instead of a falling one (agreeing with E02) and that there was a rhythm issue in line 3; however, no detailed description of the rhythm issue was provided in his comment. To this speech, Rater E05 assigned accentedness scores around the middle point of the scale, as presented in Figure 4.6.

Table 4.20

E05's Commentary on K16's Reading

Line	Transcription	Segmentals	Prosody
1	Please call Stella.	- "call" was said as "col"	- intonation at the end of line 1 was neutral
2	Ask her to bring these things with her from the store.	- "her" was said as "har" - "these" was said as "dese"	
3	Six spoons of fresh snow peas,	- "fresh" had a /u/ at the end of it	- rhythm in "six spoons" was off…
4	five thick slabs of blue cheese,	- "thick" was said with a /d/ sound - "of" was said like "ofa"	
5	and maybe a snack for her brother Bob.		
6	We also need a small plastic snake		
7	and a big toy frog for the kids.	- "frog" was said with the wrong vowel sound - "kids" was said with the wrong /i/	
8	She can scoop these things into three red bags,	- "these" was said with a /d/	
9	and we will go meet her Wednesday at the train station.		

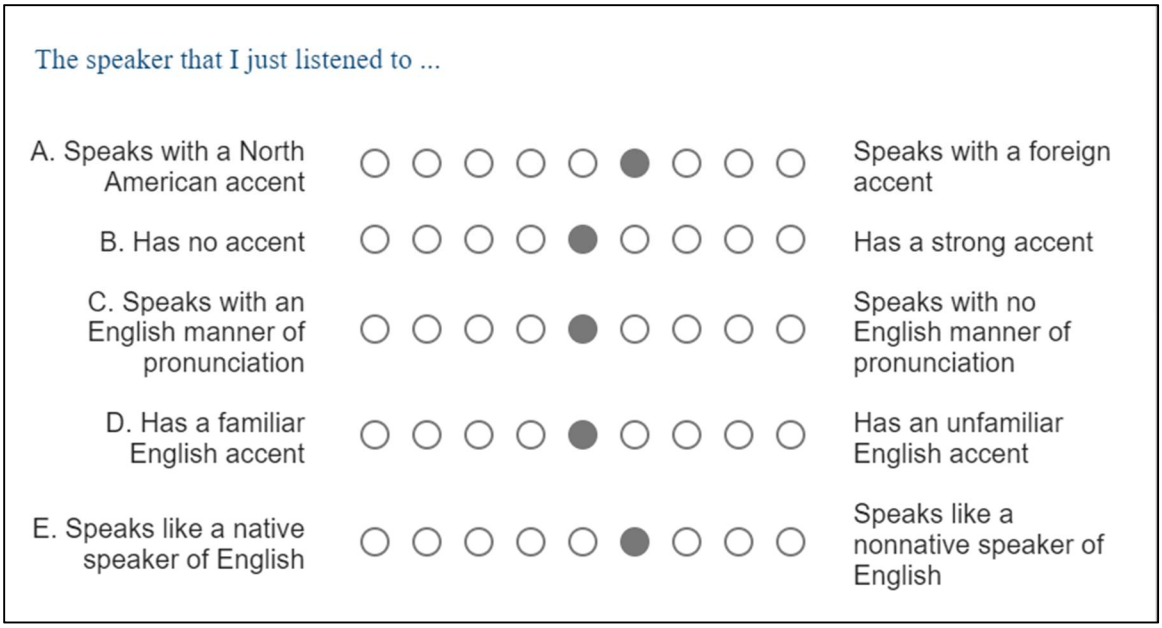


Figure 4.6 E05's Accentedness Ratings of K16's Reading

In the assessment of narration (see Table 4.21), most of E05's comments were about segmental issues which he described by transcribing incorrectly pronounced words (e.g., line 8: "leaking" was said as "licking"). These descriptions were relatively more specific than the ones provided by E02 that were reported previously. Also, similarly to the rating results from the reading sample, E05 agreed with E02 regarding the segmental issues perceived in the words *apartment* and *rented* in line 1, *please* in line 4, *these* in line 7, and *furthermore* and *leaking* in line 8. On the other hand, his comments on nonsegmental issues were somewhat limited. He left only one comment regarding the rhythm in line 2, which, again, he described as "off." Rater E05 assigned poorer accentedness scores to this speech (clustered near the higher end of the rating scales, as presented in Figure 4.7), as compared to the scores that he assigned to the reading of the

same speaker. In other words, K16's narration sounded more foreign-accented to E05 (than did the reading of the same speaker).

Table 4.21

E05's Commentary on K16's Narration

Line	Transcription	Segmentals	Prosody
1	I'm tenant living in the apartment that you rented yesterday.	- "apartment" sounded like "uppertment" - "rented" was said as "lento"	
2	Unfortunately, there are so many problems in the apartment.	- "unfortunately" was missing a syllable. I think the "a" was not pronounced.	- rhythm in "so many" was off
3	The most important, I mean, the most severe thing is that there is a rat.	- "severe" was said as "seever(e)"	
4	Please handle this problem.	- "please" was said "prease"	
5	And then the second thing is, um, the window is broken.		
6	It's very hard to handle for me.		
7	So please handle these problems.	- "these" was said as "this"	
8	And furthermore, the sink is leaking.	- "furthermore" was said like "foothermore" - "leaking" was said as "licking"	

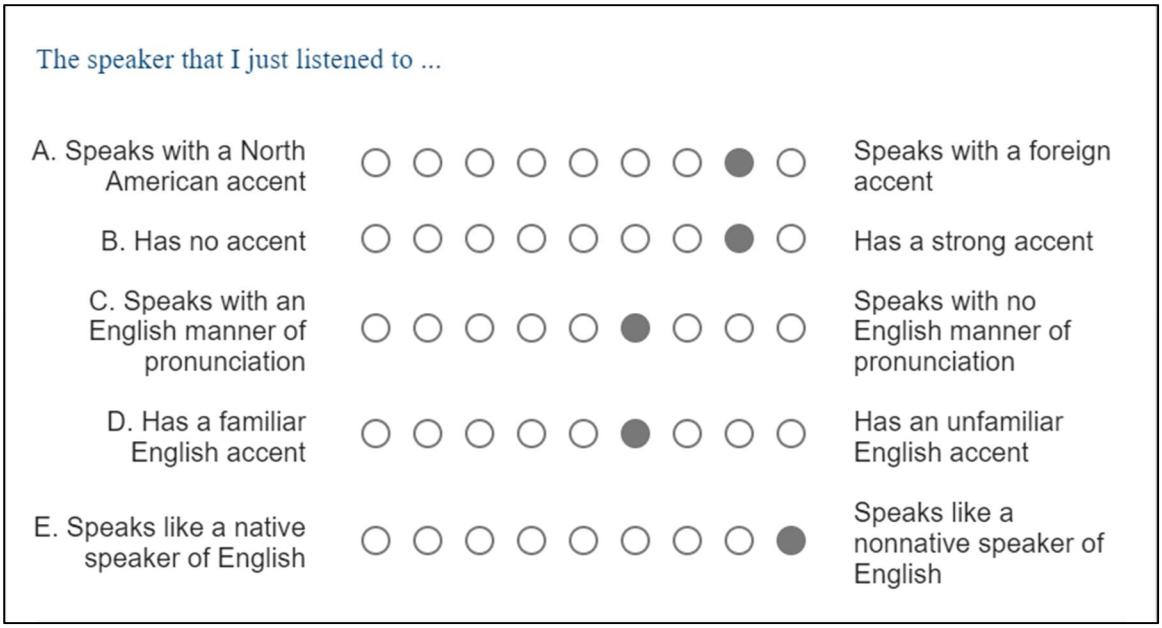
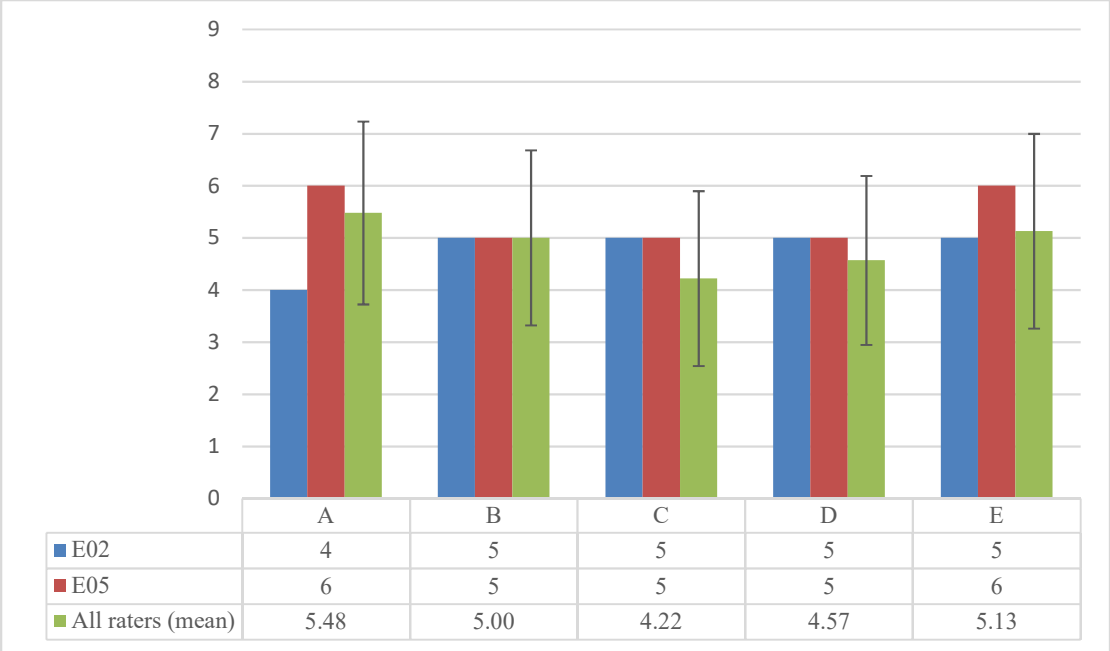


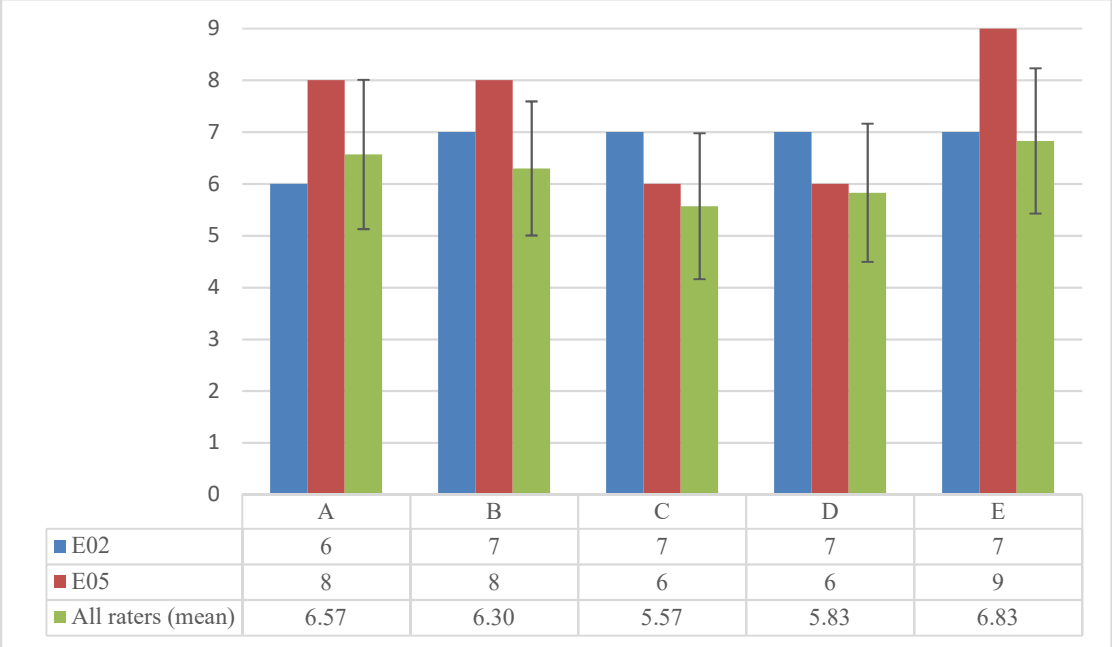
Figure 4.7 E05's Accentedness Ratings of K16's Narration

In the current analyses of the rating data of E02 and E05 on K16, the narration samples were consistently judged to sound more foreign-accented (see Figures 4.8 and 4.9). However, in the ratings of the entire speaker samples, E02 (and most of the other raters in this study) assigned higher accentedness scores to the reading samples than to the narration samples on average, while E05 provided an opposite judgment. It is also intriguing that in the scalar assessments of the selected samples (i.e., K16's reading and narration), E05 was overall slightly more severe than was E02. The reason for such a difference in rater severity is not known from the current analyses, although speculation can be made that it is related to E05's experience in ESL/EFL teaching, linguistics, or familiarity with Korean, or a combination of all of these factors. Or from a certain point during rating, there may have been a sort of practice or familiarity effect playing a role in his rating of the L2 samples.



	A	B	C	D	E
E02	4	5	5	5	5
E05	6	5	5	5	6
All raters (mean)	5.48	5.00	4.22	4.57	5.13

Figure 4.8 K16's Accentedness Scores on Reading



	A	B	C	D	E
E02	6	7	7	7	7
E05	8	8	6	6	9
All raters (mean)	6.57	6.30	5.57	5.83	6.83

Figure 4.9 K16's Accentedness Scores on Narration

Yet another important difference between the two raters pertains to the pronunciation domains that the raters attended to, as expressed in their analytical comments. Many of the comments of E02 in reading dealt with prosody, which the speakers might have paid less attention to during "reading" a given text; however, overall, E02 was found to have noticed and described pronunciation problems in all pronunciation domains. Although some of his comments were not quite specific, they still showed the amount of effort that E02 put into during rating. On the other hand, E05 was inclined to point out consonant and vowel issues in both reading and narration samples, and he left a relatively smaller number of comments on the other domains. In addition, although it was not shown in the rating data of K16, he was, in fact, found to have commented on the issues that were not directly related to pronunciation or provided holistic evaluations (e.g., "Odd… Sounded like a nonnative English speaker and yet his pronunciation was pretty good" [E05_K38_N]) with greater frequency than did E02. It may be that he was not as capable as E02 of analytically assessing L2 pronunciation, probably due to his limited teaching/linguistic experience and minimal familiarity with Korean, or that he chose to leave few comments and sped up to complete the rating tasks due to fatigue. In fact, a piece of evidence for the former possibility was provided by E05 himself, in his personal post-task reflection in which he wrote as follows:

> *[…] I apologize for the lack of proper knowledge of phonetics in my answers. […] I really would like to be able to use the correct wording to describe and analyze properly what I was hearing. (Oct, 21, 2017)*

To summarize, both similarities and differences were found in the data of the two raters E02 and E05, both of whom had varying extents of teaching and linguistic background. Both raters left smaller numbers of comments for the narration samples than

for the reading samples, but the current analyses did not determine a direct association between the amount of commentary the raters made for the speaker and their scalar assessment of accentedness. A closer look at the two raters' accent perception data revealed further details about the discrepancy between the raters' quantitative assessment and qualitative assessment of L2 speech samples. Although the two raters assigned seemingly similar accentedness scores to the same speech samples, their ways of identifying and describing the segmental and prosodic deviations were qualitatively different.

The current analyses of the two selected raters and one L2 speaker were essential in that they allowed for a more fine-grained observation and understanding of the rater behaviors in this study. The raters overall showed some consistency in distinguishing nativelike accents from stronger Korean accents, yet displayed varied behaviors in analytically describing the pronunciation deviations that contributed to their accent judgments. The following subsection reports on even more focused analyses of the perceptions of, and the judgments made about, a particular foreign accent feature and its acoustic properties.

4.1.4. Focused analyses 2: Ratings of a foreign accent feature

Performance data of three L2 speaker participants—one from each L2 experience group—were selected for the current analyses. Speaker K12 was a master's student studying psychology at Columbia University. She spent her middle school years in Connecticut and has been in contact with native English speakers through instruction. She was rated as the least foreign-accented among the speaker participants in this study; her total score ranked first among the 40 speakers. Speakers K30 and K39 were

international students enrolled in an affiliated school of State University of New York. They were in the final month of their freshman year at the time of data collection, and they had been taking intensive ESL courses that were designed to help international students' academic achievement in American colleges. Speaker K30's intensive use and study of L2 English was thus a year long, whereas K39 had lived in the United States and New Zealand, each for one year, prior to starting college. Speaker K39 was rated slightly better than K30 in terms of accentedness, but both speakers' scores were ranked below the 20th. The speakers' background information and their accentedness scores are also available in Table 4.22.

Table 4.22

Summary of Backgrounds and Performance of the Selected L2 Speakers (K12, K30, & K39)

Speaker	K12	K30	K39
Gender	Female	Female	Female
Age	30	20	20
L2 experience in years	7	1	3
L2 experience group	3	1	2
Accentedness Ratings:			
Reading			
Mean (SD)	7.48 (4.66)	34.52 (5.23)	30.39 (6.47)
Min-Max	5-26	22-43	17-42
Mdn	6	35	31
Narration			
Mean (SD)	7.74 (3.97)	32.57 (7.35)	30.39 (6.47)
Min-Max	5-20	15-42	17-42
Mdn	5	34	31
Total accentedness score rank (out of 40)	1	34	26

Notes: SD: standard deviation; Min: minimum; Max: maximum; Mdn: median

For the current analyses, one sentence was extracted from the reading data of each of the three speakers: *We also need a small plastic snake.* While K12 did not receive any commentary on her reading of the extracted sentence, more than half of the raters in this study commented on the segmental and prosodic issues of K30 and K39. In-depth examination of the rating data, however, revealed that there was a variation among the raters when they were making judgments about what they perceived as a phonological feature of foreign accents. One of such cases involves the consonant /l/ occurring in coda position, commonly known as velarized "dark l." Coda /l/ was perceived to have been replaced by approximant /r/, but there were also comments describing that approximant /r/ was inserted before coda /l/. In other words, for the same stimulus (e.g., *call*), the raters provided comments that were coded as segmental substitution (l » r) and as segmental insertion (r / _l$). Such variations were also observed in the raters' comments concerning K30's and K39's production of the words *also* and *small* in the extracted sentence.

Table 4.23

Commentary on Coda /l/ of K30 and K39

			K30		K39
Token	Accent feature	Counts of comments	Counts (%) of commenting raters	Counts of comments	Counts (%) of commenting raters
also	r insertion	5	5 (21.74%)	2	1 (4.35%)
	l » r	1	1 (4.35%)	9	5 (21.74%)
small	r insertion	2	2 (8.70%)	3	3 (13.04%)
	l » r	1	1 (4.35%)	1	1 (4.35%)

As shown in Table 4.23, five raters thought that K30 was inserting [r] in *also*, and only one rater (E17) reported that coda /l/ sounded more like [r]. Below are the actual comments left by these raters (Example 94–99):

Insertion of [r]
Example 94: "also" was said as "urlso" [E05_K30_R]
Example 95: "erlso" instead of "also" [E08_K30_R]
Example 96: "also" » "arlso" [E09_K30_R]
Example 97: sounded like "erlso" [E22_K30_R]
Example 98: "urlso" for "also" [E23_K30_R]

Substitution l » r
Example 99: "also" as "arso" [E17_K30_R]

Concerning coda /l/ in *small*, both [r]-insertion and substitution of /l/ with /r/ were reported, the former by two raters (E08, E15) and the latter by one (E17) (see Examples 100–102).

Insertion of [r]
Example 100: "smerl" instead of "small" [E08_K30_R]
Example 101: "small" with /r/ [E15_K30_R]

Substitution l » r
Example 102: "small" as "smar" [E17_K30_R]

Similar variations in comments were found in Speaker K39's production of *also* and *small* (see Table 4.23 for summary). In the word *also*, the l » r substitution was reported by three raters (E08, E17, and E22), and [r]-insertion by two raters (E05 and E23). Concerning the word *small*, two raters (E15 and E17) thought that coda /l/ sounded like [r], while three other raters (E03, E05, E23) perceived that there was an inserted [r] in coda.

For more objective examination, the sound samples of K30 and K39 were submitted to acoustic analyses, which revealed properties of English /r/ in the words *also* and *small*. An /r/ sound is typically identified in a spectrogram by a lowered F3 at the

transition due to retroflex articulation, and by a low F1 due to lip rounding (Ladefoged, 2005). The spectrogram of K30 (Figure 4.10) shows the steep transition of the F3 in the first syllable of *also* and in *small*. This result supports the judgment of the raters who were reported to have perceived an [r] sound in coda. The sudden rising of F3, while the F1 and F2 remained low near the bottom, before the syllable boundary of *al* (as annotated in the bottom tier) indicates that there is also a velarized lateral coda. This transition is somewhat faint in *small*, suggesting greater audibility of [r] in coda.

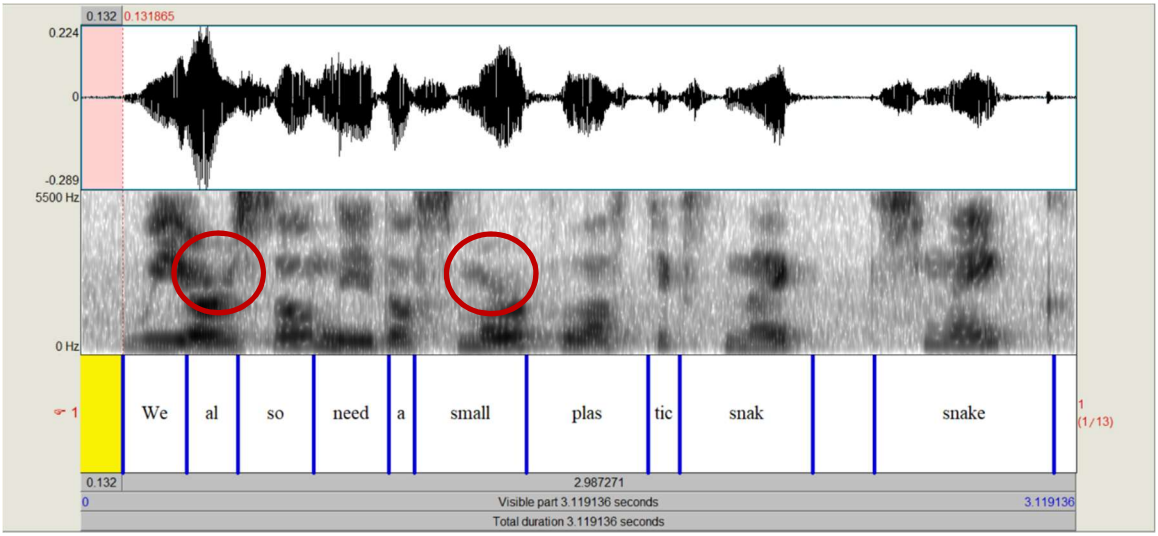


Figure 4.10 Spectrogram of K30

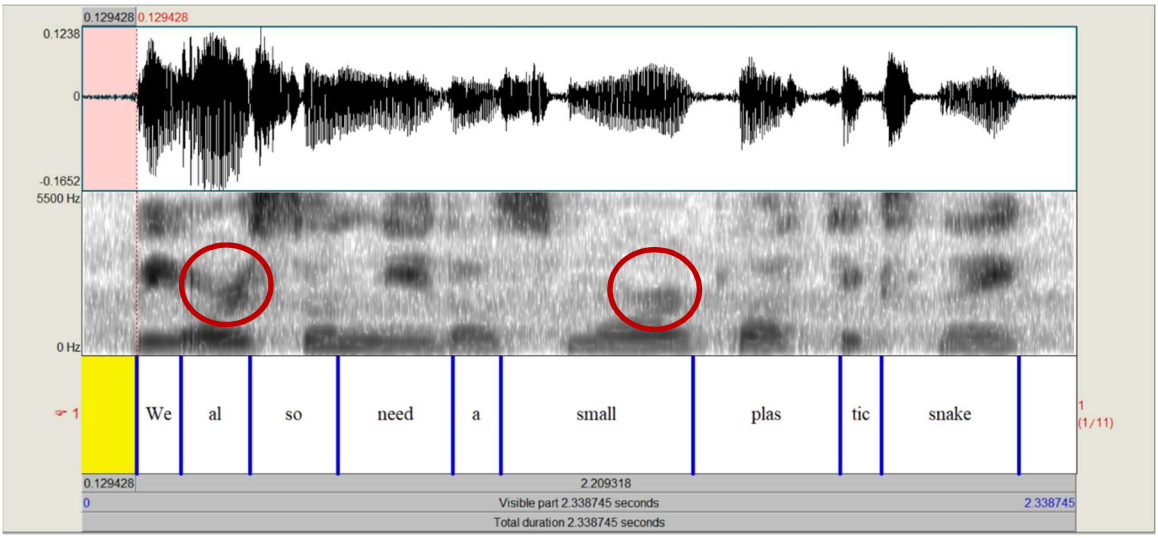


Figure 4.11 Spectrogram of K39

The spectrogram of K39's reading (Figure 4.11) also shows a steep valley (i.e., falling then rising) of the F3 in the first syllable of *also*: some part of the coda of *al* showed [r]-like acoustic characteristics and then it ended with [l]. On the other hand, in *small*, the F3 remains low until the syllable boundary, a phenomenon which may be an indication that coda was pronounced as [r].

For comparison purposes, the sentence reading of K12 was acoustically analyzed as well. Figure 4.12 shows that the position of the F3 is consistently high throughout the syllable *al* and going further up for transition to coda, while the F2 remains low, all indicating a back tongue position for lateral [l]. The F3 remains high in the rime of *small*, in contrast to what is observed in the performance of K30 and K39 (Figures 4.10 and 4.11). In short, coda /l/ in the production of K12 had the acoustic qualities of English lateral /l/ which may have made her sound the least foreign-accented to the raters.

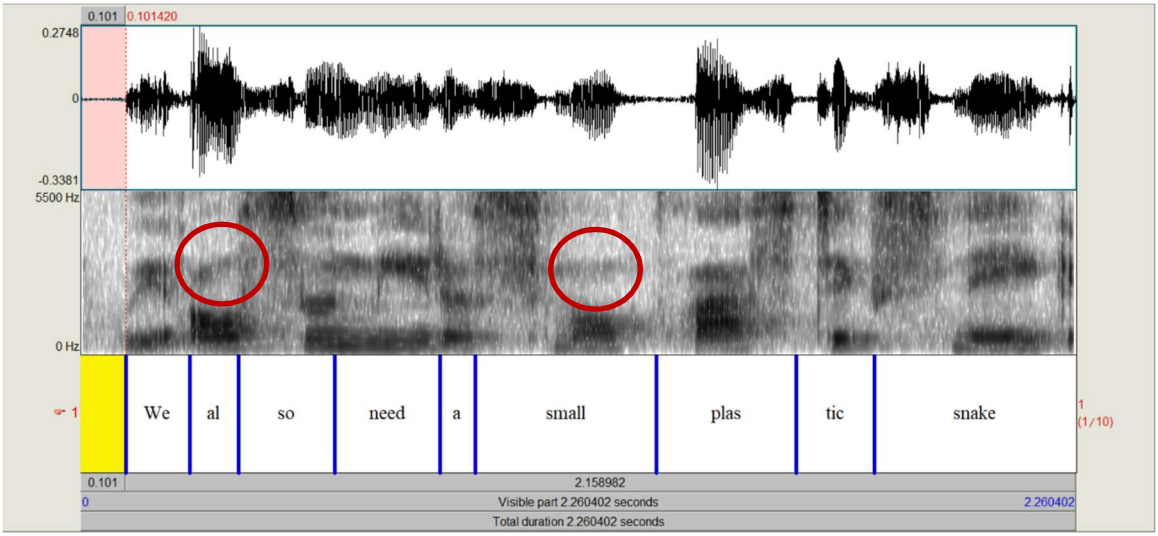


Figure 4.12 Spectrogram of K12

In summary, findings of the current focused analyses illustrated the variance among the raters and the subjective and complex nature of raters' perception and judgments with respect to foreign accents. In this particular case involving coda /l/, K30 and K39 did produce coda some part of which manifested acoustic properties of [r]; however, not all of the raters in this study pointed it out as a contributor of accentedness or wrote about it in their analytical comments. Even if they did, rarely were they sensitive to exactly what was going on in the segmentals they heard, despite their experience in ESL/EFL teaching and linguistics. Such intricacy regarding accent characteristics and accentedness was revealed through multiple qualitative and quantitative analyses of the individual participants' performance and judgments.

4.2. Phase 2

The purpose of the second phase of the study was twofold. First, it sought to measure and examine the L1 Korean-L2 English speakers' L2 phonological awareness by means of a set of phonological awareness tasks. Next, it aimed to explore the relationship between phonological awareness and foreign accentedness using the scores obtained from the L2 speaker participants, and ultimately, find answers to the second and third research questions of the current study: *Is there a relationship between foreign accent and L2 phonological awareness? If so, what is the nature of that relationship?* This section is organized as follows: first, it reports the L2 speaker participants' performance on L2 phonological awareness tasks. Then, it describes results of the correlational analyses of the phonological awareness and accentedness data from Phase 1. Lastly, the section examines the results from selected participants to delve into the relationship between phonological awareness and foreign accent.

4.2.1. Performance on L2 phonological awareness tasks

The L2 speakers overall showed varying degrees of L2 phonological awareness. Table 4.24 summarizes descriptive statistics of the scores that the L2 speakers obtained from Task 1. The score on Task 1 was the sum value of the scores that each participant earned from two judges by achieving segmental accuracy and prosodic accuracy in repeating 20 pseudowords. The highest possible score was therefore 40 (= 20 (items) * 2 (judges)) for each domain (segmentals or prosody). The scores on segmental and prosodic accuracy were different in central tendency and distribution. The mean and median scores on prosodic accuracy (M = 38.18, SD = 2.69; Mdn = 39.50) were

considerably higher than those on segmental accuracy (M = 27.10, SD = 8.26; Mdn = 28). The prosodic accuracy scores were extremely negatively distributed ($p < .01$, as assessed by Shapiro-Wilk Test of Normality, df = 40), whereas the segmental accuracy scores were in a normal distribution ($p = .28$).

Table 4.24

Descriptive Statistics of Phonological Awareness Task 1

Task 1: Pseudoword repetition		Group 1	Group 2	Group 3	Group Total
Segmental accuracy	Mean	23.93	22.90	32.50	27.10
	SD	6.73	8.10	6.80	8.26
	Min	13	9	19	9
	Max	34	32	40	40
	Mdn	25.50	26.00	34.50	28
Prosodic accuracy	Mean	37.14	37.40	39.56	38.18
	SD	3.23	3.10	.73	2.69
	Min	28	32	38	29
	Max	40	40	40	40
	Mdn	37.50	38.50	40.00	39.50

Notes: SD: standard deviation; Min: minimum; Max: maximum; Mdn: median

Differences in L2 experience groups were examined by a parametric test for the segmental data and a nonparametric test for the prosodic data. The mean segmental accuracy scores were submitted to a one-way analysis of variance (ANOVA), and a significant effect of L2 experience was found: $F(2, 37) = 7.738$, $p = .002$, $\eta^2 = .295$. A Tukey post-hoc test found that the segmental accuracy scores of the L2 speakers with six or more years of L2 experience (i.e., Group 3; 32.50 ± 6.80) were statistically significantly higher than those of the speakers in Group 2 (22.90 ± 8.10, $p < .01$) and Group 1 (23.93 ± 6.73, $p < .01$). The difference between Groups 1 and 2 was not

significant (p = .94). On the other hand, a Kruskal-Wallis H test found that the distribution of the prosodic accuracy scores was not the same across the three groups $\chi^2(2) = 7.364$, p = .025, η^2 = .184, with a mean rank scores of 15.64 for Group 1, 18.35 for Group 2, and 26.09 for Group 3. A post-hoc test with Bonferroni correction showed that the difference between Groups 1 and 3 was significant (Adjusted p = .027). The differences between other group pairs were not significant. The results indicate that the L2 speakers with more than five years of L2 experience repeated L2 pseudowords with greater accuracy, as compared to the L2 speakers with less L2 experience.

Task 2 was a phoneme perception task designed to measure the L2 speakers' awareness of L2 segmentals (phonemes). The mean of the total *d'* scores calculated in Task 2 were normally distributed (p = .16, as assessed by Shapiro-Wilk test), ranging from .89 to 2.92, and the mean value was 1.83 (SD = .55) (see Table 4.25). A one-way ANOVA showed a significant effect of L2 experience among the *d'* measures of the L2 speakers: $F(2, 37) = 9.749$, $p < .001$, $\eta^2 = .345$. A Tukey post-hoc test revealed that the *d'* of the L2 speakers in Group 3 (2.22 ± .53) were statistically significantly higher than those of the speakers in Group 2 (1.60 ± .37, $p < .01$) and in Group 1 (1.54 ± .43, $p < .01$). The difference in *d'* scores between Groups 1 and 2 was not significant (p = .93). Table 4.26 displays the mean *d'* scores that the participants obtained for the individual segmentals. The L2 speakers showed sensitivity to English sounds including /u/, /v/, and /ʃ/, but low sensitivity to /w/, /ð/, and /i/.

Table 4.25

Descriptive Statistics of Phonological Awareness Task 2

Task 2: Perception of segmentals		Group 1	Group 2	Group 3	Group Total
d'	Mean	1.60	1.54	2.22	1.83
	SD	.37	.43	.53	.55
	Min	.89	.89	1.31	.89
	Max	2.38	2.16	2.92	2.92
	Mdn	1.60	1.55	2.16	1.82

Notes: SD: standard deviation; Min: minimum; Max: maximum; Mdn: median

Table 4.26

Analysis of Items of Task 2

d' rank	Target	*d'*	Hit rate	FA rate
1	u	4.14	.91	.01
2	v	3.96	.99	.12
3	ʃ	3.56	.96	.04
4	θ	3.21	.94	.05
5	g	2.80	.95	.13
6	o	2.62	.96	.20
7	b	2.53	.86	.08
8	f	2.46	.69	.03
9	t	2.36	.87	.11
10	ʧ	2.03	.96	.40
11	r	2.02	.81	.13
12	l	1.83	.80	.16
13	s	1.60	.89	.35
14	ʤ	1.38	.69	.19
15	æ	1.18	.90	.54
16	z	1.15	.64	.21
17	j	1.09	.89	.55
18	w	.99	.87	.55
19	ð	.93	.70	.34
20	i	.69	.71	.45

Task 3 measured the L2 speakers' awareness of L2 lexical stress. As suggested by the central tendency information (M = 17.30, SD = 3.06, Mdn = 19) as well as the results from the normality test ($p < .01$) displayed in Table 4.27, the mean scores were in a

highly negatively skewed distribution. Thus, a nonparametric, Kruskal-Wallis H test was run in order to examine between-group differences in Task 3 performance. Although the test rejected the null hypothesis that the Task 3 score distribution is the same across the three groups ($\chi^2(2) = 7.655$, $p = .02$), Bonferroni-corrected multiple comparisons showed that the between-group differences were not significant: Adjusted $p = .05$ between Groups 2 and 3, Adjusted $p = .07$ between Groups 1 and 3, and Adjusted $p = 1.000$ between Groups 1 and 2.

Table 4.27

Descriptive Statistics of Phonological Awareness Task 3

Task 3: Perception of lexical stress	Group 1	Group 2	Group 3	Group Total
Mean	16.43	15.90	18.94	17.30
SD	3.13	3.78	1.53	3.06
Min	12	9	14	9
Max	20	20	20	20
Mdn	17.50	17.00	19.00	19

Notes: SD: standard deviation; Min: minimum; Max: maximum; Mdn: median

Task 4 measured awareness of L2 intonation. Similarly to the results from Task 3, the mean scores on Task 4 were in a negatively skewed distribution ($M = 12.95$, $SD = 1.63$, $Mdn = 13$, $p < .01$ as assessed by a Shapiro-Wilk test). Table 4.28 presents a summary. A Kruskal-Wallis H test did not find a significant group effect on Task 4 performance ($\chi^2(2) = 4.235$, $p = .12$), indicating that the amount of L2 experience did not exert a significant influence on the L2 speakers' awareness of L2 intonation measured in this study.

Table 4.28

Descriptive Statistics of Phonological Awareness Task 4

Task 4: Perception of intonation	Group 1	Group 2	Group 3	Group Total
Mean	12.14	13.30	13.44	12.95
SD	1.88	1.34	1.37	1.63
Min	9	11	11	9
Max	15	15	15	15
Mdn	12.00	13.50	14.00	13.00

Notes: SD: standard deviation; Min: minimum; Max: maximum; Mdn: median

Task 5 collected free responses of the L2 speakers who verbally explained pronunciation differences between four pairs of NS-NNS speech samples. The responses (681 entries in total) were coded according to the type of phonological features noticed: segmentals, prosody, and unspecified at the word level. The level of complexity and the accuracy of each entry were also coded following the coding scheme presented in Table 3.10.

To ensure the objectivity of analysis, the second coder from Phase 1 coded 25% of the entire entries (n = 170, randomly selected), and a Cohen's Kappa was run to compute the inter-coder reliability for the 170 entries. There was a high agreement between the two coders' judgments: $\kappa = .902$, $p < .001$ (95% CI .83 to .96). Because Task 5 was devised to measure explicit L2 phonological knowledge, a composite score for each participant was calculated by combining the numbers of entries at the levels of understanding (i.e., Level 2) and of understanding with metalanguage (i.e., Level 3). Examples of the coded entries are presented in Table 4.29. Preliminary coding results are summarized in Table 4.30.

Table 4.29

Samples of Verbal Explanation Entries

Entry example, translated from Korean	Noticed feature	Level of complexity
T5Entry#399 (K04) *Frog*… /f/ was not fricative. It was a plosive /p/.	Segmental	3 (understanding + metalanguage)
T5Entry#199 (K40) The <oo> sound in *spoons*, and the <ea> sound in *peas*. The vowel length… there was a difference in vowel length. The Korean speaker didn't lengthen these vowels as much as the American speaker did. Definitely shorter. The American goes 'spoon [mimicry],' while he (Korean) goes 'spun.'	Segmental	3 (understanding + metalanguage)
T5Entry#317 (K27) In *five*, the second speaker (Korean) said… something like 'pive.' It sounded something like /p/.	Segmental	2 (understanding)
T5Entry#538 (K29) /f/ in *frog* and *for* sounded very Korean.	Segmental	1 (noticing)
T5Entry#671 (K38) He (Korean) did well until 'Wednes'… but in 'day,' … in the 'ay' part. Ah, it's hard to say. I just feel it's unnatural.	Segmental	0 (perception)
T5Entry#368 (K38) They sounded pretty similar in the word *slabs*. He (Korean) said it quite well. But the intonation was different. The American speaker said 'slabs [HL],' and it was falling. But the Korean said 'slabs [LH].' He raised it and paused.	Prosodic	3 (understanding + metalanguage)
T5Entry#55 (K14) The American speaker goes 'six [speaking in a HL tone] spoons of fresh snow peas,' but the Korean speaker goes 'six spoons [LH] of fresh snow piers.'	Prosodic	2 (understanding)
T5Entry#442 (K16) Something sounds very flat.	Prosodic	1 (noticing)
T5Entry#136 (K29) He (Korean) didn't sound as melodic as the other one (American) to my ears.	Prosodic	0 (perception)
T5Entry#231 (K09) He (Korean) could have just said *of blue cheese* [fast]. But he was clearly saying 'obu' [mimicking a Korean accent].	Unspecified	1 (noticing)
T5Entry#614 (K22) In *Wednesday*, in the 'Wednes' part, they sound slightly different.	Unspecified	0 (perception)

Note: Relevant matalanguage is underlined.

Table 4.30

Task 5 (Verbal Explanation) Entries after Coding

(Raw Frequency and Percentage)

Complexity levels/accuracy	Segmentals	Prosody	Unspecified	Total
Level 0 (perception)	2 (.29%)	11 (1.62%)	25 (3.67%)	38 (5.58%)
Level 1 (noticing)	57 (8.37%)	62 (9.10%)	52 (7.64%)	171 (25.11%)
Level 2 (understanding)	212 (31.13%)	124 (18.21%)	7 (1.03%)	343 (50.37%)
Level 3 (understanding + metalanguage)	43 (6.31%)	63 (9.25%)	1 (.15%)	107 (15.71%)
Inaccurate	15 (2.20%)	1 (.15%)	6 (.88%)	22 (3.23%)
Total	329 (48.31%)	261 (38.33%)	91 (13.36%)	681 (100%)

A majority of the verbal explanations were at the understanding level of complexity (343; 50.37%). It is also indicated in the data that the L2 speakers sometimes demonstrated their analytic and metalinguistic knowledge of L2 phonology (107; 15.71%). Mere perception of pronunciation differences also occurred and was coded as low-level awareness (38; 5.58%). Nearly half (329; 48.31%) of the entries concerned the segmental differences between the native and nonnative speech samples. The prosodic differences frequently received attention as well (261; 38.33%).

The Task 5 scores are aggregated values of the numbers of individual entries made at Levels 2 and 3. As seen in Table 4.31, the mean scores of Task 5 were 6.38 (*SD* = 4.04) on segmentals, 4.67 (*SD* = 2.85) on prosody, and .20 (*SD* = .52) on unspecified elements. All of these scores were in a skewed distribution ($p < .01$, but for the prosody scores, $p = .04$). No between-group differences were found in the mean scores, according to a series of Kruskal-Wallis H tests: $\chi^2(2) = 4.965$, $p = .08$ on the segmental scores, and

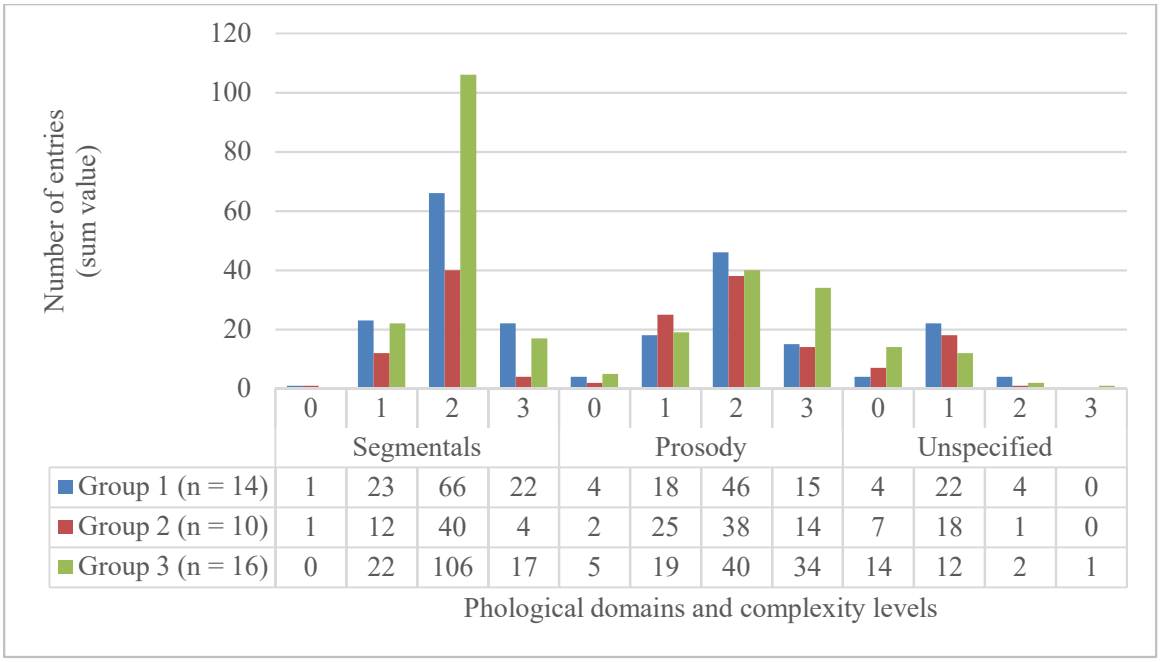
$\chi^2(2) = .658$, $p = .72$ on the prosody scores. The scores on unspecified elements were not submitted to a statistical analysis, because the size of the data was very small. The overall results from Task 5 indicate that the L2 speakers were able to verbally describe and explain NS-NNS pronunciation differences to a similar extent, irrespective of the amount of experience in L2. The findings from Task 5 are illustrated in Figure 4.13.

Table 4.31

Descriptive Statistics of Phonological Awareness Task 5

Task 5: Verbal explanation		Group 1	Group 2	Group 3	Group Total
Segmentals	Mean	6.29	4.40	7.69	6.38
	SD	4.75	2.76	3.72	4.04
	Min	1	0	3	0
	Max	19	9	18	19
	Mdn	5.00	4.00	8	6.00
Prosody	Mean	4.36	5.20	4.63	4.67
	SD	2.95	2.70	2.99	2.85
	Min	0	2	1	0
	Max	9	9	13	13
	Mdn	3.50	5.00	4.00	4.00
Unspecified	Mean	.29	.10	.19	.20
	SD	.61	.32	.54	.52
	Min	0	0	0	0
	Max	2	1	2	2
	Mdn	0	0	0	0

Notes: SD: standard deviation; Min: minimum; Max: maximum; Mdn: median



	0	1	2	3	0	1	2	3	0	1	2	3
Group 1 (n = 14)	1	23	66	22	4	18	46	15	4	22	4	0
Group 2 (n = 10)	1	12	40	4	2	25	38	14	7	18	1	0
Group 3 (n = 16)	0	22	106	17	5	19	40	34	14	12	2	1

Figure 4.13 Entries Made by L2 Speakers

A Spearman's rank order correlation was used to determine the relationships among the five phonological awareness tasks. Some tasks were significantly and positively correlated with one another, as shown in Table 4.32. Non-significant correlations were found between the prosody score from Task 5 and all the other measures of L2 phonological awareness in the current study. In other words, the L2 speakers' performance on verbal explanation of L2 prosodic features and rules was not significantly associated with their abilities to accurately perceive or repeat L2 sounds (Tasks 1 through 4).

In addition, in exploring the relationships among the components concerning L2 segmentals and prosody in Tasks 1 and 5, the study found that the segmental components of Task 1 and Task 5 were significantly correlated to each other ($r_s = .43$, $p < .01$), whereas no such significant association was found between the prosodic components of

the two tasks (r_s = .07, p > .05) (see Table 4.32). This suggests that the relationship between the ability to use L2 phonology and the ability to provide (metalinguistic) explanation of L2 phonology is intricate and thus needs elaborated discussion.

Table 4.32

Intercorrelations between Phonological Awareness Measures

Measure	1S	1P	2	3	4	5S	5P
1S. Task 1 (segmentals)	-	.63**	.78**	.51**	.63**	.43**	.25
1P. Task 1 (prosody)		-	.50**	.53**	.49**	.23	.07
2. Task 2			-	.51**	.54**	.38	.01
3. Task 3				-	.34*	.29	-.10
4. Task 4					-	.47**	.26
5S. Task 5 (segmentals)						-	.05
5P. Task 5 (prosody)							-

Notes: *p < .05, **p < .01

4.2.2. Correlations between L2 phonological awareness and accentedness

Next, in order to answer the second and third research questions and determine the relationship between L2 phonological awareness and foreign accent, correlations between measures of L2 phonological awareness and accentedness rating scores—a quantifiable measure of foreign accent—were calculated (see Table 4.33, Figures 4.14 and 4.15). A Spearman's correlation revealed that accentedness scores from both reading and narration were inversely and significantly correlated with each of the phonological awareness measures, but not with the prosody scores from Task 5. The lack of associations between Task 5 (prosody) and accentedness scores is illustrated in the scatter plots below—Figures 4.14(g) and 4.15(g).

Table 4.33

Correlations between L2 Phonological Awareness and Accentedness

Speaking task	1S	1P	2	3	4	5S	5P
Reading	-.81**	-.57**	-.75**	-.35*	-.54**	-.47**	-.22
Narration	-.82**	-.71**	-.71**	-.43**	-.59**	-.40*	-.25

Notes: *$p < .05$, **$p < .01$

Then, stepwise multiple regression analyses were carried out in order to determine the relative contribution of the different components of L2 phonological awareness to accentedness rated in reading and narration. The reading scores and the narration scores were a dependent variable in two separate regression analyses. The seven L2 phonological awareness measures were entered as independent variables. Normality and homoscedasticity of the residuals were inspected from the scatter plots and probability plots of the regression standardized residuals. Presence of outliers was inspected by computing Cook's distances and was determined as having a standardized residual score of +/- 3 or a Cook's distance score over three times of the mean score (.028 in reading and .039 in narration). Among the forty L2 speaker participants, one participant (coded as K18) was detected as an outlier in both reading and narration. The Durbin-Watson values were 1.888 in reading and 1.903 in narration, indicating that the data were not auto-correlated.

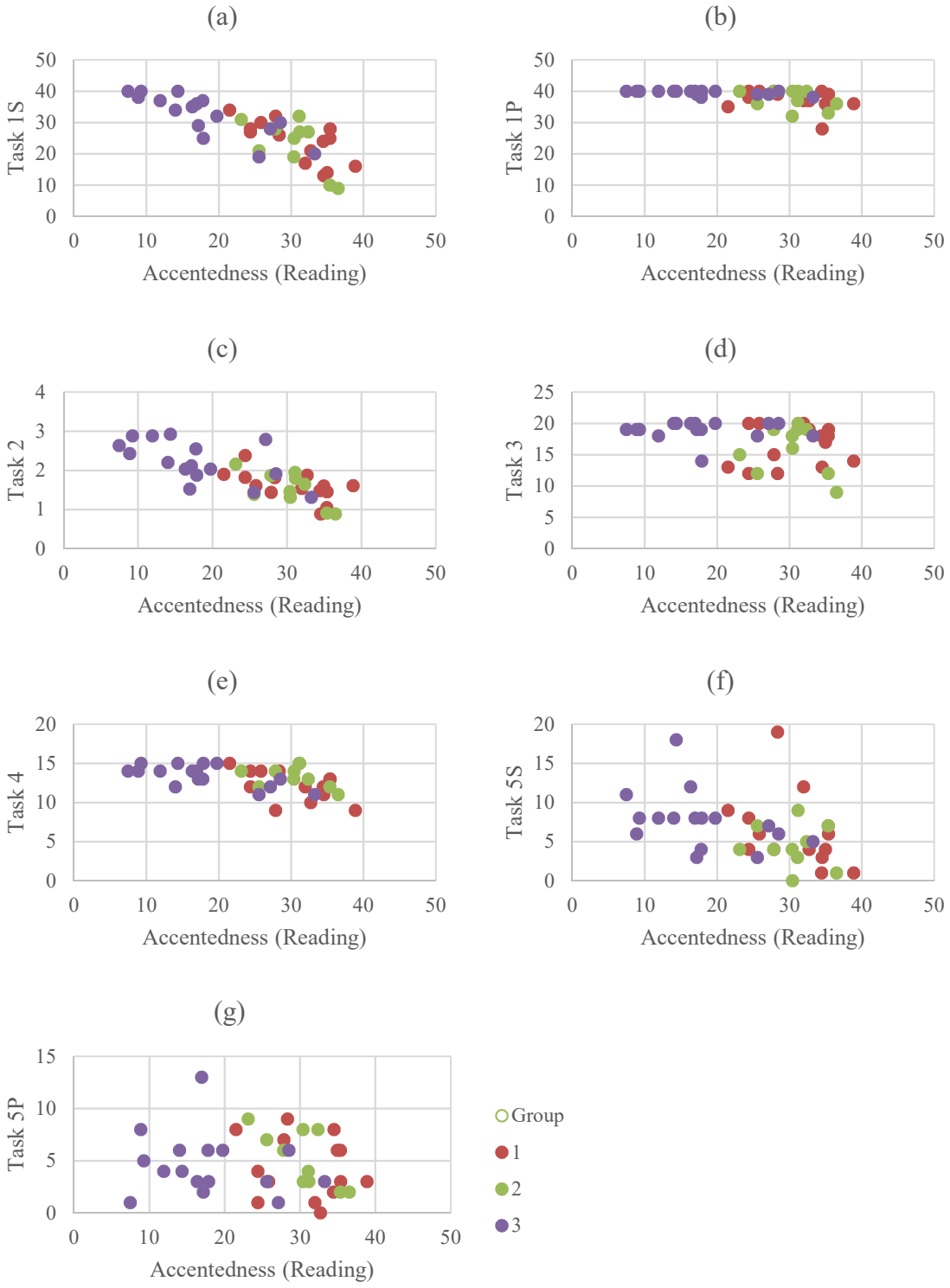


Figure 4.14 Mean Accentedness Scores in Reading and Measures of L2 Phonological Awareness

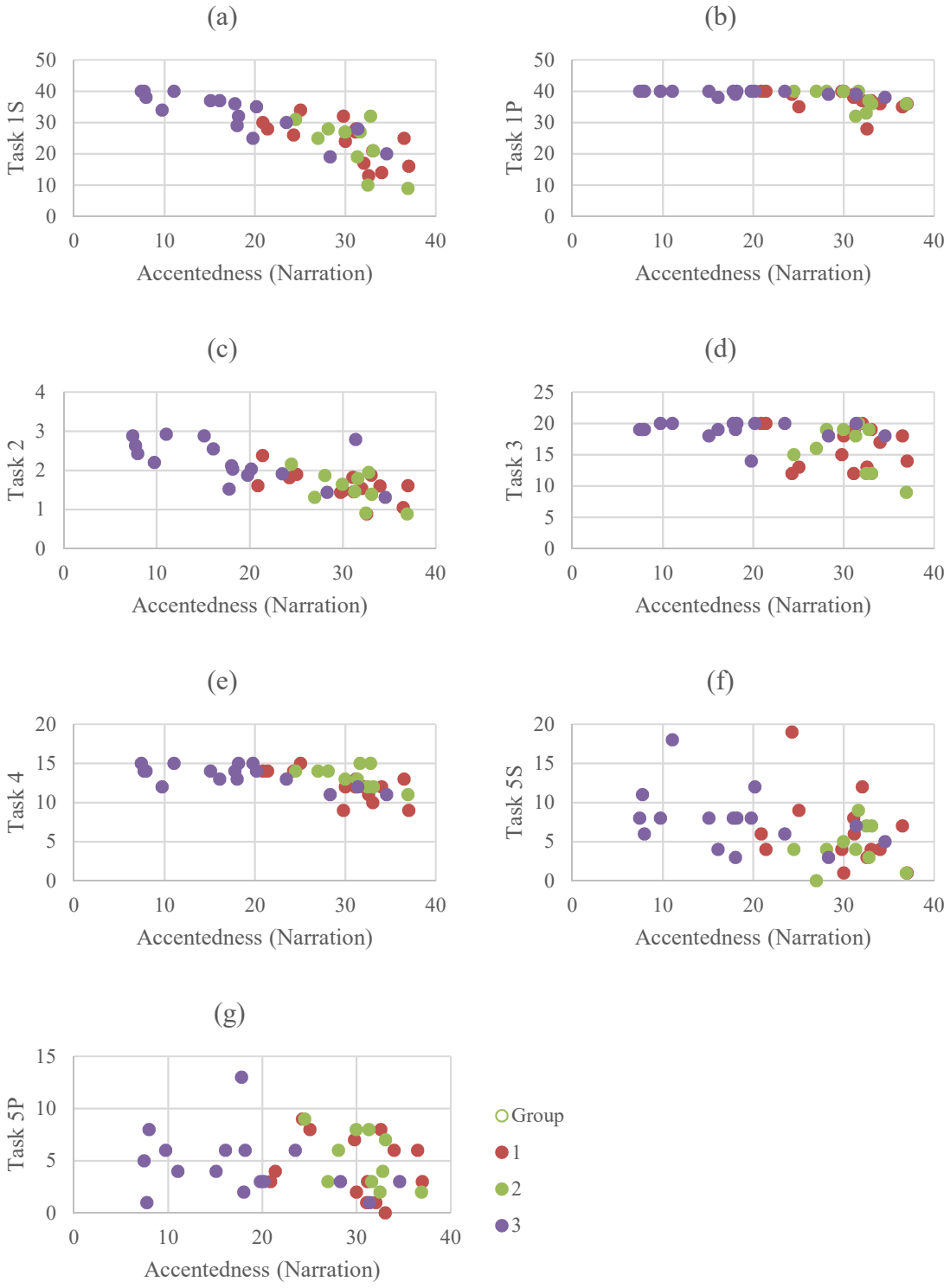


Figure 4.15 Mean Accentedness Scores in Narration and Measures of L2 Phonological Awareness

Table 4.34 summarizes stepwise multiple regression analyses of phonological awareness variables (Tasks 1 through 5) on accentedness in reading and narration. The significance level for all tests was .05. The regression models generated in the analysis of reading was statistically significant (e.g., in the final model, $F(2, 37) = 43.50$, $p < .001$). Segmental accuracy on Task 1 (delayed repetition), which best correlated with the dependent variable, alone accounted for 65.6% of the variance in accentedness in reading. Performance on Task 2 was entered in the second and final step, accounting for an additional 4.6% ($\beta = -5.379$) of the variance. The remaining phonological awareness variables did not contribute to the regression models and thus were excluded in the final step. On the other hand, in the analysis of narration, also presented in Table 4.34, segmental accuracy on Task 1 was the only strong predictor variable in the regression model: $F(1, 38) = 68.353$, $p < .001$. The other variables exerted no significant effects on the rating and were removed from the model.

Table 4.34

Stepwise Multiple Regression: L2 Phonological Awareness and Accentedness

Predictors	R^2	β	95% CI
Reading			
Step 1			
Task 1 (segmentals)	.656	-.845**	[-1.046, -.644]
Step 2			
Task 2 (*d'*)	.702	-5.379*	[-9.943, -.816]
Narration			
Step 1			
Task 1 (segmentals)	.643	-.847**	[-1.055, -.640]

Notes: *$p < .05$, **$p < .01$

The regression models overall indicated that the relationship between L2 phonological awareness and accentedness is quite nuanced. Although most measures of L2 phonological awareness were inversely and significantly correlated with accentedness ratings in reading and narration, implicit knowledge of L2 segmentals, demonstrated by use of L2 segmentals—particularly by accurately repeating L2 phonological input, was the largest predictor for accentedness. Other L2 phonological awareness measures in the current study, including the ones that tap into awareness of L2 prosodic features, were found to play relatively little roles.

4.2.3. Focused analyses 3: Five segmentals /r, ð, z, i, æ/

The study conducted even deeper analyses that aimed at characterizing the relationship between foreign accent and awareness of particular L2 segmentals. In doing this, first, the L2 speaker participants were ranked according to the sum value of the accentedness scores they received in reading and narration. The bottom 25% (i.e., almost nativelike, NL) and the top 25% (i.e., heavily accented, HA) were selected from the cohort (n = 10 each). The mean accentedness scores of the two groups were 13.41 (*SD* = 3.62) for the NL group, and 35.14 (*SD* = 1.62) for the HA group. The NL group was composed of speakers in Group 3 (average amount of L2 experience: 12.3 years); on the other hand, the HA group consisted of speakers of a wide range of L2 experience: seven from Group 1, two from Group 2, and one from Group 3. Then, these participants' performance on Tasks 2 and 5 was examined along with the qualitative comments that they received in Phase 1. Each of the participants' data was expected to be classified under one of the cases presented in Table 4.35.

Table 4.35

Classification of Eight Possible Cases

Case	The participant provided verbal explanation about an L2 segmental during Task 5 (i.e., score > 0). If not, score = 0.	The participant demonstrated knowledge of the segmental and showed sensitivity to it during Task 2 (i.e., $d' > 0$). If not, $d' \leq 0$.	The participant sounded accent-free, targetlike in using the segmental and received no comments. If five or more[4] NS raters noticed his or her pronunciation of the segmental, the production is considered as nontargetlike or accented.
Measured/ observed construct	Explicit, verbalizable knowledge of L2 phonology	Non-verbalizable L2 phonological knowledge (less explicit and more implicit than in Task 5)	Features that contributed to accentedness, perceived by NS listeners
A	Score > 0	$d' > 0$	Targetlike (Not commented)
B	Score > 0	$d' > 0$	Nontargetlike (Noticed by five or more raters)
C	Score > 0	$d' \leq 0$	Targetlike
D	Score > 0	$d' \leq 0$	Nontargetlike
E	Score = 0	$d' > 0$	Targetlike
F	Score = 0	$d' > 0$	Nontargetlike
G	Score = 0	$d' \leq 0$	Targetlike
H	Score = 0	$d' \leq 0$	Nontargetlike

If an L2 speaker exhibits his or her ability to accurately notice and verbally explain a certain sound (segmental) in L2 phonology (i.e., earns a positive score in Tasks 2 and 5) and is judged by native listeners to be able to use that segmental in a targetlike way (Case A), it can be assumed that the speaker's knowledge of that segmental (e.g., targetlike use of the segmental) has been built into his or her interlanguage system and has been successfully used in L2 speech production. However, the speaker may still be perceived to make nontargetlike production, sounding foreign-accented (Case B). In Cases C and D, the speaker exhibits explicit knowledge but the status of implicit

[4] The criterion was arbitrarily determined under the assumption that the number 5 can be a borderline for harsh raters and the other, less severe raters in this study: 5 raters = 21.74% of the cohort in this study

knowledge is not clear. The speaker may either produce a targetlike (i.e., accent-free) sound (Case C) or fail to do so (Case D). In Cases E and F, the speaker does not verbalize anything about the target L2 sound, or even if something is verbalized, it is not accurate, while there is an indication that implicit knowledge is somehow present and working, or that the speaker may simply have some intuition about it. It is also possible that explicit knowledge has gradually eroded over time after implicit knowledge was built through practice. The speaker may sound accent-free when using that segmental (Case E) or may not (Case F). The speaker in Case G would have produced a targetlike segmental somehow, but perhaps by chance—not based on his or her knowledge. Another possibility is that the speaker's existing knowledge was not tapped into by the tasks. The last case, H, would be the case in which learning of the segmental has not taken place.

Five L2 segmentals were selected for the analyses: /r, ð, z, i, æ/. They were chosen because (a) they were among the segmentals that, according to the majority of the raters in Phase 1, sounded foreign-accented (see Table 4.3), and (b) they appeared not only in the stimuli of Task 2 but also in those of Task 5, hence, the L2 speakers were given an opportunity to verbalize their knowledge of these segmentals. In Task 2, the entire population's mean *d'* scores for these segmentals were below 2, except for /r/ (between 2 and 3; see Table 4.26). For each segmental, the number of L2 speakers belonging to each case was counted, and between-group differences were examined. Results are summarized in Tables 4.36 and 4.37. Between-group comparisons for the selected segmentals are illustrated in Figure 4.16.

Table 4.36

Classification of the NL Group's Performance across Tasks

Case	/r/	/ð/	/z/	/i/	/æ/	Total
A	5	0	8	4	5	22
B	0	1	1	1	0	3
C	0	0	0	2	0	2
D	0	0	0	2	0	2
E	5	3	1	0	3	12
F	0	1	0	0	2	3
G	0	4	0	0	0	4
H	0	1	0	1	0	2

Table 4.37

Classification of the HA Group's Performance across Tasks

Case	/r/	/ð/	/z/	/i/	/æ/	Total
A	0	0	0	0	0	0
B	0	1	1	1	0	3
C	0	0	2	0	0	2
D	1	0	1	2	0	4
E	0	0	2	0	5	7
F	6	9	1	2	1	19
G	1	0	1	0	2	4
H	2	0	2	5	2	11

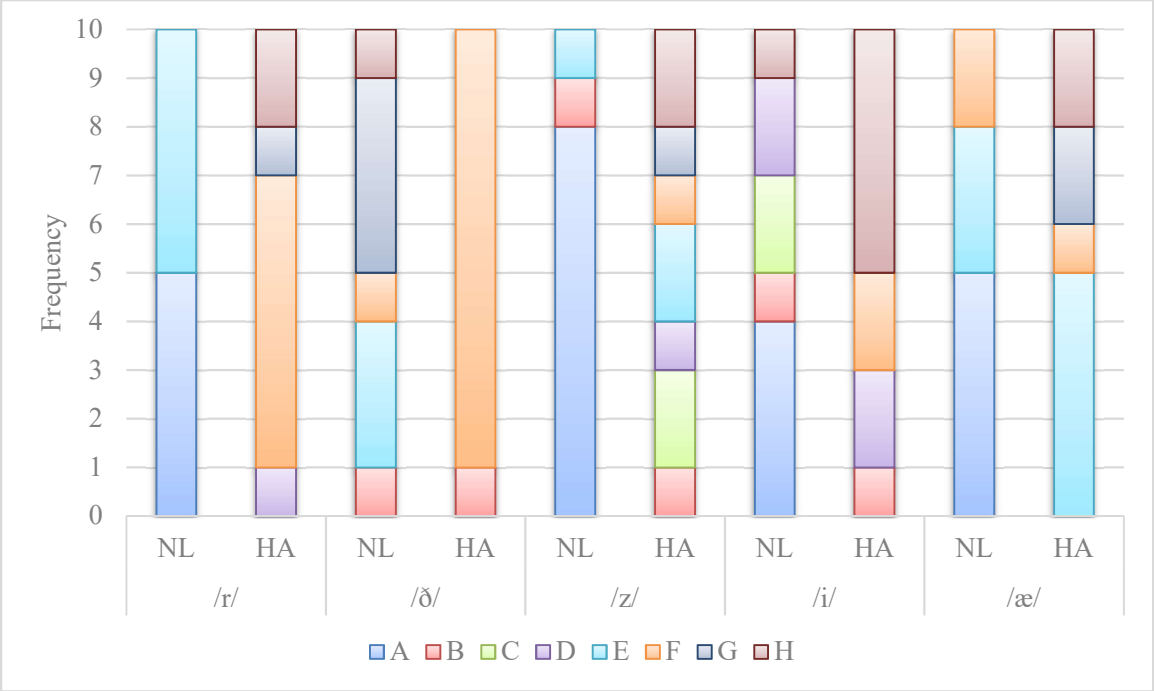


Figure 4.16 Comparison of the Performance of the NL and the HA Groups

As presented in Tables 4.36 and 4.37 and Figure 4.16, each of the closely examined segmentals showed a result different from those of the other segmentals. Starting with /r/, the NL speakers were judged to produce the sound in a targetlike way while successfully performing on both Tasks 2 and 5 (Case A) or only on Task 2 (Case E). This indicates that the NL speakers overall had some implicit knowledge of /r/ and that half of them were able to verbalize the characteristics and rules about the segmental in addition. On the other hand, most of the HA speakers were judged to be foreign-accented when pronouncing /r/ in Phase 1. Six of them could demonstrate use of the segmental, but their explicit knowledge of it seemed to be limited (Case F). The HA speakers showed a similar pattern with /ð/. Nine of them fell into Case F: they were able to use (in Task 2, perceive and notice) /ð/ with accuracy, but their performance on Task 5

and the speaking tasks for accentedness ratings was not successful. A rather interesting result was obtained from the NL speakers. Three of them fell into Case E, suggesting that they did not have the explicit kind of awareness about /ð/ but only some implicit knowledge of it, but their production of /ð/ was assessed as targetlike. Four of the NL speakers did not show their awareness about /ð/ (or their awareness was not tapped into either by Task 2 or Task 5), but they still showed targetlike performance, falling into Case G.

Moving onto /z/, which most of the NL speakers had good awareness about (and fell into Case A), the HA speakers were almost evenly distributed across the cases. The HA group's overall knowledge of the target segmental seemed to be incomplete and inconsistent. Next, most of the NL speakers were able to verbally describe and explain the features of /i/, and some of them showed their implicit knowledge of the segmental as well (Cases A and B). However, the NL speakers' pronunciation of /i/ was often assessed as nontargetlike (Cases B, D, and H). On the other hand, none of the HA speakers produced accent-free /i/, and their knowledge of /i/ was inconsistent. Half of them fell into Case H, which suggests that neither explicit nor implicit knowledge of /i/ had been built into their interlanguage system yet. Finally, the differences between the NL and the HA groups were once again evident in the data of /æ/. The NL speakers overall had some awareness about the segmental (at least implicitly), but their /æ/ was sometimes judged to be foreign-accented: i.e., Only two of them fell into Case F. The HA speakers, on the other hand, clustered in Cases E through H—that is, their explicit knowledge of /æ/ was not shown in their performance on Task 5. Six of the HA speakers demonstrated their

ability to use /æ/ in Task 2, and five of them produced it in a targetlike way (Case E); however, it remains doubtful if the other HA speakers had any awareness about /æ/.

In summary, a large number of the NL speakers were found to fall into either Case A or Case E. In the former case, both explicit knowledge (as measured through Task 5) and implicit knowledge (as measured through Task 2) of target segmentals were shown and the speakers were judged to produce these segmentals without detectable foreign accent. In the latter case, speakers showed only implicit knowledge of the target segmental and pronounced it without foreign accent. However, fewer instances were observed where only explicit knowledge was present (Cases C and D) or where no tacit or verbalized knowledge was measured (Cases G and H). On the other hand, the HA group showed a different behavioral pattern. In general, their accuracy rates for the target segmentals were lower than those of the NL group, as manifested by high frequencies of Cases F and H. The HA speakers' knowledge of the target segmentals seemed to be largely implicit, and their verbalization of the rules and characteristics of the target segmentals was limited, as compared to the NL speakers. It is also important to note that for each target segmental, the NL and HA speakers' performance patterns slightly differed, implicating that individual sounds are unique and different in learnability.

4.2.4. Focused analyses 4: Individual L2 speakers

The relationship between accentedness and phonological awareness appears to be intricate and nuanced. But findings of this study suggest that the L2 speaker participants might be an important source of variation as well. Therefore, additional analyses were conducted in the spirit of achieving more understanding of the behavior of the L2

speakers in this study and the connection between their accents and L2 phonological awareness.

For the current analyses, four L1 Korean-L2 English speaker participants were selected. Two of them, K16 and K30, appeared previously in the focused analyses conducted in Phase 1. The two other L2 speaker participants, K02 and K11, were randomly selected. The background information of the selected L2 speakers, collected using the questionnaires described in Chapter III, is summarized in Table 4.38. In the following, each selected speaker's performance is described.

Table 4.38

Summary of Backgrounds and Performance of the Selected L2 Speakers (K02, K11, K16, & K30)

Speaker	K30	K16	K11	K02
Gender	Female	Male	Female	Male
Age	20	31	34	31
L2 experience in years	1	1	11	18
L2 experience group	1	1	3	3
Accentedness Ratings:				
Reading				
Mean (SD)	34.52 (5.23)	24.39 (7.80)	16.35 (7.52)	11.91 (3.53)
Min-Max	22-43	5-37	7-32	5-20
Mdn	35	26	13	12
Narration				
Mean (SD)	32.57 (7.35)	31.09 (5.20)	20.17 (7.36)	15.09 (6.68)
Min-Max	15-42	20-40	8-35	5-30
Mdn	34	30	20	15
Total accentedness score rank (out of 40)	34	20	10	6
Counts of comments:				
Reading				
Segmentals	67	59	31	26
Prosody	32	32	25	34
Fluency	0	3	0	1
Holistic	0	0	4	4
Irrelevant	0	1	1	0
Narration				
Segmentals	69	63	27	19
Prosody	69	22	23	19
Fluency	1	1	0	0
Holistic	2	0	4	4
Irrelevant	3	1	1	0
Total counts of comments	243	182	116	107
Phonological awareness:				
Task 1S	13	27	35	37
Task 1P	28	38	40	40
Task 2	.89	1.82	2.03	2.88
Task 3	13	12	20	18
Task 4	11	12	14	14
Task 5S (Levels 2 & 3)	3	8	12	8
Task 5P (Levels 2 & 3)	8	1	3	4
Task 5U (Levels 2 & 3)	1	0	0	0

Notes: SD: standard deviation; Min: minimum; Max: maximum; Mdn: median

Speaker K30. As previously introduced in Section 4.1.4, Speaker K30 was an international student enrolled in an American college. Her intensive use and study of L2 English was a year long, while her first exposure to English took place in her middle school in Korea.

In Phase 1 of the study, K30 earned 34.52 in reading and 32.57 in narration, and her total score was ranked 34th among the forty speakers (Thus, she was one of the HA speakers in Section 4.2.3). She received a total of 243 comments from the raters, which was the highest number among the speakers selected for the current analyses. More than half of these comments were about her performance in reading. The most commented pronunciation problems had to do with the use of L2 segmentals, particularly vowels such as /æ/ (Example 103), /ɔ/, and the tense-lax pair of /i, ɪ/ (Examples 104–105), and consonants /θ, ð/ (Example 106). She was also perceived to have difficulty in using liquid consonants /l, r/, particularly coda /l/, as shown in Section 4.1.4 (Examples 94–102). On the other hand, few comments were made about syllable simplification phenomena such as vowel epenthesis or consonant deletion. All of her lexical stress errors were a result of her tendency to stress the wrong syllable within a multisyllabic word (Example 107). As for the phrasal stress/rhythm, she was often heard to speak in nontargetlike rhythm by stressing function words (Example 108) and some content words that should not be focused within utterance. Finally, most of the comments made about her use of L2 intonation described that she tended to use a rising tone in the end of a sentence, making a statement sound like a question or a statement of uncertainty (Example 109).

Example 103: “ask” was said as “esk” [E05_K30_R]
Example 104: First vowel in “windows” is /i/ [E22_K30_N]

Example 105: “these” as “this” [E17_K30_R]
Example 106: “the” as “da” [E09_K30_R]
Example 107: “mayBE” [E01_K30_R]
Example 108: “at THE train station” [E16_K30_R]
Example 109: rising in the end of sentence [E15_K30_N]

In Phase 2, which administered L2 phonological awareness tasks, K30 scored 41 out of 80 on Task 1 (pseudoword repetition): 13 in segmental accuracy and 28 in prosodic accuracy. These subscores were below average of the entire cohort (M = 27.10 (SD = 8.26) in segmental accuracy and M = 38.18 (SD = 2.69) in prosodic accuracy). It seems that she struggled when repeating the pseudowords that were aurally presented to her for Task 1. Her low performance score in prosodic accuracy is particularly noticeable, considering the fact that the speakers in this study overall had negatively skewed scores in prosodic accuracy in this task. In Task 2 (perception of segmentals), her overall sensitivity score (*d'*) was .89. She showed particularly low sensitivity to the following English sounds: /i, ɪ, æ, j, r, l, ð/ (see Table 4.39). All of these sounds except for /j/ were commented to be problematic in her production speech assessed in Phase 1. This indicates that K30 did not have awareness of these segmentals at the time of data collection.

Table 4.39

K30's Performance on the Selected Items of Task 2

Selected target	K30			All L2 speakers (Mean)		
	d'	Hit rate	FA rate	*d'*	Hit rate	FA rate
i	-3.29	.05	.95	.69	.71	.45
æ	-.55	.90	.97	1.18	.90	.54
r	1.71	.66	.10	2.02	.81	.13
l	-1.64	.50	.95	1.83	.80	.16
ð	.85	.90	.66	.93	.70	.34
θ	1.71	.66	.10	3.21	.94	.05

In Task 3 (perception of lexical stress), K30 failed to correctly identify the stressed syllable in seven items (pseudowords). Five of these words were trisyllabic. In these words, she perceived the primary stress to be placed on the penultimate syllable, even when the stress was in fact on the first or last syllable: e.g., *LItholect* perceived as *liTHOlect*; *trumpeTINE* perceived as *trumPEtine*. Her performance on Task 3 parallels with her production of lexical stress assessed in Phase 1. Her score on Task 4 (perception of prosody) was 11 out of 15. K30 showed weakness particularly in interpreting the intended meaning of an ambiguously written noun phrase (e.g., *black hole, hot dog*) based on its prosodic cues—rhythm and intonation.

Finally, in Task 5 (verbal explanation), she was able to verbalize some NS-NNS differences in the stimuli, but she mostly attended to the prosodic differences such as intonation, pauses, and rhythm, sometimes pointing at a nonspecific pronunciation issues: e.g., "I think the pronunciation of *cheese* in *blue cheese* was wrong" (T5Entry#337). In the entries discussing segmental differences, her primary focus seemed to be on particular consonants like /f, z/, but not on vowels: she did not attend to any of the segmentals to which she showed low sensitivity in Task 2. Her lack of knowledge of these L2 segmentals may have resulted in nontargetlike production, as analytically described by the raters in Phase 1, and this may have contributed to her strong accentedness.

K30 rarely used relevant metalanguage in her verbal responses. Only in one entry was her expressed L2 phonological awareness at the level of understanding with metalanguage (i.e., Level 3):

- "Here (the native speech), *train station* sounded clear, but in the second sample (the nonnative speech), there was a strong linking between the words, as if the /s/ in the second word was pulled and added to the first word, and the rhythm made it sound like *trains tation*" (T5Entry#647).

K30's scores on Task 5—the aggregated value of the number of entries made at Levels 2 and 3—were 3 in segmentals, 8 in prosody, and 1 in unspecified words.

Speaker K16. As previously introduced in Section 4.1.3, Speaker K16 had been in the United States for approximately one year at the time of data collection. His total accentedness score was ranked 20th. He received 182 comments from the 23 raters. Most of these comments concerned his nontargetlike production of L2 vowels, such as mixed usage of /o, ɔ/ (Example 110), /i, ɪ/ (Examples 39, 111), /ɛ, æ/ (Example 112), and [ə, ʌ]. The most problematic consonants were /l/ and /r/, which were often used interchangeably (Example 113), and /ð/, which was often replaced with /d/ (Example 114). Many raters commented that he tended to add a vowel after a coda consonant (Example 115). In addition, he was perceived to omit word-final /t/ in coda consonant clusters. Most of his lexical stress issues were a result of his tendency to place primary stress on the wrong syllable (Example 116). He was sometimes heard to produce nontargetlike rhythm by stressing unfocused words, as well as nontargetlike intonation by speaking in flat tone (Example 117).

Example 110: "call" was said as "col" in line 1 [E23_K16_R]
Example 111: vowel in "meet" not lengthened enough [E17_K16_R]
Example 112: "snack" » "sneck" [E09_K16_R]
Example 113: Uses a mix between /l/ and /r/ in "handle" [E08_K16_N]
Example 114: "these" was said with a *d* in line 8 [E23_K16_R]
Example 115: "fresh-i" [E04_K16_R]
Example 116: "SEvere" instead of "seVERE" [E22_K16_N]
Example 117: intonation at the end of line 1 was neutral [E05_K16_R]

In Phase 2 of the study, K16 scored 65 out of 80 on Task 1, which approximated the mean score of the entire cohort (*M* = 65.28, *SD* = 10.20). The subscores were 27 in segmental accuracy and 38 in prosodic accuracy. In Task 2, his overall sensitivity score (*d'*) was 1.82, which, again, was approximate to the average of all of the Korean speaker participants in this study (*M* = 1.83, *SD* = .55). He was particularly weak at identifying and discriminating the vowels /i, ɪ/ and the consonants /s, z/ (see Table 4.40). The former ones, /i, ɪ/, were among the sounds that he pronounced with a detectable Korean accent in Phase 1. Interestingly, he performed fairly well with the consonants /l, r, ð/, which were problematic in his reading and narration.

Table 4.40

K16's Performance on the Selected Items of Task 2

Selected target	K16			All L2 speakers (Mean)		
	d'	Hit rate	FA rate	*d'*	Hit rate	FA rate
i	-3.29	.05	.95	.69	.71	.45
æ	1.71	.90	.33	1.18	.90	.54
r	3.12	.97	.10	2.02	.81	.13
l	1.64	.95	.50	1.83	.80	.16
ð	1.71	.90	.33	.93	.70	.34
s	.00	.50	.50	1.60	.89	.35
z	1.64	.50	.05	1.15	.64	.21

On the other hand, K16 scored 12 on Task 3. A closer examination of his responses revealed that in five items, he selected a syllable that received the secondary stress, not the primary stress, which was the target feature of the task. For instance, he perceived the stress was placed on the antepenultimate syllable in *repastaFIQUE*, and on the first syllable in *saPIrical*. This suggests that while he was aware of stress in English (if only vaguely so), he did not have the sort of phonological knowledge that could

distinguish between the primary stress and the secondary stress. This limited knowledge of L2 lexical stress may also account for his incorrect stress placement reported by the raters in Phase 1. On Task 4, he scored 12 out of 15. While this score was not significantly lower than the mean of the participants, his data showed that he had weak L2 prosodic awareness—especially, awareness of sentence-level intonation and pauses in English—and thus was not able to identify the intended meaning of the ambiguously written question such as "What's in the tea, honey?" Lastly, in Task 5, he demonstrated some knowledge of L2 phonology and scored 8 in segmentals and 1 in prosody. Most of his entries focused on the segmental differences in NS-NNS speech samples as in the following entries:

- "In *frog for*, all /f/ sound is pronounced as /p/. *Frog* as 'progu'" (T5Entry#441).
- "I'm not quite sure, but in *fresh*, I hear something close to /l/. (Which sound?) In the /r/ sound" (T5Entry#65).

In contrast, few prosodic differences were verbally explained: e.g., "Something sounds very flat" (T5Entry#442). His performance on the tasks in Phase 2 overall indicate that his awareness of L2 prosody—especially, explicit knowledge of L2 prosody—was somewhat limited, as compared to L2 segmental awareness.

Speaker K11. Speaker K11, age 34, arrived in the United States as an adult, at age 23. She completed college education in California and had been working as a registered nurse in New York since 2015. Her first exposure to English was in the beginning of her middle school years. She had also taken some ESL classes during her college years. Her recent score on the TOEIC test was 990, which is a perfect score.

Her accentedness scores were 11.91 in reading and 15.09 in narration (i.e., she sounded a bit more accented in her spontaneous speech), and the total score was ranked

10th—at 25 percentile (She was thus included in the NL group in Section 4.2.3). The raters left a total of 116 comments in which they pointed out various pronunciation issues detected in her speech. Some of these comments described that K11's vowels /i, ɪ, ɛ, æ/ were occasionally nontargetlike (Examples 118–119). Other comments noted that she often pronounced English /θ, ð/ as /t, d/ (Example 120). However, there were very few comments that pointed to her production of other segmentals. A smaller number of comments were made with respect to her use of L2 prosody. According to these comments, she did not speak with natural rhythm or intonation when taking the reading task (Example 121). The task effect might have been in play here, as she also gave an informal reflection immediately after the reading task that she was not sure to what level of clarity the text had to be read aloud. Her use of prosody in narration was less commented on, yet it appeared that she had a tendency to end a phrase or a statement with a rising tone, which was a frequently observed phenomenon in Korean-accented speech.

Example 118: "six" pronounced with vowel /i/ [E22_K11_R]
Example 119: "meet" sounds like "mit" [E04_K11_R]
Example 120: "there" was said with a *d* in line 11 [E05_K11_R]
Example 121: "three red bags" is unnecessarily stressed [E07_K11_R]

Then, her performance on the phonological awareness tasks in Phase 2 was examined. On Task 1, she scored 35 in segmental accuracy and 40 in prosodic accuracy: although she could not repeat some pseudowords with segmental accuracy, she was able to maintain and reproduce the prosodic characteristics of the pseudowords in the task. In Task 2, she had difficulty in perceiving and identifying the vowels /i, ɪ, æ/—the vowels that she could not correctly produce in Phase 1 (see Table 4.41). Her score on the /i-ɪ/ pair was particularly low—in fact, not different from that of K16, who had far less

experience in L2 use. In addition, she showed low sensitivity to /ð/ and /t/, which were also problematic in her speech in Phase 1.

Table 4.41

K11's Performance on the Selected Items of Task 2

Selected target	K11			All L2 speakers (Mean)		
	d'	Hit rate	FA rate	*d'*	Hit rate	FA rate
i	-3.29	.05	.95	.69	.71	.45
æ	.85	.90	.67	1.18	.90	.54
r	3.12	.97	.10	2.02	.81	.13
l	3.29	.95	.05	1.83	.80	.16
ð	.85	.90	.67	.93	.70	.34
t	.00	.50	.50	2.36	.87	.11

Her performance on Tasks 3 and 4 was near perfect. The only item that challenged her in Task 4 required her awareness of sentence-level intonation and pauses in a question containing a list of content words: e.g., "Do you have a pain, nausea, or dizziness?" Finally, in Task 5, K11 was able to detect a variety of NS-NNS differences in the stimuli. Although she could not come up with much metalanguage, she was able to provide detailed verbal explanation of the L2 phonological rules relevant to the task stimuli (Level 2). Her scores on Task 5 were 12 in segmentals and 3 in prosody. An interesting finding is that she expressed her knowledge of the vowel pair /i, ɪ/ and the vowel /æ/ in her entries. These entries are as follows:

- "Here in *cheese*, the American speaker pronounced it as 'cheese, but this one (Korean) said, 'chiz'" (T5Entry#241).
- "In *kids*, this part [pointing at the vowel] needs to be short, it has to be 'kids.' *For the kids*. No need to lengthen it. But the Korean speaker used a long vowel" (T5Entry#427).
- "In *slabs*, the 'a' part is full and long. That's how I hear that sound in the American's speech. But here (Korean), it sounds shorter" (T5Entry#240).

The above excerpts, as well as the data obtained through Task 2, indicate that K11 had verbalizable knowledge of these target sounds, but that knowledge was not readily available in her perception and production of L2. In the current analysis, K11's awareness of some L2 segmentals was found to be a bit limited, although she showed nativelike performance in using other components in the L2 phonological system.

Speaker K02. Speaker K02, age 31, immigrated to the United States with his family at age 13 and grew up in Virginia. He had been working at the U.S. Department of Treasury after graduating Cornell. His first contact with English was instruction that he received in primary school in Korea. He reported that he had taken some ESL lessons when he first moved to the United States. He also reported that Korean was the primarily spoken language in his family.

K02 earned 16.35 in reading and 20.17 in narration. As it was the case in the data of K11 earlier, K02's spontaneous speech was perceived to be more accented than was his reading speech. The total accentedness score of K02 was ranked 6th among the forty participants in this study (He was also a NL speaker in Section 4.2.3). The NS raters left a total of 107 comments on his performance in Phase 1. Despite his prolonged experience in L2 use, the problematic segmentals in his speech were not much different from those commonly commented in other speakers' data: /i, ɪ, ɛ, æ/ and /ð/ (Examples 122–124). Some comments pointed to the mixed use of the vowels /u, ʊ/ as well (Example 40). However, the quantity of these comments and the number of the raters who gave these comments were not large, when compared with the data of K16 or K30 (see Table 4.38). This suggests that the segmental deviations in his speech were negligible to the majority of the raters in this study. Meanwhile, the comments also revealed that he made a very

small number of lexical issues, being confused about primary versus secondary stress (Example 125). In addition, some raters commented that he had some unnatural rhythm and intonation in the reading speech, but such comments were rare in his narration data.

Example 122: Second vowel in "condition" is /i/ [E08_K02_N]
Example 123: Line 6: "trash" » "tresh" [E03_K02_N]
Example 124: *th* in "these" line 2 [E13_K02_R]
Example 125: lexical stress in line 1 "stella" [E12_K02_R]

The performance of K02 on the tasks in Phase 2 was overall above the average. In Task 1, he earned 37 in segmental accuracy and 40 in prosodic accuracy; that is, he was able to repeat most of the pseudoword stimuli with both segmental and prosodic accuracy. His *d'* scores (Task 2) were overall high, he had a slightly lower Hit Rate with the items involving the consonants /s, z/ (see Table 4.42). Additionally, he was very cautious while taking this task, as evident in many scribbles and marks that he left on the answer sheet. While this may be an illustration of his personality or test-taking strategy, it is clear that he was very alert at the time of taking this task.

Table 4.42

K02's Performance on the Selected Items of Task 2

Selected target	K02			All L2 speakers (Mean)		
	d'	Hit rate	FA rate	*d'*	Hit rate	FA rate
i	3.29	.95	.05	.69	.71	.45
æ	1.71	.90	.33	1.18	.90	.54
r	3.12	.97	.10	2.02	.81	.13
l	3.29	.95	.05	1.83	.80	.16
ð	3.12	.90	.33	.93	.70	.34
s	1.65	.95	.50	1.60	.89	.35
z	.85	.33	.10	1.15	.64	.21

In Task 3, he made two incorrect responses by choosing the secondarily-stressed syllables rather than the primarily-stressed ones. This, along with the analytical comments he received in Phase 1, seems to suggest that although he is able to produce near-nativelike L2 speech, his knowledge about L2 lexical stress is not quite targetlike yet. In Task 4, the only item that he struggled with had to do with meaning identification of an ambiguously written noun phrase based on phrase-level intonation: e.g., *hot dog*. Finally, in Task 5, he provided verbal explanation of a variety of NS-NNS pronunciation differences, both segmental and prosodic. His total score on Task 5 was 12 (8 on the segmentals and 4 on the prosody). While his metalanguage was not dense, he thoroughly described the pronunciation characteristics at phrasal level, then at word level, then at segmental level. While his sensitivity to /s, z/ was not nativelike as measured through Task 2, he attended to these sounds occurring in the stimuli of Task 5 (e.g., *six spoons, kids, cheese*), as expressed in some of his verbal response entries.

A noticeable strategy observed in his Task 5 data was repetition: every time he heard a stimulus and until he provided his verbal response (i.e., while he was processing the aural input and formulating his response), he continuously repeated the stimulus aloud. This strategy of repetition was indeed used by a few of the speaker participants in this study, but most attempts were made at lower volume—more like a quiet monologue or whisper. On the other hand, K02's repetitions were loud and unhesitant throughout the study. If such repetition had been one of his L2 learning strategies indeed, his extraordinary performance on Task 1 might be accounted for to some extent: it might have helped him develop skills such as perception and processing of the L2 auditory input, identification of the segmental and prosodic features in the input, retaining the

information in phonological memory, and retrieval and articulation of that input. Lastly, in his post-task reflection, K02 stated that he had a strong motivation to master L2 (English) sounds at age 15, due to a traumatic experience of confusing *bleeding* with *breathing*. Although motivation was not included as a variable in the current study, it is likely that his self-estimated, high level of motivation to learn L2 phonology has also influenced his L2 phonological development.

Summary. This section analyzed and reported the performance data of four selected L1 Korean-L2 English speakers on the speaking tasks in Phase 1 and the L2 phonological awareness tasks in Phase 2. A closer examination of the data revealed that the Korean speakers' awareness of some segmentals or prosodic features may not fully develop even after having many years of experience in L2 use and instruction, which may result in the trace of a Korean accent in their L2 speech.

Thus far, this chapter has presented the results obtained from the qualitative and quantitative analyses of the data collected in Phases 1 and 2 of the current study. In the following chapter, the findings are discussed in light of the research questions posed in Chapters I and III. The conclusion of the study follows.

V – DISCUSSION

The main purposes of the current dissertation study were twofold. First, it aimed to advance the understanding of what is considered as a foreign accent by investigating the phonological and phonetic variables underlying the listeners' perception of a foreign accent. Secondly, the study aimed to explore the relationship between foreign accentedness and L2 phonological awareness, more specifically whether L2 learners with different degrees of foreign accent vary in their awareness of L2 phonology. This chapter discusses the results from Phases 1 and 2 while re-addressing the research questions posed in the study.

5.1. Foreign Accent (RQ1)

Phase 1 of this study investigated the construct of foreign accent through quantitative and qualitative analyses of the accentedness rating data and answered the first research question of the study: *What constitutes a foreign accent?* Answering this question required analytical and metaphonological assessment of a foreign accent—a Korean accent of English in this study. The study thus employed raters experienced in English phonology and ESL/EFL teaching. These criteria were proposed for collecting extensive and metalinguistic descriptions of a Korean accent so that pronunciation problems in L2 speech could be meticulously scrutinized. The raters attended to a variety of nontargetlike phonological and phonetic features and phenomena detected in L2 speech.

Analyses of the written comments of the raters revealed a wide range of the segmental and prosodic deviations constituting a foreign accent, consistent with previous research that examined the phonological influences on accent (e.g., Anderson-Hsieh et al., 1992; Kang et al., 2010; Munro & Derwing, 1999; Trofimovich & Isaacs, 2012). Many of the reported deviations, particularly segmental substitutions and vowel insertions, could be attributed to the influence of the pre-existing perceptual and articulation settings of the speakers' L1, Korean. This result largely aligns with the previous research on L1 influence, namely *transfer*. The segmental deviations that were most reported in this study involved nontargetlike pronunciation of segmentals that do not exist in the phonemic inventory of the Korean language (e.g., consonants /θ, ð/ or a tense-lax vowel pair /i/-/ɪ/), empirically confirming the CAH (Lado, 1957). These sounds tended to be substituted with the sounds of their Korean counterparts having a similar place/manner of articulation, or alternate sounds shared by both the Korean and English phonemic inventories, as evidenced by the high rates of segmental substitution in Section 4.1. However, not all substitution patterns were explained by the CAH. For example, nearly half (48%) of the raters described that flap [ɾ] in some samples seemed to have been phonetically realized as [r], which exists in English but not in Korean. This pattern is explained by the SLM (Flege, 1995), which claims that L2 sounds equivalent to L1 sounds are prone to perceptual difficulty: to Korean speakers, the extent of perceptual dissimilarity between L1 (Korean) [ɾ] and L2 (English) [r] is greater than that between L1 [ɾ] and L2 [ɾ]. Hence, [ɾ] is more likely to be mapped to [r] in the Korean speakers' interlanguage.

High occurrences of vowel insertion and consonant deletion in the data suggest that L1 phonology poses a barrier to acquiring English phonotactics. Results of L1 transfer were also found in the prosodic patterns. Data suggest that Korean-accented speakers may have defaulted to the tonal patterns that are dominant in their L1 (e.g., the Korean rising patterns LHLH, HHLH; Kim & Kim, 2001) for all words. The overall pitch range was not as wide as that of native English speakers. This supports the previous findings that accentedness of L2 English speech deteriorates when the F0 range is narrower than that of the native English speech (Kang, 2010; Sereno et al., 2016). This outcome is not limited to only the pitch peak, contour, and range, but also includes the speech rhythm, or phrasal stress, and even tone sequences that mark the phrasal boundaries, which determine pauses of the utterance (Trofimovich & Baker, 2006). In short, many of the accent features reported in the raters' comments seem to be due to the influence of L1 phonology on the L2 speakers' interlanguage phonology, in line with previous research findings on L1 transfer (e.g., Hong et al., 2014; Riney et al., 2005; Tsukada et al., 2005).

However, not all reported accent features could be attributed to L1 influence. Substitutions of [o, ɑ, ɛ] for [ə], for example, seem to be suitably explained by orthographic influence, referred to as *spelling pronunciation* (Altenberg & Vago, 1983). Being a neutral vowel, [ə] is represented by diverse graphemes, and because the orthographic system in English is relatively more opaque than Korean (i.e., one letter (or a letter cluster) may correspond to more than one sound; e.g., English letter <a> as in *cat, game,* and *father*), L2 speakers with limited knowledge of vocabulary or vowel reduction in English would likely have pronounced the vowel as it is spelled, yielding nontargetlike

pronunciation. Moreover, spelling pronunciation phenomena were detected not only in the reading data but also in the narration data (e.g., pronouncing the /ɑ/ vowel spelled as <o> (as in *problem*) as it is spelled), suggesting that for some learners, the mental representation of an L2 phonological item is affected by its orthographic form and its phonetic realization depends primarily on orthographic information in lieu of phonological information (Piske et al., 2001).

Although it was not the main focus of this study to examine the elicitation-type effects on L2 accentedness, the study sought to investigate the accentedness judgments made in two sets of the L2 speech samples: reading and narration. The average accentedness scores assigned to the two sets had no statistically significant difference, and were consistent with the previous studies that reported no effect of elicitation-type on the accentedness ratings (e.g., Munro & Derwing, 1994). However, the raters made more comments in the reading samples than in the narration samples. The difference was particularly large in the comments that reported nontargetlike use of the L2 rhythm: the proportion of comments coded for the subcategory of phrasal stress (within the main category of prosodic issues) showed a 21.12% difference across sample sets. While taking the reading task, the L2 speakers might have been highly cautious and self-conscious and attempted to place focal stress on every word in the given text (not just on content words but even on function words), which may have further interrupted successful reduction and linking. The speakers' ability/inability to parse meaningful units or understanding of the text content may have caused them to assign phrasal breaks in locations that would be considered inappropriate in native English prosody. Furthermore, this may have affected the overall fluency perceived by the raters (Polyanskava et al.,

2017). In short, as shown in previous studies (e.g., Huang & Jun, 2011; Munro & Derwing, 1995), the speakers' reading ability could be a potential confounding factor. On the other hand, when the speakers produced extemporaneous utterances, relatively fewer rhythm-related problems were detected. A few of the participants reflected immediately after the tasks (although it was not a protocol of the current study) that they focused on enunciating the sounds in the reading text to achieve accuracy. This suggests that some L2 speakers consciously prioritize certain phonological features while reading aloud, although more systematic retrospective data are required to confirm this possibility.

Issues regarding the comments classified as "unspecified" need to be discussed as well. These comments appeared across the coded categories, but were substantial in the category of intonation issues (i.e., 513 out of 1,228 comments). This suggests that even for the raters with linguistics and teaching backgrounds, prosodic features were uneasy to pinpoint or analytically describe. Findings from the focused analyses of the two selected raters (Section 4.1.3) also reflect the fact that prosodic issues, as compared to segmental issues, are difficult to describe without reference to sophisticated terminology or using orthographic representations (cf. Trofimovich & Isaacs, 2012). The raters may have found it relatively easier to analyze and describe the segmental issues in the current data, considering the lower occurrence rate of the "unspecified" comments in the category of segmental substitutions (i.e., 338 out of 1,372 comments). The "unspecified" comments, although excluded in the analyses that aimed to characterize a Korean accent, remain an indication of the raters' awareness of (and attempt to describe) the accent features in the L2 speech samples.

It is also important to note that the accent features and the accentedness scores were determined by how listeners rated L2 pronunciation. The data obtained and analyzed in Phase 1 were based on the raters' perception and judgments made on numerical scales and in reports (i.e., analytical comments). Overall, the 23 raters showed high inter-rater reliability for accentedness ratings, a result which was satisfactory in running quantitative analyses of the rating data. On the other hand, the raters' comments were conveniently coded into the four main categories and their corresponding subcategories. Qualitative analyses of the comments revealed the differences among the raters that were not detected by the quantitative analyses. They revealed that the decision-making processes of the raters vary, as shown in previous research (e.g., Hayes-Harb & Hacking, 2015; Hsieh, 2011), and that the raters' judgments cannot be categorical. The fact that the stimuli for perception were auditory and not as concrete as visual stimuli seems to have made the rating process even more complex and dynamic. Individual differences in the raters' hearing, sensitivity, orientation, severity, and many other traits could not be controlled despite the study's attempt to employ its rater participants by setting specific criteria (thus minimizing the methodological issues reviewed in Section 3.1): native English speakers experienced in ESL/EFL teaching and training in linguistics. The raters appear to have paid differential attention to the various segmental and prosodic deviations reported in this study. Furthermore, as revealed by focused analyses, not one L2 pronunciation phenomenon was unanimously judged to be an accent feature. For instance, the [r]-colored coda /l/ in L2 samples, despite its acoustic evidence as shown in Section 4.1.4, did not hit a threshold level of many raters at which it begins to affect accentedness.

The results from Phase 1 not only shed light on the phonological features the raters attended to, but also highlighted *how* these raters processed and explained these features while approaching the construct of foreign accent (e.g., Raters E02's and E05's assessments of K16 in Section 4.1.3). The overall findings suggest that analyzing data via both quantitative and qualitative methods is essential, particularly when the data are multifaceted and composed of the subjective judgments of humans, who can hardly be strictly homogeneous. More discussion on the methodological limitations of the study appears in Chapter VI.

5.2. L2 Phonological Awareness (RQ2 and RQ3)

Phase 2 sought to answer the second and third research questions: *Is there a relationship between foreign accent and phonological awareness? If so, what is the nature of that relationship?* Phase 2 delved into L2 phonological awareness of the L2 speakers, and then explored how it is related to their L2 pronunciation performance assessed in Phase 1.

The L2 speaker participants first took tasks that were purported to tap into their awareness of the L2 (i.e., English) phonology, in different domains of pronunciation and at gradient levels of implicitness/explicitness. Each L2 speaker had a certain level of L2 phonological awareness in both segmental and prosodic domains. In the data of L2 segmental awareness, the speakers with relatively less L2 experience (Group 1) performed similar to, or sometimes even outperformed, those with more L2 experience (Groups 2 and 3) in verbalizing NS-NNS pronunciation differences, with or without the relevant metalanguage (Task 5). It is possible that for adult L2 speakers, the amount of

L2 experience is not a significant factor of their explicit knowledge in L2 phonology. Some understanding of L2 segmentals seems to have been established in the L2 speakers from early stages of L2 learning. This is consistent with Venkatagiri and Levis (2007), who found no relationship between explicit L2 phonological awareness and language experience (number of years of L2 study and months in the L2 country). In the current study, all the L2 speakers self-reported in the background questionnaire that they had received some formal instruction in L2, mostly in EFL settings. Thus, it is likely that, from the beginning stages, they might have been exposed to learning materials in which cross-linguistic (here, Korean-English) pronunciation differences were highlighted to a certain extent. By contrast, the group differences emerged in the tasks of relatively implicit L2 segmental knowledge (i.e., Tasks 1 and 2). This kind of knowledge might be an outcome of explicit knowledge automatized through massive practice (DeKeyser, 2003), as L2 experience is accumulated over a substantial period of time (e.g., 5+ years). After subsequent exposure to more exemplars and perhaps experiencing communication breakdowns (as in K02's anecdote: *breathing* vs. *bleeding*), the L2 speakers might have become able to use some distinctive segmental features so that they could approximate targetlike pronunciations.

The data of L2 prosodic awareness showed a different pattern from that of the segmental data. The L2 speakers, in general, manifested high degrees of prosodic awareness. The scores on Task 3 (perception and noticing of lexical stress) and Task 4 (making judgments based on the prosodic features in L2 auditory input) were in negatively skewed distributions, and the amount of L2 experience was not a significant factor in these performance scores. Similarly, in Task 1, many of the L2 speakers

repeated English pseudowords with a high rate of prosodic accuracy, showing a negatively skewed distribution. The results obtained from the prosodic awareness data suggest that implicit knowledge of L2 prosody, measured through perception and/or reproduction of prosodic features in the L2 auditory input (as in Tasks 1, 3, and 4) can be achieved to a certain level without prolonged L2 experience. As research on early language development shows, prosodic features are the most (perceptually) salient to, and learned first by, young babies (Menyuk, Liebergott, & Schultz, 2014). In a similar vein, certain L2 prosodic features might have been particularly salient to L2 learners and resulted in initial learning gains even after short-term exposure (e.g., de Bot & Malifert, 1982). In that case, some tasks of prosodic awareness might have not been sufficiently challenging to the L2 speakers, consequently yielding a ceiling effect. One should also note that Tasks 3 and 4 required the use of L2 prosody at a receptive level but did not require the speakers to produce L2 prosody. Task 1 did require both perception and (re-)production, but its task items were lexical items (pseudowords), the prosodic accuracy of which was simply determined by the placement of lexical stress or salient pitch accent, which could have also been achieved by the speakers with limited L2 experience. However, the significant difference between the speakers with the most L2 experience (Group 3) and those with the least L2 experience (Group 1; i.e., two years at most) indicates that L2 prosodic learning appears to be gradual and that nativelike, not just "acceptable," production of L2 prosody may require multiple years of use, as pointed out by previous research (Trofimovich & Baker, 2006).

On the other hand, it is not known whether there was a ceiling effect in the prosodic data of Task 5 because the data were entries tallied from open-ended responses.

In other words, there was no absolute, maximum possible score on Task 5. Although Group 3 showed a higher degree of metalinguistic knowledge of L2 prosody than the other groups (see Figure 4.13), the overall differences were not statistically significant, suggesting that regardless of the amount of their L2 experience, the speakers performed similarly in analyzing and verbalizing prosodic differences in NS-NNS speech samples, as was the case in the segmental data. The speakers were in general able to express their explicit knowledge of how a certain L1-accented L2 sound (in a word/phrase) was phonologically/phonetically different from the nativelike norm or how it should be pronounced to be nativelike. To summarize, the L2 speakers in this study overall had some degree of awareness of L2 prosody, as demonstrated by its use and verbal explanation, while there is a possibility of a ceiling effect in some tasks.

An important point to note here is that, as criticized by Ellis (2005, 2009), one can never be certain that the L2 speaker participants actually used the type of knowledge—explicit or implicit, though gradient—that each of the phonological awareness tasks intended to tap into. Ellis also points out that the ability to verbalize knowledge may in part be related to the amount of metalanguage the learner has. This may have been the case in the L2 speakers' performance on Task 5: the range of the metalinguistic terms used by the L2 speakers was not extensive. Therefore, the study focused on the verbal report entries made at Levels 2 (Understanding) and 3 (Understanding with metalanguage) in order to zero in on the explicit aspects of L2 phonological knowledge. Still, the study was not free from some limitations which are outlined in Chapter VI.

In exploring the relationship between L2 phonological awareness and accent, the study conducted a series of statistical analyses. First, the correlational analyses revealed

an inverse correlation between the nonnative speakers' L2 phonological awareness and accentedness, a quantifiable measure of a foreign accent. The finding extends the previous research that observed a relationship between the awareness of L2 sounds and L2 pronunciation accuracy (e.g., Baker & Trofimovich, 2006; Kivistö-de Souza, 2015; Venkatagiri & Levis, 2007). The relationship, however, was not straightforward in this study. In both reading and narration samples, the correlations between accentedness and performance on Task 1 (both segmental and prosodic accuracy) and between accentedness and Task 2 were stronger than those between accentedness and performance on most of the remaining phonological awareness measures. Performance on Task 5 prosody was an exception, in that correlations between that measure and accentedness were small and non-significant in both reading and narration. This indicates that not all measures of phonological awareness were related to accentedness.

Stepwise multiple regression analyses revealed that performance on only the segmental component of Task 1 and on Task 2 accounted for accentedness in reading (70.2% of the variance), and only the former was included in the regression model for narration (accounting for 64.3% of the variance). The measures related to L2 prosodic awareness as well as explicit, verbalizable knowledge of L2 phonology were excluded in the regression models. This indicates that accentedness may be more dependent on the speakers' awareness of L2 segmentals—particularly, how they perceive, notice, and analyze (and reproduce, if in Task 1) L2 segmentals. In other words, the findings from this study show a relationship between the non-verbalizable aspects of L2 phonological awareness (i.e., implicit knowledge) and the perceived degree of foreign accent in L2

speech. The processes underlying L2 pronunciation are implicit in the sense that L2 speakers can perform without attention to these processes.

Yet, the causality of this relationship needs to be established with caution. The general scrutiny about L2 speech is that perception (i.e., auditory and cognitive processing to connect sound to meaning) is the basis for establishing a mental representation of L2 sounds, and further on, for producing those sounds (e.g., Baker & Trofimovich, 2006; Flege, 1995). But then, one must also note a possibility that a learner's perceptual abilities in L2 may be higher than their actual production abilities (e.g., Saito, 2013). As noted by Kivistö-de Souza (2015), a clearer picture of the causality of the relationship between L2 phonological awareness and accent can be obtained in studies with a longitudinal design in which L2 learners' pronunciation and phonological awareness are periodically tested throughout their learning processes. While this was not feasible in the design and data of the current study, additional focused analyses were conducted instead to probe the relationship between foreign accent and L2 phonological awareness—particularly, segmental awareness—observed in the current data.

The focused analyses revealed that the nativelike (NL) speakers, occupying the bottom 25% in accentedness rating scores, either had both implicit and explicit knowledge of the target segmentals or only had implicit knowledge, but they were still able to make targetlike production (Cases A and E). The latter case (Case E) suggests that for these learners, explicit knowledge of the selected segmentals was not that essential. On the other hand, a large number of instances of Case F were observed among the heavily-accented (HA) speakers, who occupied the top 25% in accentedness rating scores, which indicates that their implicit knowledge was not always translated into the

actual pronunciation. Their knowledge may have been unstable or incomplete to be used to produce targetlike pronunciation. Case F may also be associated with the gap between perception and production, caused by limited articulation abilities. The HA group also showed a higher occurrence rate of Case H (i.e., neither implicit nor explicit knowledge of a segmental is developed, and accurate production cannot take place), which indicates that some of these L2 speakers had minimal or almost no awareness of the L2 segmentals selected for the analyses. As nothing could be translated into the production of those segmentals, these speakers (Speaker K30 can be one example) may have relied on the pre-existing (i.e., L1) phonological system in them and ended up sounding heavily accented.

The overall results suggest that for the selected segmentals, L2 speakers were able to have targetlike pronunciation even when they had only the implicit kind of knowledge of these sounds. Relatively lower rates of Case C in both the NL group and the HA group implicate that verbalizable knowledge of the L2 segmentals alone might not be sufficient in the nativelike production of these sounds. The findings from Speakers K02 and K11 can be an example here: both speakers showed high consciousness while attending to the L2 sounds in Task 5, but their performance differed in Task 2 and also in actual L2 speech. In the former, K11 was seen to have somewhat limited implicit knowledge of the sounds which she could surely attend to in Task 5 (e.g., the /i-ɪ/ contrast). Her production of these sounds in reading and narration was also deviant from nativelike production. On the other hand, K02's implicit knowledge of the L2 segmentals seemed to be stable and accurate, and only few raters detected a foreign accent in the segmentals in his speech. Based on these findings, explicit knowledge rather seems to play an ancillary role in the

development of L2 phonology, and it might have exerted the beneficial influences on perception by making relevant features in the input and the learners' linguistic competence salient (N. Ellis, 1994).

To what extent implicit knowledge interfaces with explicit knowledge is not clearly known from the current data, since the current study did not intend to test an interface position (see Ellis (2009), for a suggested design for an experimental study testing an interface position). Only speculations can be made as to how the learners' L2 phonological awareness and its explicit and implicit aspects have been developed. From the positions advocating the existence of interface (e.g., DeKeyser, 1998; R. Ellis, 1994), implicit knowledge of L2 phonology—especially of the NL speakers, all of whom had extensive L2 experience, or greater opportunity for practice—might be a proceduralized kind of what they learned as declarative rules, while transformation to the opposite direction is possible as well: knowledge that originated as implicit could be transformed into explicit through conscious analysis of output generated by means of implicit knowledge (Bialystok, 1994). Or, as Han and Finneran (2014) noted, L2 acquisition cannot be explained by a singular relation between explicit and implicit knowledge. Leaving this matter open for future research, the current study demonstrated that the implicit and explicit aspects of L2 phonological awareness are essential (although not equally) in L2 pronunciation and the degree to which L2 speakers are perceived to be targetlike or foreign-accented.

Furthermore, the focused analyses underscored the fact that individual sounds are unique and different in learnability. For instance, /ð/, which was found to be highly problematic in Phase 1, was indeed a tricky sound for both groups. Most of the HA

speakers pronounced it in a deviant way, although they demonstrated some sensitivity to that sound (Case F). On the other hand, the NL speakers were in general able to show targetlike performance pronouncing /ð/, although their awareness of the sound seemed limited (Cases E and G). Another example is /i/, as revealed by the focused analyses of individual participants. The vowel /i/ seemed to be particularly difficult to perceive, notice, and produce, even for speakers with prolonged L2 experience (e.g., Speaker K11: 11 years).

These findings generate a follow-up question for future research as to whether a certain segmental (or phonological feature) is different from the others in terms of learning difficulty or achievement rate. It might be that some sort of threshold might exist for individual components of a language (segmentals and prosody); that is, some features could be noticed and picked up relatively easily from regular input, while others that are perceptually less salient—either universally or as a result of L1 experience influencing and limiting perception—might require extra help (Kivistö-de Souza, 2015; Long, 2015).

A relevant issue in L2 research and pedagogy concerns what should be taught and how it should be taught in order to help L2 learners improve their pronunciation, particularly their foreign accent. Unfortunately, in the real world, existing accent reduction programs have received much criticism from researchers, while remaining popular among L2 learners, for spreading materials and curricula that have no empirical basis, such as reciting tongue twisters with a marshmallow between the lips (Derwing & Munro, 2009). However, if the primary goal of pronunciation instruction is improvement in accentedness, an adequate and necessary step in preparing for a lesson should be the characterization of learners' foreign accents, accompanied by a contrastive analysis of the

learners' L1 and L2 as well as the accent evaluation provided by reliable judges who can represent the learners' potential interlocutors.

In the quantitative data of this study, the conscious aspect of L2 phonological knowledge was seen not as essential as the implicit kind of knowledge in L2 speech performance. However, the findings of the study certainly do not underestimate the role of explicit phonological knowledge in the development of L2 pronunciation. From the views of the strong-interface and weak-interface positions, explicit knowledge can be a potential source of implicit knowledge (DeKeyser, 2003), or facilitate learning by, for instance, helping learners' perception (and noticing) of relevant features in the available input (N. Ellis, 1994). Also, explicit pronunciation instruction may be more feasible and realistic, especially for late L2/FL learners who have limited opportunities for implicit pronunciation learning. Research on implicit pronunciation instruction has been scarce, and rarely were the instructional (e.g., recasts) effects reported (Papachristou, 2014). On the other hand, a growing number of studies have examined the role of explicit pronunciation instruction on L2 perception and production (e.g., Kissling, 2013; Lambacher et al., 2005; Ramírez Verdugo, 2006; Saito, 2013). Whether pronunciation instruction—be it implicit or explicit—is facilitative to development of L2 phonological awareness *per se* remains to be elucidated.

VI – CONCLUSION

As noted in the beginning of this dissertation study, L2 pronunciation is a social phenomenon that is integral to both communication and the interlocutors' judgment about L2 speakers. Approaching a nativelike level in L2 phonology, however, is widely known to be difficult for late L2 learners. Despite much personal efforts or prolonged experience in L2 use, many learners retain traces of foreign accents in their L2 speech. The current study examined this construct of foreign accent, focusing on what constitutes, or underlies the listeners' perception of, a foreign accent, and it further probed the relationship between individual speakers' accentedness and their awareness of the phonological system of the L2. This chapter summarizes the main findings of the study and discusses its limitations and directions for future research.

6.1. Summary of Findings

To explore the relationship between foreign accent and phonological awareness, the study first characterized a foreign accent, the construct commonly understood as nonnative deviation from nativelike pronunciation norms. The raters in this study, all experienced in teaching English, identified various segmental and prosodic deviations contributing to a Korean accent in English. Notable consonantal features of Korean-accented English include: a tendency to substitute dental fricatives /θ, ð/ with stops (or alternatively, fricative /s/), mixed use of liquid consonants /l, r/, substitution of labials [p, b] for labiodentals /f, v/, affrication of the voiced palatal fricative, and insufficient

voicing of voiced obstruents. The most notable vowel issue was lack of distinction between tense and lax vowels (e.g., /i, ɪ/), among other vowel mismatches. Korean-accented utterances are also characterized by words with an added vowel after coda, diphthongs realized as separate vowels, inaccurately placed lexical and phrasal stress, and fluctuating intonation (pitch contours).

Most of these deviations could be predicted by the contrastive analysis of the learners' L1 and L2 phonologies, supporting the influence of L1 phonology on L2 pronunciation documented in previous research (e.g., Lado, 1957; Moulton, 1962). Furthermore, there were certain deviations that drew greater attention of the raters. In other words, some deviation types were overtly detectable (e.g., /i/ » /ɪ/), while others were negligible to many of the raters (e.g., /z/ » /ʤ/). The finding suggests that an examination of a foreign accent should not simply focus on pronunciation accuracy but rather investigate how the listeners perceive the pronunciation deviations.

The study also found some variance in quantity, quality, and content of the raters' commentary, which highlights that foreign accent is a complex construct for individual listeners to approach. Nevertheless, the high inter-rater reliability in the scalar assessments of accentedness supports previous research findings that listeners, in general, possess an internalized notion of foreign-accented speech and to what extent it differs from native speech (e.g., Trofimovich & Isaacs, 2012).

L2 learners' phonological awareness was measured using five tasks that purported to tap into the learners' awareness of L2 phonology at gradient levels of implicitness/explicitness. The relationship between foreign accent and phonological awareness was examined using quantitative analyses and then qualitative analyses of

individual learners' performance. The study found a relationship between foreign accent and phonological awareness, but the relationship was found to be intricate and nuanced.

The accentedness scores, a quantifiable measure of a foreign accent, were inversely correlated with most of the measures of L2 phonological awareness. Accentedness was most strongly correlated with the measure for implicit knowledge of L2 segmentals; in other words, the learners with greater implicit, non-verbalizable knowledge of L2 segmentals produced speech that sounded less foreign-accented to listeners. On the other hand, explicit knowledge of L2 segmentals and implicit knowledge of L2 prosody were moderately correlated with accentedness. The correlation between accentedness and explicit knowledge of L2 prosody was low.

Qualitative analyses further revealed that learners having limited or no awareness of particular L2 segmentals were likely to pronounce these sounds in a deviant (accented) manner. Most of these segmentals are not in the phonemic inventory of the learners' L1 (e.g., English /i, æ, ð/). On the other hand, learners with the implicit knowledge of these segmentals (i.e., those who can perceptually discern, for instance, the /i-ɪ/ contrast) were able to produce these sounds in a targetlike (or accent-free) manner. However, explicit, verbalizable knowledge of these segmentals alone was not sufficient for targetlike pronunciation. No such a relationship was found in the knowledge of L2 prosody. Even though the L2 learners could perceive the prosody in the L2 input and verbalize relevant prosodic features and rules, the prosody in their L2 speech was judged to have been influenced by the prosodic pattern of their L1 (e.g., the LHLH tone sequence).

The learners' L2 phonological awareness, as well as pronunciation, can develop as the amount of L2 experience is accumulated; however, certain phonological features

may remain persistently challenging even after extensive experience in the L2, resulting in a trace of a foreign accent. The role of phonological instruction on the development of L2 phonological awareness still remains to be investigated.

6.2. Limitations and Directions for Future Research

The current study had several methodological limitations that need to be taken into consideration when interpreting its findings. One of the notable limitations resides in the fact that the providers of the main data (i.e., accentedness ratings) were not fundamentally homogenous. The study posed certain background criteria upon recruiting the rater participants and focused on the ratings provided by native English speakers with backgrounds in English language teaching and linguistics. Examples of native and nonnative speech samples were provided during the norming. Nevertheless, the raters seemed to have had their subjective rating criteria and interpretations of the accent features in the samples. Qualitative analyses of the commentary and some selected raters' behaviors revealed that the raters' background characteristics at a micro level—for example, teaching ESL for 17 years (E02) versus one year (E05)—may result in differences in their tolerance for accents and consistency in accentedness assessments, the domains of pronunciation they attend to, and their capability in providing analytical assessments, to list a few. In short, a wide range of issues and factors commented by the raters may stem from individual differences. The inclusion of potential mediating variables (e.g., the raters' attitudes toward nonnative speech; Winke & Gass, 2013) between the raters' background characteristics and their rating behaviors may provide additional insights to the rating behaviors. While it is impossible to establish a sample

comprising perfectly homogeneous participants, due to the subjective nature of human judgments, deeper analyses of the individual participants' responses can mitigate this limitation when interpreting subjective judgment data.

Further complicating the issue of rater heterogeneity was the fact that the raters had diverse linguistic backgrounds, ranging from monolingual in English to having acquired an additional language other than their L1, English (while proficiency levels varied). Furthermore, two of the raters identified themselves as heritage speakers of Korean, and one rater self-reported as fluent in Korean. The majority of the raters were therefore multilingual speakers who, in Cook's (1999) view, are cognitively different from monolingual speakers. Such differences, be they subtle or not, may emerge in their perception and/or production of sounds of the languages they speak. For instance, L1 French-L2 English speakers were found to have a slightly longer VOT for /t/ in their L1 than French monolinguals (Flege, 1987). Likewise, in the present study, the individual raters' varying states of multicompetence may have affected how they conceptualized and operationalized the notion of "foreignness" or "nativelikeness" while assessing L2 speech samples. The judgments of the raters, therefore, are difficult to generalize, and this limitation must be noted in interpreting the findings of the study.

Another methodological limitation is that the study overly relied on the reports voluntarily provided by the participants. Two types of verbal reports were collected and analyzed in this study. In Phase 1, the raters typed on the rating form their analytical comments of accent features. In Phase 2, the L2 speakers verbally described and explained the NS-NNS pronunciation differences in Task 5, and their oral responses were transcribed. Such verbal report data, however, may not reflect the actual thoughts or

knowledge that the participants had in mind at the time of participation. It is also not known from the current data what decision-making processes were involved when the participants were providing their reports; for instance, the participants may have reported everything they perceived from the stimuli, or only pointed at certain (probably salient) ones. Whether the participants were equipped with some metalanguage, as discussed in Section 5.2, could have been an issue as well. Another serious problem lies in the fact that both reports used in this study did not have a limit in time or number of words, and the participants could choose to not provide any responses; therefore, an absence of reports cannot be equated with evidence of zero awareness (Ortega, 2009). Nevertheless, verbal reports have served as an important source of information about the participants' cognitive processes, information that would otherwise be unavailable (Bowles, 2008). Thus, careful collection, coding, and analyses of the verbal report data are most desired.

A related limitation pertains to the way in which the data were coded. For analysis purposes, the data in this study, comprising participants' complicated processing as well as linguistic information, were examined according to the coding schemes or guidelines adapted from previous studies. However, it was not always feasible to classify the data into a certain category over the other. In such cases, I had to make arbitrary decisions. For example, in Phase 1, instances of mixing the words *snake* and *snack* appearing in adjacent lines in the reading text were arbitrarily judged as a reading mistake (and were thus coded as "irrelevant to pronunciation"), rather than as an accent feature, "substitution of /æ/ for /ej/," since it was not feasible to trace the L2 speaker participants to obtain a reflective report on such instances. As such, the distinction between errors and mistakes might have not been consistent (cf. Corder, 1971), and this might have resulted

in an inconsistency in coding in the study. Furthermore, the categorical manner of analyses may have filtered what was embedded in the original data. This may have limited the understanding of the constructs in the study and their relationship.

Some other limitations lie in the methods employed for data collection. As stated in the previous chapters, the raters' participation was made online (i.e., remotely from the researcher) and self-paced. In other words, the rating tasks could not be administered by the researcher on-site, and the raters had freedom in terms of time, location, and duration of rating, and importantly, the amount of commentary that they voluntarily provided in describing the features of a foreign accent. This was different from most previous studies in which the researcher and rater(s) had to meet and complete tasks in one or multiple sittings (e.g., Munro & Derwing, 1994, 1995; Saito et al., 2017). The foremost goal of using an online rating system was to promote raters' participation in the study because the volume of the L2 speech data each rater was required to assess was large (80 samples). Indeed, many rater participants reported that the online system was convenient to use; however, there were drawbacks as well, including difficulties in engaging the raters in participation and familiarizing them with the rating system. It was also impossible to estimate the actual amount of time each rater spent on the rating or the number of replays he or she had throughout the rating tasks. Thus, the current study had to rely on the data (i.e., ratings on numerical scales and analytical commentary) *saved* by the rater participants. These data might not be sufficient in future studies that examine, for instance, specific rating behaviors. Use of technology cannot be discouraged, but finding a balancing point between modern and traditional data collection methods would be a crucial step in future research.

Yet another limitation of the study concerns the use of questionnaires as a means of data collection. The study relied on self-reported responses from questionnaires that were used to collect the participants' background information data. Thus, the information might not have been strictly objective. For instance, it was impossible to confirm whether the L2 speakers provided underestimated or overestimated values in reporting their L2 learning experiences (operationalized by means of LOR). Collection of such personal data had to rely on the participants' own estimation and judgment. Also, as mentioned in Section 4.2.4, it was impossible to know simply from the questionnaire responses the quality and quantity of L2 input that the speakers were exposed to (or greatly influenced by) in their L2 learning history (cf. Piske et al., 2001). For instance, L2 speakers with richer and more meaningful opportunities for practice, feedback, and interaction with native speakers would have been more likely to acquire a nativelike accent than those whose affiliation with the target language community was limited (Moyer, 2004, 2014).

Another example of self-reported information data is the raters' familiarity with a Korean accent. The data were collected using a multiple-choice question in a questionnaire for which the rater participants provided an estimate of the frequency of their contact with Korean speakers. But this method could not fully reflect their familiarity with the Korean language. Such a limitation could have been mitigated by administering a relatively objective accent identification task—such as the one in Huang (2013) in which native English speakers were asked to identify the origin of the speaker of the speech sample—to the rater participants before they were informed that they would be evaluating the L1 Korean speakers' accentedness.

Lastly, both of the speaking tasks in this study elicited monologic L2 speech, one type of which, reading, is often subject to criticism for not being naturalistic and lacking ecological validity (Levis & Moyer, 2014). A foreign accent, however, is an essential aspect of speech that may impose a barrier to L2 communication and interaction with interlocutors, as noted in the beginning of this study. In light of this sociolinguistic importance of the constructs of foreign accent and accentedness, a necessary and useful step of future research would be the exploration of accent features in authentic communication and varying interactional dynamics, and further on, probing whether or to what extent such features are detrimental to the success in L2 communication.

Many of the above limitations, particularly those pertaining to the subjective and uncontrollable nature of the verbal data or judgments, would probably be best mitigated by performing deeper, qualitative analyses of the data in conjunction with quantitative analyses. This way would enable us to find, for instance, even more detailed characteristics of a foreign accent that were not yet revealed in the categorical analyses in this study, and further on, help enhance the validity of the data. Much still remains underexplored about the rater-internal decision-making processes involved in the qualitative and quantitative assessments of foreign accents as well as the learner-internal processes involved in perceiving and producing L2 sounds. There is surely more to be revealed about the intricacies of foreign accent and phonological awareness in future research that would use improved research designs and testing instruments in a number of other L1-L2 combinations and research settings.

REFERENCES

Abercrombie, D. (1949). Teaching pronunciation. *English Language Teaching, 3*(5), 113-122.

Adams, M. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Alba-Salas, J. (2004). Voice onset time and foreign accent detection: Are L2 learners better than monolinguals? *Revista Alicantina de Estudios Ingleses, 17*, 9-30.

Altenberg, E. P. & Vago, R. M. (1983). Theoretical implications of an error analysis of second language phonology production. *Language Learning, 33*(4), 427-447.

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmental, prosody, and syllable structure. *Language Learning, 42*(4), 529-555.

Association for Language Awareness. (2012). *About*. Retrieved from http://www.languageawareness.org/web/ala/web/about.

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*(6), 575-589.

Baker, W. & Trofimovich, P. (2006). Perceptual paths to accurate production of L2 vowels: The role of individual differences. *IRAL – International Review of Applied Linguistics in Language Teaching, 44*(3), 231-250.

Baker-Smemoe, W. & Haslam, N. (2013). The effect of language learning aptitude, strategy use and learning context on L2 pronunciation learning. *Applied Linguistics, 34*(3), 435-456.

Bardovi-Harlig, K. (1986). *Pragmatic determinants of English sentence stress*. Bloomington, IN: Indiana University Linguistics Club.

Benson, P. & Lor, W. (1999). Conceptions of language and language learning. *System, 27*(4), 459-472.

Bialystok, E. (1994). Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition, 16*(2), 157-168.

Boersma, P. & Weenink, D. (2017). *Praat: Doing phonetics by computer* [Computer program]. Version 6.0.25, retrieved 13 February 2017 from http://www.praat.org/

Bohn, O-S. & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition, 14*(2), 131-158.

Bongaerts, T. (1999). *Nativelike-ness of pronunciation in naturalistic post-critical period second language acquisition*. Paper presented at the 12th World AILA Congress, Tokyo.

Bongaerts, T., Planken, B., & Schils, E. (1995). Can late learners attain a native accent in a foreign language? A test of the critical period hypothesis. In D. Singleton & Z. Lengyel (Eds.), *The age factor in second language acquisition* (pp. 30-50), Clevedon, UK: Multilingual Matters.

Bongaerts, T., van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition, 19*(4), 447-465.

Bowles, M. A. (2008). Task type and reactivity of verbal reports in SLA: A first look at a L2 task other than reading. *Studies of Second Language Acquisition, 30*(3), 359-387.

Brennan, E. M. & Brennan, J. S. (1981). Accent scaling and language attitudes: Reactions to Mexican American English speech. *Language and Speech, 24*(3), 207-221.

Calla McDermott, W. L. (1986). *The scalability of degrees of foreign accent*. Doctoral Dissertation. Cornell University.

Carroll, J. & Sapon, S. (1959). *Modern Language Aptitude Test (MLAT): Manual*. New York: The Psychological Corporation.

Cassady, J. C., Smith, L. S., & Huber, L. K. (2005). Enhancing validity in phonological awareness assessment through computer-supported testing. *Practical Assessment, Research & Evaluation, 10*(18), 1-13.

Chien, C-N., Kao, L-H., & Wei, L. (2008). The role of phonological awareness development in young Chinese EFL learners. *Language Awareness, 17*(4), 271-288.

Choi, Y. H. (2007). A study of students' attitude towards World Englishes and non-native English teachers. *English Teaching, 62*(4), 47-68.

Chomsky, N. (1976). *Reflections on language*. London: Temple Smith.

Chung, K. K. H., McBride-Chang, C., Cheung, H., & Wong, S. W. L. (2013). General auditory processing, speech perception and phonological awareness skills in Chinese-English biliteracy. *Journal of Research in Reading, 36*(2), 202-222.

Cook, V. (1999). Going beyond the native speaker in language teaching. *TESOL Quarterly, 33*(2), 185-209.

Corder, S. P. (1971). Idiosyncratic dialects and error analysis. *IRAL – International Review of Applied Linguistics in Language Teaching, 9*(2), 147-159.

Dalton, C. & Seidlhofer, B. (1994). *Pronunciation*. Oxford, UK: Oxford University Press.

Davies, A. (2003). *The native speaker: myth and reality*. Clevedon: Multilingual Matters.

de Bot, K. & Malifert, K. (1982). The teaching of intonation: Fundamental research and classroom applications. *TESOL Quarterly, 16*(1), 71-77.

DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. J. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42-63). New York, NY: Cambridge University Press.

DeKeyser, R. (2003). Implicit and explicit learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 313-348). Oxford: Blackwell.

Derwing, T. M. & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39*(3), 379-397.

Derwing, T. M. & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching, 42*(4), 476-490.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning, 48*(3), 393-410.

Derwing, T. M. & Rossiter, M. J. (2003). The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2 accented speech. *Applied Language Learning, 13*(1), 1-17.

Derwing, T. M., Thomson, R. I., Foote, J. A., & Munro, M. J. (2012). A longitudinal study of listening perception in adult learners of English: Implications for teachers. *The Canadian Modern Language Review, 68*(3), 247-266.

Donmall, G. (1985). *Language Awareness (NCLE Papers and Reports), 6*. London: CILT.

Dörnyei, Z. & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589-630). Oxford: Blackwell.

Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning, 27*(2), 315-330.

Elgort, I. (2011). Deliberate learning and vocabulary acquisition in a second language. *Language Learning, 61*(2), 367-413.

Ellis, N. (1994). Implicit and explicit learning: An overview. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 1-16). London: Academic Press.

Ellis, R. (1993). Second language acquisition and the structural syllabus. *TESOL Quarterly, 27*(1), 91-113.

Ellis, R. (1994). A theory of instructed second language acquisition. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 79-114). London: Academic Press.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition, 27*(2), 141-172.

Ellis, R. (2009). Measuring implicit and explicit knowledge of a second language. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing, and teaching* (pp. 31-64). Clevedon, UK: Channel View Publications.

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly, 39*(3), 399-423.

Flege, J. E. (1981). The phonological basis of foreign accent: A hypothesis. *TESOL Quarterly, 15*(4), 443-455.

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

Flege, J. E. (1995). Second-language speech learning: theory, findings, and problems. In Strange, W. (Ed.), *Speech perception and linguistic experience* (pp. 233-277). Timonium, MD: York Press.

Flege, J. E. (2008). Language contact in bilingualism: phonetic system interactions. In J. Cole & J. I. Hualde (Eds.), *Laboratory phonology 9* (pp. 353-382). Berlin: Mouton de Gruyter.

Flege, J. E., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Tsukada, K. (2006). Degree of foreign accent in English sentences produced by Korean children and adults. *Journal of Phonetics, 34*(2), 153-175.

Flege, J. E. & Fletcher, K. L. (1992). Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustic Society of America, 91*(1), 370-387.

Flege, J. E. & Hillenbrand, J. (1984). Limits on phonetic accuracy in foreign language speech production. *Journal of Acoustical Society of America, 76*(3), 708-721.

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

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language, 41*(1), 78-104.

Fromkin, V., Rodman, R., & Hyams, N. (2003). *An introduction to language* (7th ed.). Boston, MA: Wadsworth.

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory, 2*(2), 103-127.

Goldsmith, J. A. (1978). English as a tone language. *Communication and Cognition, 11*, 453-476.

Goswami, U. & Bryant, P. (1990). *Phonological skills and learning to read*. East Sussex: Erlbaum.

Goswami, U., Gerson, D., & Astruc, L. (2009). Amplitude envelope perception, phonology and prosodic sensitivity in children with developmental dyslexia. *Reading and Writing, 23*(8), 995-1019.

Gut, U. (2009). *Non-native speech: A corpus-based analysis of phonological and phonetic properties of L2 English and German*. Frankfrut: Peter Lang.

Hahn, I. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly, 38*(2), 201-223.

Han, Z-H. (2004). *Fossilization in adult second language acquisition.* Clevedon, UK: Multilingual Matters.

Han, Z-H. & Finneran, R. (2014). Re-engaging the interface debate: Strong, weak, none, or all? *International Journal of Applied Linguistics, 24*(3), 370-389.

Hayes-Harb, R. (2014). Acoustic-phonetic parameters in the perception of accent. In J. M. Levis, A. Moyer, & U. Jessner (Eds.), *Social dynamics in second language accent* (pp. 31-51). Berlin: Walter de Gruyter.

Hayes-Harb, R. & Hacking, J. (2015). Beyond rating data: What do listeners believe underlies their accentedness judgments? *Journal of Second Language Pronunciation, 1*(1), 43-64.

Hong, H., Kim, S., & Chung, M. (2014). A corpus-based analysis of English segments produced by Korean learners. *Journal of Phonetics, 46*, 52-67.

Hsieh, C-N. (2011). Rater effects in ITA testing: ESL teachers' versus American undergraduates' judgments of accentedness, comprehensibility, and oral proficiency. *Spaan Fellow Working Papers in Second or Foreign Language Assessment, 9*, 47-74.

Huang, B. H. (2013). The effects of accent familiarity and language teaching experience on raters' judgments of non-native speech. *System, 41*(3), 773-785.

Huang, B. H., Alegre, A., & Eisenberg, A. (2016). A cross-linguistic investigation of the effect of raters' accent familiarity on speaking assessment. *Language Assessment Quarterly, 13*(1), 25-41.

Huang, B. H. & Jun, S-A. (2011). The effect of age on the acquisition of second language prosody. *Language and Speech, 54*(3), 387-414.

Huang, B. H. & Jun, S-A. (2015). Age matters, and so may raters: Rater differences in the assessment of foreign accents. *Studies in Second Language Acquisition, 37*(4), 623-650.

Hummel, K. M. (2009). Aptitude, phonological memory, and second language proficiency in nonnative adult learners. *Applied Psycholinguistics, 30*(2), 225-249.

Ioup, G., Boustagui, E., El Tigi, M., & Moselle, M. (1994). Re-examining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition, 16*(1), 73-98.

Iverson, G. K. & Sohn, H-S. (1994). Liquid representation in Korean. In Y-K. Kim-Renaud (Ed.), *Theoretical issues in Korean linguistics* (pp. 1-19). Stanford, CA: CSLI.

Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation program syllabus for English as an international language. *Applied Linguistics, 23*(1), 83-103.

Jesney, K. (2004). *The use of global foreign accent rating in studies of L2 acquisition.* Calgary, AB: University of Calgary Language Research Centre Reports.

Jilka, M. (2000). Testing the contribution of prosody to the perception of foreign accent. *Proceedings of New Sounds (4th International Symposium on the Acquisition of Second Language Speech)*, (pp. 199-207). Amsterdam.

Juffs, A. (1990). Tone, syllable structure and interlanguage phonology: Chinese learners' stress errors. *IRAL – International Review of Applied Linguistics in Language Teaching, 28*(2), 99-117.

Jun, S-A. (1996). *The phonetics and phonology of Korean prosody: Intonational phonology and prosodic structure*. New York, NY: Garland.

Kabak, B. & Idsardi, W. J. (2007). Perceptual distortions in the adaptation of English consonant clusters: Syllable structure or consonantal contact constraints? *Language and Speech, 50*(1), 23-52.

Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System, 38*(2), 301-315.

Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal, 94*(4), 554-566.

Kang, Y. (2009). The role of phonological awareness in Korean elementary EFL learners' word reading. *English Teaching, 64*(2), 29-45.

Kashiwagi, A. & Snyder, M. (2010). Speech characteristics of Japanese speakers affecting American and Japanese listener evaluators. *Teachers College, Columbia University Working Papers in TESOL and Applied Linguistics, 10*, 1-14.

Kennedy, S., Blanchet, J., & Trofimovich, P. (2014). Learner pronunciation, awareness, and instruction in French as a second language. *Foreign Language Annals, 47*(1), 79-96.

Kennedy, S. & Trofimovich, P. (2010). Language awareness and second language pronunciation: A classroom study. *Language Awareness, 19*(3), 171-185.

Kim, J. S. & Kim, S-H. (2001). Remarks on Korean speakers' realization of English intonation: Focusing on declarative sentences. *Pan-Pacific Association of Applied Linguistics, 5*(2), 187-207.

Kissling, E. M. (2013). Teaching pronunciation: Is explicit phonetics instruction beneficial for FL learners? *The Modern Language Journal, 97*(3), 720-744.

Kivistö-de Souza, H. (2015). *Phonological awareness and pronunciation in a second language.* Doctoral dissertation. University of Barcelona.

Koda, K. (1998). The role of phonemic awareness in second language reading. *Second Language Research, 14*(2), 194-215.

Kopečková, R. (2018). Exploring metalinguistic awareness in L3 phonological acquisition: The case of young instructed learners of Spanish in Germany. *Language Awareness, 27*(1-2), 153-166.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.

Labov, W. (1995). Resyllabification. In R. Van Hout & F. Hinskens (Eds.), *Proceedings of the International Workshop on Language Variation and Linguistic Theory* (pp. 145-179). Nijmegen.

Ladefoged, P. (2005). *Vowels and consonants: An introduction to the sounds of languages.* (2nd ed.). Malden, MA: Blackwell.

Lado, R. (1957). *Linguistics across cultures*. Ann Arbor, MI: University of Michigan Press.

Lambacher, S. G., Martens, W. L., Kakehi, K., Marasinghe, C. A., & Molholt, G. (2005). The effects of identification and production of American English vowels by native speakers of Japanese. *Applied Psycholinguistics, 26*(2), 227-247.

Lee, B., Guion, S. G., & Harada T. (2006). Acoustic analysis of the production of unstressed English vowels by early and late Korean and Japanese bilinguals. *Studies in Second Language Acquisition, 28*(3), 487-513.

Lee, H-Y. (1996). *Korean phonetics*. Seoul: Taehaksa (in Korean).

Lenneberg, E. H. (1967). *Biological foundations of language*. New York, NY: Wiley.

Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 25-42). Ann Arbor, MI: University of Michigan Press.

Leow, R. P. (1997). Attention, awareness, and foreign language behavior. *Language Learning, 47*(3), 467-505.

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.

Levis, J. & Moyer, A. (2014). Introduction. In J. M. Levis, A. Moyer, & U. Jessner (Eds.), *Social dynamics in second language accent* (pp. 12-20). Berlin, DEU: Walter de Gruyter.

Loewen, S. (2014). *Introduction to instructed second language acquisition*. Florence, KY: Taylor and Francis.

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *The handbook of second language acquisition* (pp. 413-468). San Diego, CA: Academic Press.

Long, M. (2003). Stabilization and fossilization in interlanguage development. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 487-535). Oxford: Blackwell.

Long, M. (2015). *Second language acquisition and task-based language teaching.* Chichester: John Wiley and Sons, Inc.

MacKay, I., Flege, J.E., Piske, T., & Schirru, C. (2001). Category restructuring during second-language (L2) speech acquisition. *Journal of Acoustical Society of America, 110*(1), 516-528.

Magen, H. S. (1998). The perception of foreign-accented speech. *Journal of Phonetics, 26*(4), 381-400.

Major, R. C. (2001). *Foreign accent: The ontogeny and polygemy of second language phonology*. Mahwah, NJ: Erlbaum.

Major, R. C. (2007). Identifying a foreign accent in an unfamiliar language. *Studies in Second Language Acquisition, 29*(4), 539-556.

Major, R. C. (2008). Transfer in second language phonology: A review. In J. G. Hansen Edwards & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 63-94). Philadelphia, PA: John Benjamins.

McBride-Chang, C., Bialystok, E., Chong, K., & Yanping, L. (2004). Levels of phonological awareness in three cultures. *Journal of Experimental Child Psychology, 89*(2), 93-111.

Mennen, I. (2008). Phonological and phonetic influences in non-native intonation. In J. Trouvain & Gut, U. (Eds.), *Non-Native Prosody: Phonetic Description and Teaching Practice* (pp. 53-76). Berlin: Mouton de Gruyter.

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary acquisition. *British Journal of Developmental Psychology, 91*(1), 3-19.

Menyuk, P., Liebergott, J. W., & Schultz, M. C. (2014). *Early language development in full-term and premature infants*. New York, NY: Psychology Press.

Missaglia, F. (1999). Contrastive prosody in SLA—an empirical study with adult Italian learners of German. In J. Ohala, Y. Hsegawa, M. Ohala, D. Granville, & A. C. Bailey (Eds.), *Proceedings of the 14th International Congress of Phonetic Science* (vol. 1, pp. 551-554). Berkeley, CA: University of California.

Mora J. C., Rochdi, Y., & Kivistö-de Souza, H. (2014). Mimicking accented speech as L2 phonological awareness. *Language Awareness, 23*(1-2), 57-75.

Moulton, W. G. (1962). *The sounds of English and German*. Chicago, IL: University of Chicago Press.

Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition, 21*(1), 81-108.

Moyer, A. (2004). *Age, accent and experience in second language acquisition: An integrated approach to critical period inquiry*. Clevedon, UK: Multilingual Matters.

Moyer, A. (2014). The social nature of L2 pronunciation. In J. M. Levis, A. Moyer, & U. Jessner (Eds.), *Social dynamics in second language accent* (pp. 22-44). Berlin: Walter de Gruyter.

Moyer, A. (2015). *Foreign accent: The phenomenon of non-native speech*. New York, NY: Cambridge University Press.

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

Munro, M. J. (1995). Nonsegmental factors in foreign accent. *Studies in Second Language Acquisition, 17*(1), 17-34.

Munro, M. J. (2003). A primer on accent discrimination in the Canadian context. *TESL Canada Journal, 20*(2), 38-51.

Munro, M. J. (2008). Foreign accent and speech intelligibility. In J. G. Hansen Edwards & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 193-217). Philadelphia, PA: John Benjamins.

Munro, M. J. & Derwing, T. M. (1994). Evaluations of foreign accent in extemporaneous and read material. *Language Testing, 45*(1), 73-97.

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

Munro, M. J. & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 49*(s1), 285-310.

Munro, M. J. & Derwing, T. M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate. *Studies in Second Language Acquisition, 23*(4), 451-468.

Munro, M. J. & Derwing, T. M. (2011). The foundations of accent and intelligibility in pronunciation research. *Language Teaching, 44*(3), 316-327.

Munro, M. J., Derwing, T. M., & Flege, J. E. (1999). Canadians in Alabama: a perceptual study of dialect acquisition in adults. *Journal of Phonetics, 27*(4), 385-403.

Munro, M. J., Flege, J. E., & MacKay, I. R. A. (1996). The effects of age of second language learning on the production of English vowels. *Applied Psycholinguistics, 17*(3), 313-334.

Oller, J. W. & Ziahosseiny, S. M. (1970). The contrastive analysis hypothesis and spelling errors. *Language Learning, 20*(2), 153-189.

Ortega, L. (2009). *Understanding second language acquisition*. London: Hodder Education.

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research, 5*(3), 261-283.

Papachristou, V. (2014). *The effective of explicit or implicit pronunciation teaching to Greek learners of English: the case of acquisition of English vowels*. Doctoral dissertation. Aristotle University of Thessaloniki.

Pennington, M. C. & Richards, J. C. (1986). Pronunciation revisited. *TESOL Quarterly, 20*(2), 207-225.

Piasecka, L. (2011). The role of phonemic awareness in the development of L1 and L2 reading. In J. Arabski & A. Wojtaszek (Eds.), *Acquisition of L2 phonology* (pp. 108-120). Bristol, UK: Channel View Publications.

Pickering, L. & Baker, A. (2014). Suprasegmental measures of accentedness. In J. M. Levis, A. Moyer, & U. Jessner (Eds.), *Social dynamics in second language accent* (pp. 75-94). Berlin: Walter de Gruyter.

Pimsleur, P. (1966). *Pimsleur Language Aptitude Battery*. The Psychological Corp.

Piske, T. (2008). Phonetic awareness, phonetic sensitivity and the second language learner. In J. Cenoz & N. H. Hornberger (Eds.), *Encyclopedia of language and education*, (2nd ed.), *Vol. 6: Knowledge about language* (pp. 155-166). New York, NY: Springer.

Piske, T., MacKay, I., & Flege, J. E. (2001). Factors affecting degree of foreign accents in an L2: A review. *Journal of Phonetics, 29*(2), 191-215.

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.

Posner, M. I. & Peterson, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*, 25-42.

Pullen, P. C. & Justice, L. M. (2003). Enhancing phonological awareness, print awareness, and oral language skills in preschool children. *Intervention in School and Clinic, 39*(2), 87-98.

Purcell, E. T. & Suter, R. W. (1980). Predictors of pronunciation accuracy: a reexamination. *Language Learning, 30*(2), 271-287.

Ramírez Verdugo, D. (2006). A study of intonation awareness and learning in non-native speakers of English. *Language Awareness, 15*(3), 141-159.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. *Quarterly Journal of Experimental Psychology, 55A*, 1339-1362.

Riney, T. J. & Flege, J. E. (1998). Changes over time in global foreign accent and liquid identifiability and accuracy. *Studies in Second Language Acquisition, 20*(2), 213-243.

Riney, T. J., Takada, M., & Ota, M. (2000). Segmentals and global foreign accent: the Japanese flap in EFL. *TESOL Quarterly, 34*(4), 717-737.

Riney, T. J. & Takagi, N. (1999). Global foreign accent and voice onset time among Japanese EFL speakers. *Language Learning, 49*(2), 275-302.

Riney, T. J., Takagi, N., & Inutsuka, K. (2005). Phonetic parameters and perceptual judgments of accent in English by American and Japanese listeners. *TESOL Quarterly, 39*(3), 441-466.

Robertson, C. & Salter, W. (2007). *The Phonological Awareness Test 2 (PAT 2)*. East Moline, IL: Linguisystems.

Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning, 45*(2), 283-331.

Robinson, P. (2001). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287-318). Cambridge: Cambridge University Press.

Robinson, P. (2005). Attention and Memory during SLA. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 631-678). Oxford: Blackwell.

Saito, K. (2011). Identifying problematic segmental features to acquire comprehensible pronunciation in EFL settings: The case of Japanese learners of English. *RELC Journal, 42*(3), 363-378.

Saito, K. (2013). The acquisitional value of recasts in instructed second language speech learning: teaching the perception and production of English /ɹ/ to adult Japanese learners. *Language Learning, 63*(3), 499-529.

Saito, K., Trofimovich, P., & Isaacs, T. (2017). Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics, 38*(4), 439-462.

Schachter, S., Christenfeld, N., Ravina, B., & Bilous. F. (1991). Speech disfluency and the structure of knowledge. *Journal of Personality and Social Psychology, 60*(3), 362-367.

Schertz, J., Cho, T., Lotto, A., & Warner, N. (2015). Individual differences in phonetic cue use in production and perception of non-native sound contrast. *Journal of Phonetics, 52*, 183-204.

Schiff, R. & Calif, S. (2007). Role of phonological and morphological awareness in L2 oral reading. *Language Learning, 57*(2), 271-298.

Schmidt, A. M. (1996). Cross-language identification of consonants. Part 1. Korean perception of English. *Journal of the Acoustical Society of America, 99*(5), 3201-3211.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.

Schmidt, R. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 165-209). London: Academic Press.

Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker (Eds.), *Proceedings of CLaSIC 2010, Singapore* (pp. 721-737). Singapore: National University of Singapore, Centre for Language Studies.

Schneiderman, E. I. & Desmarais, C. (1988). The talented language learner: some preliminary findings. *Second Language Research, 4*(2), 91-109.

Scovel, T. (1969). Foreign accents, language acquisition, and cerebral dominance. *Language Learning, 19*(3-4), 245-253.

Scovel, T. (2000). A critical review of the critical period research. *Annual Review of Applied Linguistics, 20*, 213-223.

Selinker, L. (1972). Interlanguage. *IRAL – International Review of Applied Linguistics in Language Teaching, 10*, 209-232.

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.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington D.C.: National Academy Press.

Snow, C. E. & Hoefnagel-Höhle, M. (1977). Age differences in the pronunciation of foreign sounds. *Language and Speech, 20*(4), 357-365.

Southwood, M. H. & Flege, J. E. (1999). Scaling foreign accent: direct magnitude estimation versus interval scaling. *Clinical Linguistics & Phonetics, 13*, 335-349.

Stockwell, R. & Bowen, J. (1965). *The sounds of English and Spanish*. Chicago, IL: University of Chicago Press.

Swanson, H. L., Rosston, K., Gerber, M., & Solari, E. (2008). Influence of oral language and phonological awareness on children's bilingual reading. *Journal of School Psychology, 46*(4), 413-429.

Thompson, I. (1991). Foreign accents revisited: The English pronunciation of Russian immigrants. *Language Learning, 41*(2), 177-204.

Tomlin, R. & Villa, V. (1994). Attention in cognitive science and SLA. *Studies in Second Language Acquisition, 16*(2), 183-203.

Trofimovich, P. & Baker, W. (2006). Learning second language suprasegmentals: Effects of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition, 28*(1)*,* 1-30.

Trofimovich, P. & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language & Cognition, 15*(4), 905-916.

Truscott, J. (1998). Noticing in second language acquisition: a critical review. *Second Language Research, 14*(2), 103-135.

Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Flege, J. (2005). A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics, 33*(3), 263-290.

Venkatagiri, H. S., & Levis, J. M. (2007). Phonological awareness and speech comprehensibility: An exploratory study. *Language Awareness, 16*(4), 263-277.

Verhoeven, L. (2007). Early bilingualism, language transfer, and phonological awareness. *Applied Psycholinguistics, 28*(3), 425-439.

Wagner, R., Torgesen, J., Rashotte, C., & Pearson, N. A. (2013). *Comprehensive test of phonological processing, second edition (CTOPP-2)*. Austin, TX: Pro-Ed.

Wardhaugh, R. (1970). The contrastive analysis hypothesis. *TESOL Quarterly, 4*(2), 123-130.

Weinberger, S. (2015). *Speech Accent Archive*. George Mason University. Retrieved from http://accent.gmu.edu.

Wennerstrom, A. (1994). Intonational meaning in English discourse: a study of non-native speakers. *Applied Linguistics, 15*(4), 399-420.

Whalley, K. & Hansen, J. (2006). The role of prosodic sensitivity in children's reading development. *Journal of Research in Reading, 29*(3), 288-303.

Whitman, R. L. & Jackson, K. L. (1972). The unpredictability of contrastive analysis. *Language Learning, 22*(1), 29-41.

Williams, G. J., & Wood, C. (2012). Sensitivity to the acoustic correlates of lexical stress and their relationship to reading in skilled readers. *Advances in Cognitive Psychology, 8*(4), 267-280.

Winke, P. & Gass, S. (2013). The influence of second language experience and accent familiarity on oral proficiency rating: A qualitative investigation. *TESOL Quarterly, 47*(4), 762-789.

Winters, S. & O'Brien, M. G. (2013). Perceived accentedness and intelligibility: the relative contributions of F0 and duration. *Speech Communication, 55*(3), 486-507.

Wrembel, M. (2011). Metaphonetic awareness in the production of speech. In M. Pawlak, E. Waniek-Klimczak, & J. Majer (Eds.). *Speaking and instructed foreign language acquisition* (pp. 169-182). Clevedon, UK: Multilingual Matters.

Wrembel, M. (2015). Metaphonological awareness in multilinguals: A case of L3 Polish. *Language Awareness, 24*(1), 60-83.

Yeung, S. S. & Chan, C. K. K. (2012). Phonological awareness and oral language proficiency in learning to read English among Chinese kindergarten children in Hong Kong. *British Journal of Educational Psychology, 83*(4), 550-568.

Yeung, S. S., Siegel, L. S., & Chan, C. K. K. (2013). Effects of a phonological awareness program on English reading and spelling among Hong Kong Chinese ESL children. *Reading and Writing, 26*(5), 681-704.

Yuan, J., Jiang, Y., & Song, Z. (2010). Perception of foreign accent in spontaneous L2 English speech. *Proceedings of Speech Prosody 2010 (ISCA Archive)*, 100884:1-4.

Yule, G. (2014). *The study of language* (5th ed.). New York, NY: Cambridge University Press.

Appendix A

Speaking Tasks

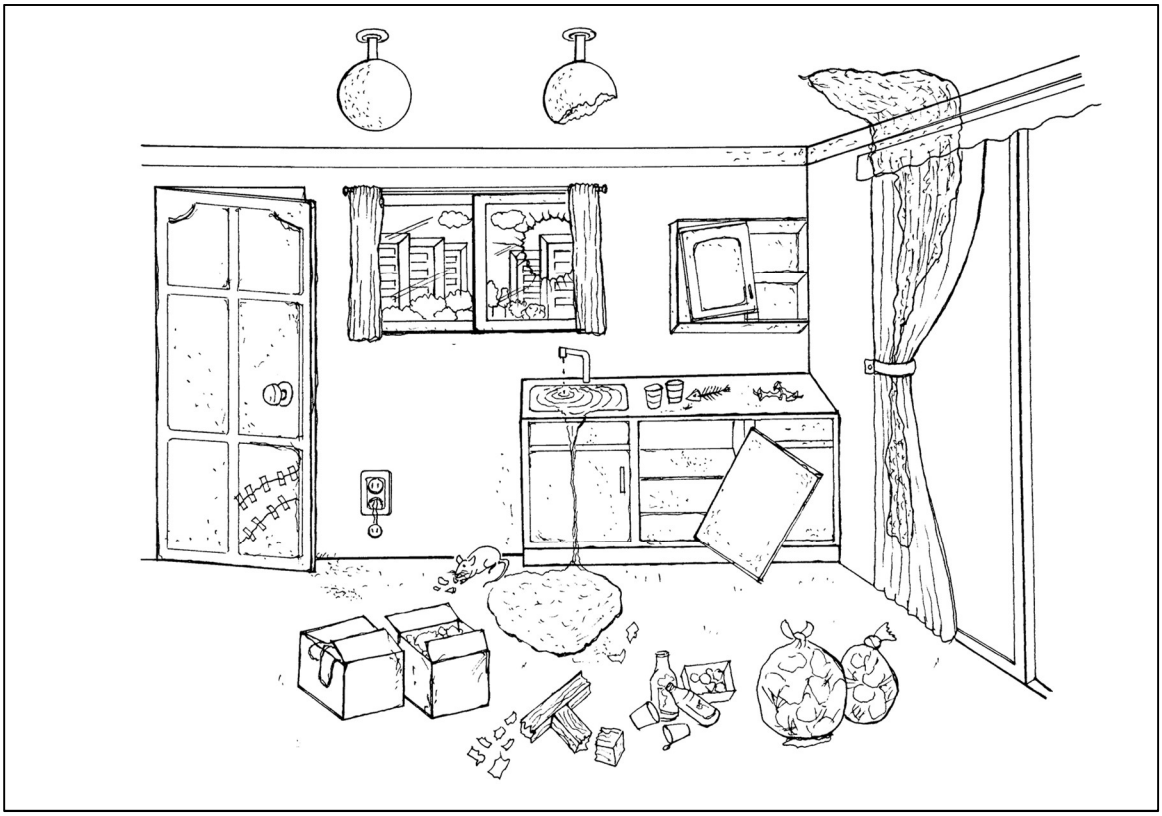
Reading

Read aloud the following passage. You may prepare for your response for 30 seconds. Recording will start after 30 seconds.

Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.

Narration

Situation: You have just moved into a new apartment and found several problems as you see below. You need to leave a voice message for the landlord complaining and describing the problems shown in the picture. Recording will start after 30 seconds.



Appendix B

L2 Phonological Awareness Tasks

Task 1

Instruction: Listen to the sentence carefully. After a short question and a signal beep, repeat the word that you heard. Imitate as closely to the original as possible.
Your response will be audio-recorded.

Example. This is *bannifer*. What is this?

1) This is *benevolate*. What is this?
2) This is *cantileen*. What is this?
3) This is *contramponist*. What is this?
4) This is *detailoring*. What is this?
5) This is *elenator*. What is this?
6) This is *frescovent*. What is this?
7) This is *horozone*. What is this?
8) This is *humberoid*. What is this?
9) This is *lannery*. What is this?
10) This is *lapidoscope*. What is this?
11) This is *lypocrisy*. What is this?
12) This is *mercusy*. What is this?
13) This is *yestrogeny*. What is this?
14) This is *reservory*. What is this?
15) This is *scudamore*. What is this?
16) This is *sinterity*. What is this?
17) This is *teometry*. What is this?
18) This is *thickery*. What is this?
19) This is *utilisk*. What is this?
20) This is *voltularity*. What is this?

Task 2

Instruction: For each round, listen carefully to the target sound followed by four individual words. As you listen, circle the letter for the word that has the target sound.
You may listen up to two times.
Use the answer sheet.

Stimulus set	Target phoneme and example (Only the examples in bold and italic were displayed on the answer sheet)	Pseudowords (Only the bold letters were displayed on the answer sheet. Stimuli were not displayed orthographically)			
Example	/p/ as in ***pop***	**A**. *speem*	**B**. *graff*	**C**. *purf*	**D**. *chisp*
1)	/b/ as in ***bob***	**A**. *bligg*	**B**. *scob*	**C**. *frouse*	**D**. *spraw*
2)	/u/ as in ***pool***	**A**. *cluft*	**B**. *mool*	**C**. *sploon*	**D**. *bru*
3)	/r/ as in ***rear***	**A**. *runched*	**B**. *lafe*	**C**. *prants*	**D**. *scier*
4)	/ð/ as in ***the***	**A**. *waint*	**B**. *doist*	**C**. *thout*	**D**. *droch*
5)	/ʤ/ as in ***judge***	**A**. *krodge*	**B**. *jowns*	**C**. *zaints*	**D**. *shround*
6)	/g/ as in ***gig***	**A** *bleck*	**B**. *garse*	**C**. *croots*	**D**. *jang*
7)	/i/ as in ***beat***	**A**. *screase*	**B**. *snibs*	**C**. *glink*	**D**. *gheam*
8)	/t/ as in ***tote***	**A**. *blit*	**B**. *stule*	**C**. *snarth*	**D**. *dorce*
9)	/s/ as in ***socks***	**A**. *thorve*	**B**. *snace*	**C**. *fless*	**D**. *shrodd*
10)	/z/ as in ***zoos***	**A**. *dwaze*	**B**. *slafe*	**C**. *buzed*	**D**. *jave*
11)	/l/ as in ***little***	**A**. *pleesh*	**B**. *kilv*	**C**. *rynt*	**D**. *croag*
12)	/ʃ/ as in ***shush***	**A**. *shaund*	**B**. *kleeche*	**C**. *gaced*	**D**. *steesh*
13)	/v/ as in ***valve***	**A**. *beadge*	**B**. *freap*	**C**. *blyced*	**D**. *vulps*
14)	/θ/ as in ***thirteenth***	**A**. *skeede*	**B**. *tharnt*	**C**. *thobs*	**D**. *thrail*
15)	/o/ as in ***so***	**A**. *stose*	**B**. *dwarst*	**C**. *scoaps*	**D**. *plost*
16)	/f/ as in ***photograph***	**A**. *feens*	**B**. *blurth*	**C**. *braff*	**D**. *heft*
17)	/æ/ as in ***can***	**A**. *ganks*	**B**. *redge*	**C**. *sness*	**D**. *vake*
18)	/ʧ/ as in ***church***	**A**. *choste*	**B**. *geach*	**C**. *kneedge*	**D**. *frash*
19)	/j/ as in ***year***	**A**. *yisk*	**B**. *eelt*	**C**. *ilv*	**D**. *yindge*
20)	/w/ as in ***wood***	**A**. *oolt*	**B**. *woosh*	**C**. *woost*	**D**. *woovs*

Task 3

Instruction: Listen carefully to the words and circle the box of the syllable that is most stressed within each word.
You may listen up to two times.
Use the answer sheet.

Example	*oligation*	o	li	ga	tion
1)	*almanical*	al	ma	ni	cal
2)	*anasytic*	a	na	sy	tic
3)	*bastionate*	bas	tio	nate	
4)	*combustulate*	com	bus	tu	late
5)	*condimented*	con	di	men	ted
6)	*draconite*	dra	co	nite	
7)	*levosnoot*	le	vos	noot	

8)	*litholect*	li	tho	lect		
9)	*misabrogate*	mis	a	bro	gate	
10)	*pristoractional*	pris	to	rac	tio	nal
11)	*recenticle*	re	cen	ti	cle	
12)	*repastafique*	re	pas	ta	fique	
13)	*retregradient*	re	tre	gra	di	ent
14)	*sanseveek*	san	se	veek		
15)	*sapirical*	sa	pi	ri	cal	
16)	*scurrilize*	scur	ri	lize		
17)	*semaphrodite*	se	ma	phro	dite	
18)	*sepretennial*	se	pre	ten	ni	al
19)	*stimulcrate*	sti	mul	crate		
20)	*trumpetine*	trum	pe	tine		

Task 4

Instruction: Listen carefully to the melody of each sentence or question and choose the appropriate response or meaning.
You may replay the speech up to two times.
Use the answer sheet.

	Displayed text	Aural stimulus	Choose one:	
e.g.,	What is he eating	Will hear two possible intonations	Response: Yes, he is.	Response: Chicken
e.g.,	Did you use a desktop a laptop or an iPad	Will hear two possible intonations	Response: Yes, I did.	Response: My iPad
1)	What is he cooking	What is he cooking?	Response: Not really.	Response: Beef stew
2)	I like chocolate chips and oatmeal	I like chocolate, chips, and oatmeal.	[chocolate] [chips] [oatmeal]	[chocolate chips] [oatmeal]
3)	Who's not here John	Who's not here? John?	(She asks John)	(Maybe John is not here)
4)	What is he washing	What, is he washing?	Response: Maybe	Response: His clothes
5)	There is a hot dog on the table	There is a hot "dog" on the table.	(Food)	(The dog is hot)
6)	What is he driving	What is he driving?	Response: No, his mother is.	Response: His mother's car
7)	Did you have some coffee tea or milk	Did you have some coffee, tea or milk? (Choose one)	Response: No, I didn't.	Response: Just coffee.
8)	What is he reading	What is he reading?	Response: Yes, he is.	Response: A new magazine

9)	What's in the tea honey	What's in the tea, honey?	(She calls the other person honey)	(She asks if there is honey in her tea)
10)	Do you have a pain nausea or dizziness	Do you have a pain, nausea, or dizziness? (Yes/no)	Response: I think so.	Response: Just nausea.
11)	What is he studying	What, is he studying?	Response: Yes!	Response: Science
12)	It could be a black hole	It could be a black "hole."	(Black hole as in the space)	(The hole is black)
13)	I had cheese bread and ham	I had cheese, bread, and ham.	[cheese] [bread] [ham]	[cheese bread] [ham]
14)	What is he moving	What, is he moving?	Response: Yes, to Canada.	Response: Just a box
15)	I will see you tomorrow	I will see you tomorrow?	(This is a question)	(This is not a question)

Task 5

Instruction: In each set, you will hear two speakers reading the same text. After listening to both, explain how the two speeches are different in terms of pronunciation.
You may take notes on the transcript.
You may listen multiple times.
You may respond in Korean.
Your free responses will be audio-recorded.

Stimulus set	Displayed transcript	Native English	Nonnative English
Example	Please call Stella	English 49 from the archive	Korean 12 from the archive
1)	Six spoons of fresh snow peas		
2)	Five thick slabs of blue cheese		
3)	And a big toy frog for the kids		
4)	Wednesday at the train station		

Appendix C

Background Questionnaires

For L1 Korean-L2 English Speakers:

Name: _____
Phone Number: _____
Email: _____
Hometown (not necessarily birthplace) in Korea: _____
Age: _____
Most recent education (check one):
high school: __
undergraduate: __
graduate: __
Do you speak any language(s) other than Korean and English?
-_____ at (circle one) advanced/intermediate/basic level

About your English learning history:
Back in Korea
How many years have you studied English? _____
Where did you receive primary English instruction? List all:
_____
Purpose(s) of learning English:
_____

Have you had any contacts with a native speaker of English regularly? If yes,
-Where: _____ Duration: _____ Origin of the NS: _____
-Where: _____ Duration: _____ Origin of the NS: _____

Have you taken any English speaking or pronunciation course? If yes,
-Where: _____ Duration: _____ Focus: _____
-Where: _____ Duration: _____ Focus: _____

Have you taken any English proficiency test? If yes,
-Name of test: _____ Score: _____ Year: _____

Here in the US
When did you move to the US, and why? _____
Purpose(s) of learning English:
_____

Have you taken any English classes (including the CLP)? If yes,
-Where: _____ Duration: _____ Focus: _____
-Where: _____ Duration: _____ Focus: _____

For NS Raters:

Name: _____
Phone Number: _____
Email: _____
Hometown (not necessarily birthplace): _____
Most recent education (check one):
high school: ___
undergraduate: ___
graduate: ___
Do you speak any language(s) other than English?
-_____ at (circle one) advanced/intermediate/basic level
-_____ at (circle one) advanced/intermediate/basic level

Describe your English (ESL and/or EFL) teaching experience. Please indicate the target audience, location, duration, and other notable facts:
_____

Describe your experience in phonetics and phonology (e.g., courses you have taken, trainings you have received). Please indicate the course name(s) (if possible), duration, and other notable facts:
_____

Have you been in contact with a native speaker of Korean (e.g., friend, student, or neighbor)? If yes, how regularly?
Yes, almost every day: ___
Yes, 3-5 times a week: ___
Yes, 1-2 times a week: ___
Yes, 1-2 times a month: ___
Yes, but not as frequently as above: ___
No: ___