# ENGLISH SYLLABLE CONFUSION AND IMITATION
# IN KOREAN BILINGUAL AND MONOLINGUAL CHILDREN AND ADULTS

BY

SOOK-YOUN YOON

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Speech and Hearing Science
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Doctoral Committee:

Associate Professor Cynthia J. Johnson, Chair
Associate Professor Jont B. Allen
Dr. David M. Gooler
Emeritus Professor Chin-Woo Kim
Emeritus Professor David P. Kuehn

Abstract

In the present study, Korean-English bilingual (KEB) and Korean monolingual (KM) children, between the ages of 8 and 13 years, and KEB adults, ages 18 and older, were examined with one speech perception task, called the Nonsense Syllable Confusion Matrix (NSCM) task (Allen, 2005), and two production tasks, called the Nonsense Syllable Imitation Task (NSIT) and the Nonword Repetition Task (NRT; Dollaghan & Campbell, 1998). The present study examined (a) which English sounds on the NSCM task were identified less well, presumably due to interference from Korean phonology, in bilinguals learning English as a second language (L2) and in monolinguals learning English as a foreign language (FL); (b) which English phonemes on the NSIT were more challenging for bilinguals and monolinguals to produce; (c) whether perception on the NSCM task is related to production on the NSIT, or phonological awareness, as measured by the NRT; and (d) whether perception and production differ in three age-language status groups (i.e., KEB children, KEB adults, and KM children) and in three proficiency subgroups of KEB children (i.e., English-dominant, ED; balanced, BAL; and Korean-dominant, KD).

In order to determine English proficiency in each group, language samples were extensively and rigorously analyzed, using software, called Systematic Analysis of Language Transcripts (SALT). Length of samples in complete and intelligible utterances, number of different and total words (NDW and NTW, respectively), speech rate in words per minute (WPM), and number of grammatical errors, mazes, and abandoned utterances were measured and compared among the three initial groups and the three proficiency subgroups. Results of the language sample analysis (LSA) showed significant group differences only between the KEBs

and the KM children, but not between the KEB children and adults. Nonetheless, compared to normative means (from a sample length- and age-matched database provided by SALT), the KEB adult group and the KD subgroup produced English at significantly slower speech rates than expected for monolingual, English-speaking counterparts.

Two existing models of bilingual speech perception and production—the Speech Learning Model or SLM (Flege, 1987, 1992) and the Perceptual Assimilation Model or PAM (Best, McRoberts, & Sithole, 1988; Best, McRoberts, & Goodell, 2001)—were considered to see if they could account for the perceptual and production patterns evident in the present study. The selected English sounds for stimuli in the NSCM task and the NSIT were 10 consonants, /p, b, k, g, f, θ, s, z, ʧ, ʤ/, and 3 vowels /ɪ, ɛ, æ/, which were used to create 30 nonsense syllables in a consonant-vowel structure. Based on phonetic or phonemic differences between the two languages, English sounds were categorized either as familiar sounds—namely, English sounds that are similar, but not identical, to L1 Korean, including /p, k, s, ʧ, ɛ/—or unfamiliar sounds—namely, English sounds that are new to L1, including /b, g, f, θ, z, ʤ, ɪ, æ/.

The results of the NSCM task showed that (a) consonants were perceived correctly more often than vowels, (b) familiar sounds were perceived correctly more often than unfamiliar ones, and (c) familiar consonants were perceived correctly more often than unfamiliar ones across the three age-language status groups and across the three proficiency subgroups; and (d) the KEB children perceived correctly more often than the KEB adults, the KEB children and adults perceived correctly more often than the KM children, and the ED and BAL subgroups perceived correctly more often than the KD subgroup.

The results of the NSIT showed (a) consonants were produced more accurately than vowels, and (b) familiar sounds were produced more accurately than unfamiliar ones, across the three age-language status groups. Also, (c) familiar consonants were produced more accurately than unfamiliar ones in the KEB and KM child groups, and (d) unfamiliar vowels were produced more accurately than a familiar one in the KEB child group, but the reverse was true in the KEB adult and KM child groups. The KEB children produced sounds correctly significantly more often than the KM children and the KEB adults, though the percent correct differences were smaller than for perception. Production differences were not found among the three proficiency subgroups. Perception on the NSCM task was compared to production on the NSIT and NRT. Weak positive correlations were found between perception and production (NSIT) for unfamiliar consonants and sounds, whereas a weak negative correlation was found for unfamiliar vowels. Several correlations were significant for perceptual performance on the NSCM task and overall production performance on the NRT: for unfamiliar consonants, unfamiliar vowels, unfamiliar sounds, consonants, vowels, and overall performance on the NSCM task. Nonetheless, no significant correlation was found between production on the NSIT and NRT. Evidently these are two very different production tasks, where immediate imitation of single syllables on the NSIT results in high performance for all groups.

Findings of the present study suggest that (a) perception and production of L2 consonants differ from those of vowels; (b) perception and production of L2 sounds involve an interaction of sound type and familiarity; (c) a weak relation exists between perception and production performance for unfamiliar sounds; and (d) L2 experience generally predicts perceptual and production performance.

The present study yields several conclusions. The first is that familiarity of sounds is an important influence on L2 learning, as claimed by both SLM and PAM. In the present study, familiar sounds were perceived and produced correctly more often than unfamiliar ones in most cases, in keeping with PAM, though experienced L2 learners (i.e., the KEB children) produced unfamiliar vowels better than familiar ones, in keeping with SLM. Nonetheless, the second conclusion is that neither SLM nor PAM consistently and thoroughly explains the results of the present study. This is because both theories assume that the influence of L1 on the perception of L2 consonants and vowels works in the same way as for production of them. The third and fourth conclusions are two proposed arguments: that perception and production of consonants are different than for vowels, and that sound type interacts with familiarity and L2 experience. These two arguments can best explain the current findings.

These findings may help us to develop educational curricula for bilingual individuals listening to and articulating English. Further, the extensive analysis of spontaneous speech in the present study should contribute to the specification of parameters for normal language development and function in Korean-English bilingual children and adults.

Dedicated to my parents, Yong-Soon Lee and Kyu-Joong Yoon,

For their endless support and faith

**어머니 아버지께서 해주신 지원과 믿음에 감사하며**

**이 논문을 바칩니다.**

ACKNOWLEDGEMENTS

Though only my name appears on the cover of this dissertation, I could never have reached the heights or explored the depths without help, support, and guidance from many people. Therefore, it is my sole pleasure to thank those who made this dissertation possible. I owe my deepest gratitude to my advisor, *Dr. Cynthia J. Johnson*, whose guidance and advice helped me throughout the research and the writing of this dissertation. Her mentorship was paramount in providing a well rounded experience consistent with my long-term career goals. She helped me with the design and analyses of the study, patiently corrected and edited my manuscript, and heartily encouraged me when I encountered hardships. Her attitude of a true scientist was exceptionally inspiring and enriched my growth as a researcher and teacher that I want to be. I could not have imagined having a better advisor and mentor for my Ph.D. studies.

I would like to thank *Dr. Jont B. Allen* for his insightful comments on and financial support for my research. His research on perceptual sound confusion has triggered interest in my current dissertation topics. My sincere appreciation also goes to *Dr. David Kuehn*, whose supervision and insightful comments from the preliminary to the concluding level enabled me to develop an understanding of scientific research. I am grateful to him for holding me to a high research standard and enforcing strict validations. His comments over the years helped me refine the literature review and develop a pilot study to the dissertation research project with his extensive knowledge with speech science. I gratefully acknowledge *Dr. David Gooler* for his valuable advice in the prospectus meeting and supervision in the organization of the dissertation. Also, I want to thank him for a discussion on his approach and research on speech perception with hearing impairments and also for inviting me as a guest lecturer on phonological disorders and multilingualism that I have enjoyed so much. I would also like to thank *Dr. Chin-Woo Kim* for all his heartily support for the past years. A discussion with him on Korean phonetics and phonology helped me a lot to justify the stimuli used in the dissertation research. Most of all, however, I am much indebted to him for his caring nature. He never forgets celebrating my birthday for past 8 years, saying "happy birthday Seok-Youn. Are you now 21 years old?" My special thanks goes to *Dr. Torrey Loucks*, who gratefully served as a committee for the prospectus meeting, for his encouragement and his valuable and durable advice on the dissertation research.

I am also indebted to many of my colleagues and friends who supported me: *Feipeng Li* and *Jihye Bae* (both from Electrical and Computer Engineering at University of Illinois at Urbana-Champaign), *Feipeng*, for the development of matlab codes for the Nonsense Syllable Confusion Matrix task, and *Jihye*, for helping me recruit participants; *Julie Mahieu* and *Michele Turk* (both from Speech and Hearing Science at the University of Illinois), for evaluating participants' production tasks; *Heeyoun Cho* and *Eunha Kim* (both from Linguistics at the University of Illinois), for their loyal friendship; *Myunghwa Choi*, for helping me recruit 18 Korean monolingual children in Korea; and *all of my participants* in the present study, for patiently participating and completing a 5- to 7-hour-long experiment.

Above all, I am grateful to *my family* for their constant support for past 8 years. I am so aware of their sincere sacrifice for my study abroad: they must have retrenched living expenditures and saved all their pennies to support me, even during the depression, without a complaint. To *my grandfather* in heaven, I thank him for calling me from Korea during his last moments to encourage me, even though I was not able to finish my doctoral degree or marry someone as he wished, before he was called to heaven.

To answer to all these thanks, I will definitely become a good researcher and professor, in order to contribute to the community even a little. I will do my best. 헛된 노력이 되지 않도록, 조금이나마 사회에 보탬이 되기 위해, 꼭 좋은 학자이자 교수가 되도록 노력하겠습니다. 私は きっと いい 学者に 教授に なります。 少しでも 社會のため 役に 立つように 勞力する つもりです。

TABLE OF CONTENTS

CHAPTER I: INTRODUCTION……………………......……………….…….………………… 1

Review of the Literature.....……………………………….………………………. 8

*Speech Perception and Production*…………………………………….. 8

*Acoustic Cues to Speech Perception*…......…………………… 8

*Speech Perception in English-Listening Children*……………… 13

*Speech Production in English-Speaking Children*………..…... 17

*Korean Consonants and Vowels*..……………………………………… 19

*Korean Consonants*……………..…………………………….. 19

*Korean Vowels*……….………………………………………… 24

*Familiar versus Unfamiliar English Sounds*………………….. 25

*Speech Perception and Production of a Second Language*…………….. 27

*First Language Influences and Second Language Experience*.... 27

*Speech Learning Model*……………………………………... 35

*Perceptual Assimilation Model*………………………………... 42

CHAPTER II: LANGUAGE SAMPLE ANALYSES……………………......………... 54

Rationale, Research Question and Hypothesis.………………………………….. 54

Method.…………………………………………………………………… 56

*Participants*………………………………………………………. 56

*Task Procedure*…………………………………………………… 57

*General Procedure*……………………………………………….. 61

Results……………………………………………………………………………….. 62

*Language Survey and Articulation and Vocabulary Tests*…………….. 62

*Comparisons of the Three Age-Language Status Groups*.... 62

*Comparisons of the Three Proficiency Subgroups*………... 64

*Language Sample Analyses*……………………………………….. 65

*Comparisons of the Three Age-Language Status Groups*..... 65

*Comparisons of the Three Proficiency Subgroups*………... 70

Summary and Conclusions……………………………………………………. 70

CHAPTER III: NONSENSE SYLLABLE CONFUSION MATRIX TASK…......…… 72

Rationale for a Perceptual Task…....………………………………………….. 72

Research Questions and Hypotheses………………………………………... 75

*Research Question 1: Familiar vs. Unfamiliar Sounds*……….…….. 76

*Research Question 2: Consonants vs. Vowels*……………..……….. 78

*Research Question 3: Group Comparisons*…………………………… 79

*Research Question 4: Sound Confusion*……………………………… 80

Method.…………………………………………………………………………..... 82

*Participants*……………………………………………………………… 82

*Stimuli*…………………………………………………………………… 82

*Task Procedure*……………………………………………………….... 84

Results…………………………………………………………………………….. 84

*Familiar vs. Unfamiliar Consonants and Vowels*…………………….. 85

*Comparisons of the Three Age-Language Status Groups*..... 85

*Select Group Comparisons*…………………….…………………… 87

*Comparisons of the Two Age Groups*.……………………… 87

*Comparisons of the Three Proficiency Subgroups*..……….. 88

*Perceptual Performance for Individual Sounds*……………………… 89

*Group Differences for Each Phoneme*……………………… 90

*Perceptual Confusions for Each Phoneme*.…………………. 92

Summary and Conclusions……………………………………………………. 94

CHAPTER IV: NONSENSE SYLLABLE IMITATION TASK……….....…………… 96

Rationale for a Production Task…………………………….………………….. 96

Research Questions and Hypotheses……………………………………………… 98

*Research Question 1: Familiar vs. Unfamiliar Sounds*……….……... 98

*Research Question 2: Consonants vs. Vowels*……………….……….. 99

*Research Question 3: Group Comparisons*…………………………… 100

*Research Question 4: Correlations between Perception and Production* 100

Method.………………………………………………………………………… 102

*Participants*………………………………………………………………. 102

*Stimuli*…………………………………………………………………….. 103

*Task Procedure*…………………………………………………………… 103

Results…………………………………………………………………..…… 104

*Familiar vs. Unfamiliar Consonants and Vowels*…………………..… 104

*Comparisons of the Three Age-Language Status Groups*..... 104

*Select Group Comparisons*………………………………………………. 106

*Comparisons of the Two Age Groups*.……………………… 106

*Comparisons of the Three Proficiency Subgroups*.……….. 108

*Correlations between Perceptual and Production Performance*.……. 108

Summary and Conclusions………………………………………………………. 109

CHAPTER V: NONWORD REPETITION TASK……….………..………………… 112

Rationale, Research Questions and Hypotheses………………………………….. 112

Method.……………………………………………………………………… 114

*Participants*……………………………………………………………….. 114

*Stimuli*…………………………………………………………………….. 114

*Task Procedure*…………………………………………………………. 114

Results……………………………………………………………………….. 115

*Comparisons of the Three Age-Language Status Groups*.................... 115

*Correlations among Perception, Production, and Phonological Awareness*………………………………………………………… 117

Summary and Conclusions……………………………………………………….. 118

CHAPTER VI: DISCUSSION………..…………………………………....……….. 120

Influence of L2 Experience and L2 Proficiency………………………………….. 120

Speech Learning Model and Perceptual Assimilation Model……………………. 123

Familiarity Interplay for Perception and Production of L2 Consonants and Vowels…………………………………………………………………….... 125
*Relations between Perceptual and Production Performance*……..…. 125
*High-scoring vs. Low-scoring Sounds*……………………………….. 128
*Sound Type: Consonants vs. Vowels in Perception*………………….. 131
*Sound Type by Familiarity in Perception*…………………………..... 134
*Sound Type: Consonants vs. Vowels in Production*.………………… 135
*Sound Type by Familiarity in Production*…………………………..... 137
Towards an Interplay Model of Processing of L2 Consonants and Vowels….…..... 138
Limitations and Future Study…………………………………………………… 141

TABLES……………………………………………………………….……….... 144

FIGURES…………………………………………………………..……..….. 178

REFERENCES………………………………………………………...……….... 183

APPENDIX A: ABBREVIATIONS AND DISTINCTIVE FEATURES..……..……..... 202

APPENDIX B: ENGLISH CONSONANT INVENTORY..........………………....…… 205

APPENDIX C: PREDICTIONS OF THE SPEECH LEARNING MODEL AND

PERCEPTUAL ASSIMILATION MODEL……………………………...……...…… 206

APPENDIX D: FAMILY LANGUAGE PREFERENCE SURVEY……………….......… 207

APPENDIX E: CONFUSION MATRICES FOR THE 10 CONSONANTS AND THE 3 VOWELS IN THE NONSENSE SYLLABLE CONFUSION MATRIX TASK IN KOREAN-ENGLISH BILINGUAL AND KOREAN MONOLINGAL CHILDREN AND ADULTS……………………………………………………………………….. 212

# CHAPTER I

# INTRODUCTION

A number of studies on bilingualism have examined how a first language (L1) interferes with perception and production of a second language (L2), which sound groups may cause L2 learners more difficulty in learning, how perception and the production of differs for adult and child L2 learners, and how L2 experience relates to perceptual and production performance of L2 speech sounds. The present study investigates which types of English L2 sounds are more challenging for Korean-English bilingual (KEB; see Appendix A for abbreviations used hereafter) children and adults and for Korean monolingual (KM) children to perceive and produce under the influence of L1 Korean, which is phonologically and phonetically different from English. The present study also examines whether the L2 sounds that were difficult to perceive and produce for KEB children—who acquire L2 early in life—are also challenging for KEB adults—who acquire L2 later in life and KM children—who are not exposed to English in naturalistic settings. In addition, the present study compares performance on a speech perception task with performance on speech production tasks. The present study includes a speech perception task, called the *Nonsense Syllable Confusion Matrix task* (NSCM; Allen, 2005; Miller and Nicely, 1955), two speech production tasks, called the *Nonsense Syllable Imitation Task* (NSIT) and the *Nonword Repetition Task* (NRT; Dollaghan & Campbell, 1998), and language sample analysis (LSA) for spontaneous speech in English.

It has been well documented that the age of L2 acquisition has a lasting effect on speech perception and production. Numerous findings in cross-linguistic speech studies have shown that perceptual and production performance of nonnative speech sounds are closely related to the age of L2 acquisition (or age of arrival, AOA) and the amount of exposure (or the length of residence, LOR; Baker et al, 2008; Best & McRoberts, 2003; Eimas et al, 1971; Johnson & Newport, 1987; Mayberry & Eichen, 1991; Ploog, 1984; Yeni-Komshian et al. 2000). In general, child L2 learners who are exposed to L2 before age 7 can achieve native-like fluency in L2; however, adult L2 learners typically attain lower L2 proficiency (Mayberry & Eichen, 1991; Ploog, 1984; Yeni-Komshian et al., 2000). Further, bilingual children who are exposed to an L2 prior to age 3 are generally considered to be simultaneous bilinguals, whose development of both languages is comparable to that of monolinguals (de Houwer, 1995; McLaughlin, 1978; Oller, Eilers, Urbano, & Cobo-Lewis, 1997; Peña, Bedore, & Rappazzo, 2003). On the other hand, bilingual children who acquire an L2 after age 3 are generally considered to be simultaneous bilinguals whose development of an L1 precedes that of an L2.

Some bilingual studies have shown that age effect on L2 proficiency is not always significant (Bialystok & Miller, 1999; Epstein, Flynn, & Martohardjono, 1996). Difficulty in achieving native-like proficiency may be attributed to several factors other than AOA, such as amount of L2 exposure, environment of L2 usage, motivation to learn L2, and so on. For L2 experience, Flege and his colleagues have shown that L2 learners with late AOA or with shorter LOR (even at the same age) perceive and produce nonnative sounds under a greater influence of L1 than L2 learners with early AOA and with longer LOR (Flege, Bohn & Jang, 1997; Gildersleeve-Neumann, Kester, Davis and Peña, 2008).

In order to examine the effect of AOA and LOR on perceptual and production performance of English L2 sounds, participants varied in age and L2 experience were recruited. KEB children, aged between 8 and 13 years, were first exposed to L2 English at a young age and had lived in the U.S. for at least 2 years. KEB adults, aged 18 or older, were first exposed to L2 English later in life and had lived in the U.S. for at least 2 years. KM children, aged between 8 and 13 years, had never been to an English-speaking community and had just begun to learn English as foreign language (FL) at school, typically for one hour per week. KEB children were further grouped into one of the three proficiency subgroups, depending on their AOA, LOR, and L2 usage at home and work, as reported by their parents. Therefore, the present study is able to examine whether perceptual and production performance of L2 is closely related to the age of L2 acquisition (or AOA) and the amount of L2 exposure (or LOR).

Literature on speech perception and production of L2 sounds argued that difficulties in listening to or articulating L2 sounds may be predicted by relations between L1 and L2 phonology. Nonetheless, explanations for performance characteristics of speech perception and production diverge into several theories. Some investigators argue that an L2 speech sound similar but not identical to one in L1 is produced less native-like because speakers do not notice minor differences between the two phones, yet an L2 sound different from or new to L1 is produced more native-like. On the other hand, other investigators argue that two L2 sounds which are discriminated from each other with a new nonnative contrast cause more confusions in speech perception than two L2 sounds which are discriminated with a nonnative contrast. According to the former studies, similarity of L2 phones to L1 phones will hinder accurate perception and production; whereas, the latter argues that similarity of L2 contrasts, which distinguish two L2 phones, will help accurate perception.

For example, Flege (1987) showed that English-speaking learners of French sounds produced the new vowel /y/ (i.e., a high front, rounded vowel) more French like than the vowel

/u/ (i.e., a high back vowel), which is similar in the two languages. Flege argued that native English speakers fail to notice subtle differences between the French /u/ and the English /u/. Therefore, similarity of an L2 phoneme to its corresponding L1 phoneme results in difficulty in producing the L2 phoneme with native-like accuracy. In contrast, Best, McRoberts, and Sithole (1988) showed that English monolingual speakers can discriminate a pair of Zulu clicks that are new to English monolingual speakers with 80% accuracy. The investigators argued that native English speakers employed a voicing contrast in English to discriminate the pair of Zulu clicks from each other. The former argues that similarity of L2 phones to L1 phones will hinder accurate production and perception; whereas, the latter argues that similarity of L2 contrasts, which distinguish two L2 phones, to L1 contrasts will promote accurate perception (and possibly production, with an extended interpretation of PAM). Therefore, the two theories make opposite predictions about perception and production of English L2 sounds. The present study will examine whether the two models can predict perceptual and production performance in KEBs and KMs.

With regard to the relation of speech perception and production in monolingual children, it has been shown that perceptual development begins earlier but its development appears to reach fully adult-like performance on certain tasks later than production development for everyday speech. Typically by ages 6 to 8, normally developing children display adult-like speech production in everyday, casual, connected speech; making no errors in pronunciation of their native language (Dodd, Holm, Hua, & Crosbie, 2003). Nonetheless, their performance on certain speech perception tasks displays different patterns from adults' until age 13 (Nittrouer, 1992; Nittrouer & Miller, 1997; Sussman, 1993; Sussman & Carney, 1989). English-listening infants can discriminate /p/ from /b/ as young as 1 month of age (Eimas, Siqueland, Jusczyk, & Vigorito, 1971) even before they can produce those sounds. Also, infants are sensitive to vowels in their native language more than to vowels in foreign languages as young as 6 months of age (Kuhl, 1992) and begin recognizing phonemes in their native language as young as 10 months of age (Werker & Lalonde, 1988).

Despite such early perceptual development, children aged from 5 to 10 years perform differently from adults on the discrimination of their L1 sounds (Nittrouer, 1992; Sussman, 1993a), and children younger than 13 years perform differently from adults on the identification of consonant in noise (Elliott, 1979). Some investigators have argued that children's voice onset time (VOT) threshold to discriminate one voiced stop from another is higher than adults'. Others have argued that children weigh available acoustic cues differently than adults. The present study does not aim to determine which theoretical view best explains speech perception in bilingual children, but merely to examine whether bilingual children show different L2 perceptual patterns from adults, under the influence of L1.

In general, children's ability to produce speech sounds of their L1 is fully developed by the age of 8 years, but their ability to perceive them on certain tasks continues to develop throughout the lifespan and never stops developing, because even adult learners or late learners showed the ability of discriminating and identifying L2 sounds to be increased with greater L2 experience. Therefore, certain aspects of phonemic perception seem to occur earlier but progress more slowly than phonemic production in everyday speech. This is argued to hold true for bilinguals, too. Unfortunately, most bilingual studies are based on production, but far fewer on perception tasks, particularly in school-aged children and even less in school-aged children who are bilingual speakers of Korean and English. The present study will fill this gap in the literature and hopefully will explain the developmental process for speech perception and production in linguistically-divergent, school-aged children.

In recent years, researchers in bilingualism and in speech-language pathology have emphasized the need for reliable language assessment for bilingual children (Roseberry-McKibbin, Brice, & O'Hanlon, 2005), because children from culturally and linguistically divergent backgrounds appear on clinical caseloads in increasing numbers. Assessment tools for evaluating language proficiency are now being developed, but mostly for Spanish-English bilinguals; rarely or not at all for Korean-English bilinguals, as the author acknowledges (S-Y. Yoon, 2010). Therefore, language sample analyses of KEB children, KEB adults, and KM children, varying in age and L2 proficiency, provides valuable information for the development of better measures and assessment tools and developmental norms in Korean-English bilingual children and adults.

The purpose of the present study is to examine (a) which English sounds are more challenging for KEBs and KMs to perceive and produce under the influence of Korean as L1; (b) whether KEB children perceive and produce English sounds differently from KEB adults or KMs; (c) whether speech perception and production of English are related with each other; and (d) whether English proficiency predicts perceptual and production performance of English nonsense syllables. The results of the present study may give a better way to predict English L2 sounds that are challenging for English language learners who learn English as L2 in the U.S., or who learn English as a foreign language (FL) in Korea to perceive and produce. In addition, the results of the present study may provide a better understanding of the effects of AOA (by a comparison between KEB children and adults) and LOR (by a comparison between KEB and KM children) on perceptual and production performance. Furthermore, language sample analyses may contribute to the development of normative data for various domains of language performance, such as syntax, semantics, morphology, pragmatics, phonetics, and phonology. Eventually, the study may help us to (a) specify parameters for normal language development and function in Korean-English bilinguals, and (b) develop educational curricula for listening to and articulating English for use with children who have just begun to learn English as a foreign language.

In this chapter, I will review literature that addresses (a) which acoustic cues are responsible for perceiving speech sounds, (b) how children perceive sounds differently from adults, (c) how children develop the ability to produce their native language, (d) how the Korean phonological system is different from that of English, (e) how speech perception and production of one's second language are influenced by one's first language, and (f) how such differences are explained in different theories. From this literature, I will conclude that spectral and durational cues are responsible for perceiving speech sounds, some investigators have claimed that children are more reliant on formant transitions and durational cues than adults, children learn some speech sounds later than others but typically produce the sounds of their native language well by 8 years of age, Korean has many fewer fricatives than English, one's first language does influence perception of a second language, and two prominent theories have been used to account for the perception and production of a second language by adults.

# Review of the Literature

## *Speech Perception and Production*

Understanding speech perception requires interdisciplinary efforts and encompasses a breadth of specialties, including sensory and cognitive psychology, linguistics, communication engineering, artificial intelligence, audiology, and speech-language pathology. In the language area, speech perception studies are essential in order to investigate language use, language development, and second language acquisition. A number of previous studies have suggested that perception of nonnative speech sounds, i.e., those from L2, is greatly influenced by language experience. For example, native speakers of Japanese find it difficult to perceive the contrast between /l/ and /r/ in English—a contrast that is very easy for listeners who are native English speakers (Flege, Takagi, & Mann, 1995; 1996). It is suggested that the loss of ability to discriminate nonnative contrasts begins as early as 10- to 12-months of age and the ability to discriminate nonnative contrasts weakens with increasing L1 experience (Best, McRoberts, & Sithole, 1988). Nevertheless, there is no clear consensus about how children develop speech perception and eventually achieve adult-like performance. This section begins with an account of how varying speech signals of English are mapped to the discrete phonemes with respect to voicing, manner, place of articulation and vowels, and then continues with a review of studies of development of speech perception and production in English monolingual infants and children.

### *Acoustic Cues to Speech Perception*

*Voicing.* Speech perception studies were vigorously investigated in the 1950s at Haskins Laboratories, which showed that speech perception is categorical. In other words, an individual does not or cannot hear speech sounds as falling halfway between /ba/ and /pa/ even though /ba/

and /pa/ lie on a physical VOT continuum (Abramson & Lisker, 1970; 1973). For example, both a synthetic /p/ with a VOT value of +120 ms and a synthetic /p/ with a VOT value of +60 ms would be identified as the same phoneme, /p/, by English listeners, and a synthetic token with a VOT of 0 or less would be identified as /b/. Nevertheless, a synthetic /p/ with a VOT value around +25 ms would be identified either as /p/ or /b/.

Even though VOT differentiates voiceless from voiced stops in English, additional acoustic information carried on the vowel following a consonant contributes to the perception of voicing on the consonant. Fundamental frequency ($F_0$) combines with VOT to specify the voicing feature of stops. Ohde (1984) showed that $F_0$ values associated with voiceless aspirated stops in English were significantly higher than $F_0$ values associated with voiced stops, whereas VOT values overlap between voiceless unaspirated and voiced stops. With regard to voicing of consonants that follow vowels, Klatt and Klatt (1990) showed that the amplitude for the first two harmonics (i.e., $H_1$ and $H_2$) is slightly lower when a vowel is followed by a voiced consonant. They also argued that the difference of the amplitude between $H_1$ and $H_2$ is greater when a vowel is followed by a voiced consonant. In production, stiffness of the vocal folds is higher at the offset of voiceless fricatives than voiced ones (Halle & Stevens, 1971). Interestingly, voicing confusion in perception is not always bidirectional. For example, voiced consonants are more likely to be identified as voiceless than vice versa (Abdelatty Ali, Spiegel & Mueller, 2001).

*Manner*. Salient acoustic cues for manner of articulation vary for consonants. VOT contributes not only to a perceived voicing feature (Abramson & Lisker, 1970; 1973) but also to a perceived laryngeal feature often noted by Korean linguists in studies of Korean obstruents (Dart, 1987; Han & Weitzman, 1970; S-A. Jun, Beckman, & H. Lee, 1998; Kagaya, 1974; C-W.

Kim, 1970). With respect to manner of articulation, the most salient acoustic cue for the perception of stops is a silence between the articulatory release and the onset of voicing. The presence of frication or turbulent noise is a cue for detecting fricative sounds. The burst followed by a silent gap in stops and the frication noise in fricatives are robust cues; thus, it is unlikely that stops and fricatives will be confused with each other in perception. Miller and Nicely (1955) showed that stops are clearly distinguishable from fricatives at SNR = +12 dB—the condition which seemed to result in the least spread of confusions (i.e., the best performance condition). In addition, Phatak and Allen (2007) showed that fricatives are more likely to be confused with one another than with other manners of articulation such as stops. In their study, the only fricative-stop or stop-fricative confusion (see Appendix B for English consonant inventory) occurred for the voiced bilabial stop /b/, which was often confused with /v/ and /ð/. Nasals or approximants are characterized by the presence of low resonance at about 250 Hz with diminished spectrum amplitude at the high frequencies (Ladefoged, 2005; Steriade, 1992). These sonorant sounds are shown to be less confusing than obstruent sounds (Miller & Nicely, 1955; Phatak & Allen, 2007).

*Place*. Place of articulation is less easily captured than manner of articulation. Formant transitions on the vowel portion, particularly for the second ($F_2$) and the third formants ($F_3$), seem to be the acoustic cues responsible for perception of place of articulation. Nevertheless, formant transitions do not show consistent patterns even for the same consonant (Delattre, Cooper, & Liberman, 1955). This is because patterns for $F_2$ and $F_3$ are affected by the following vowels, as well as by the position of the consonant in relation to the vowel—for example, consonant-vowel (CV), vowel-consonant-vowel (VCV), or vowel-consonant (VC) templates

(Liberman, Ingemann, Lisker, Delattre & Cooper, 1959). Other studies argue that the primary acoustic cue for the place feature is the gross shape of the sliced spectrum at the release of a consonant articulation and that listeners are able to integrate such invariance of spectral information (Kewely-Port, 1983; Stevens, 1995, 2000; Stevens & Blumstein, 1978). For example, the frequency value of $F_2$ at a moment right after release of a front stop consonant is higher than the $F_2$ of a more back stop consonant, as viewed from a spectral slice.

Many consonant confusion studies have shown that perception of place of articulation interacts not only with pre- or post-vocalic position of a consonant, but also with voicing and manner of articulation, which cause listeners difficulty in identifying place of articulation. Lisker and Abramson (1964) argued that *bilabial* stops have a shorter VOT than *velar* stops; indicating VOT interacts with place of articulation in perceiving voiced stops in English. Because VOT difference between voiceless and voiced cognates is smaller for bilabials than velars, bilabial voiced stops are more likely to be perceived as their voiceless cognates than are velar voiced stops (Benki, 2001). Likewise, *anterior* (see Appendix A also for terminology of distinctive features) fricatives such as /f, v, θ, ð/ are identified less correctly than *posterior* fricatives such as /s, z, ʃ, ʒ/ (Phatak & Allen, 2007). From a physiological point of view, posterior constrictions involve greater mass and a wider contact area than anterior constrictions; thus, posterior constrictions may last longer than anterior constrictions (Hardcastle, 1973). Other studies support that *dorsal* stops are perceptually more salient than *labial* or *coronal* stops (Hume, Johnson, Seo & Tserdanelis, 1998; Jiang, Chen & Alwan, 2003). From an aerodynamic point of view, posterior constrictions create a smaller air cavity and delay the initiation of vocal fold vibration for the following vowel (Stevens, 1998). All together, these views suggest anterior sounds will be more confusing than posterior sounds.

*Vowels*. A significant body of studies has discussed that spectral information such as formant frequencies (e.g., $F_1$, $F_2$, and $F_3$; Lindblom & Studdert-Kenney, 1967) and formant transitions (e.g., onglides and offglides; Nearey & Assmann, 1986) are necessary to distinguish one vowel from another. Error rates of listeners for vowel identification were affected by static formant patterns such as formant frequencies and transitions (Hillenbrand & Nearey, 1999). Several studies, however, have argued that temporal or dynamic information (e.g., duration and spectral change of vowels) also plays an important role in vowel perception (Gottfried, Miller, & Payton, 1990; Hillenbrand, Getty, Clark, & Wheeler, 1995). In relation to consonants, several studies have shown that listeners can identify vowels better when vowels are embedded in consonant-vowel-consonant (CVC) syllables than when they are isolated (Strange, Edman, & Jenkins, 1979; Gottfried & Strange, 1980; Hillenbrand et al., 1995; Hillenbrand, Clark, & Nearey, 2001). The presence of final consonants contributes to the identification of vowels more than does the presence of initial consonants (Strange, Edman, & Jenkins, 1979).

The present study examines which English sounds may cause perceptual confusion and articulatory difficulties for 8- to 13-year-old bilingual speakers of Korean and English, compared to adult bilingual speakers. It has been demonstrated in numerous studies that young children display different patterns of speech perception than adult listeners. Therefore, if children's perception operates in a different fashion than adults', this should lead to different performance on experimental tasks in the present study. Up to this point, however, whether or not children in mid- to late-elementary school have adult-like speech perception has been controversial. Next, a body of literature will be reviewed to address how young children perceive speech sounds differently from adults.

*Speech Perception in English-listening Children*

As a child develops his or her language skills, his or her listening abilities become tuned to the language to which he or she is extensively exposed. Typically developing children should be able to discriminate consonants in his or her native language by the age of 11 or 12 months (Minagawa-Kawai, Mori, Naoi, & Kojima, 2006); however, they show different speech perception processing than adults until the age of 13 years (Elliott, 1979; Neuman & Hochberg, 1983). Some studies using synthetic CV syllables that varied in VOT have shown that children's speech perception is different from adults' with respect to difference limens[1] (Elliott, Longinotti, Meyer, Raz, & Zucker, 1981; Elliott, 1986; Elliott, Busse, Partridge, Rupert, & DeGraaff, 1986; Sussman & Carney, 1989; Sussman, 1991, 1993a, 1993b). Other studies argue that children's speech perception is different from adults' with respect to a developmental weighting shift for acoustic cues (Harris, 1958; Morrongiello, Robson, Best, & Clifton, 1984; Nittrouer & Studdert-Kenney, 1987; Nittrouer, 1992, 1996a; Nittrouer & Crowther, 1998).

One early speech perception study in children, carried out by Elliot and her colleagues, showed that 6- to 8-year-old children required significantly longer VOTs than adults in order to produce the same performance on sound discrimination tasks for voicing distinctions among stop

[1] Difference Limens are the smallest difference that can be discriminated between two stimuli or a difference that is barely above the threshold of detection. Sometimes, the term is used interchangeably with the term Just Noticeable Difference (JND), which refers to the smallest detectable difference to be perceived as a different stimulus from the original one. In the present study, DL refers to different intensity between the groups of different listeners that is required to perceive the same stimulus.

consonants (Elliot et al., 1986). A study by Sussman and Carney (1989) yielded similar results. Sussman and Carney tested 5- to 10-year-old children and adults in a place discrimination task, where they labeled “change” or “no-change” in synthetic CV stimuli. For example, in a change trial, a child would hear one syllable and then a second syllable. The child was instructed to push the button if she or he noticed any change. In the study, CV stimuli were synthesized by changing the first three values of vowel formants in seven different steps on the /ba/ - /da/ continuum. The end point for /ba/ had 286 Hz, 543 Hz, and 1360 Hz for the first ($F_1$), second ($F_2$), and third ($F_3$) formants, respectively; and the other end point for /da/ had 260 Hz, 1620 Hz, and 3499 Hz for each formant. From the end point for /ba/ to that for /da/, $F_1$ decreased in seven steps; $F_2$ and $F_3$ increased in seven steps. Seven synthetic CV stimuli on the continuum also had durational differences; one set of seven synthetic CV stimuli had shorter duration in $F_2$ and $F_3$ by 50 ms than the other set.

Sussman and Carney (1989) demonstrated that larger formant frequency differences in $F_2$, $F_3$, and longer formant duration were necessary in order even for the oldest children to label “change” in stop consonants. It appears that children even at the age of 10 years require bigger change in formants and longer duration to achieve the same level of the performance with adults in discriminating sounds. In a follow-up study, Sussman (1993) found no significant difference in the selective adaptation effects induced by focused attention for children, with the same /ba/ - /da/ continuum. Evaluating distracter identification performance, she concluded that a great capacity to pay attention is not necessarily required to perceive and process acoustic information. Thus, a series of Sussman's studies supports Elliot's study, which argued that children have larger difference limens than adults.

In the study of Morrongiello, Kulig, and Clifton (1984), 5-year-old children were tested for identification of the synthetic syllables “say” vs. “stay,” that systematically varied the duration of silence (i.e., the presence of /t/) between a /s/-like noise and the vowel onset. They found that children's identification performance on the “say-stay” continuum was better than that of adults for the gap with a short duration of a silence, and proposed that, unlike adults, children pay more attention to vowel formant transitions than the silence gap to identify the /st/ cluster. Studies done by Nittrouer and her colleagues also support that children rely on different acoustic signal, such as frication noise than adults (Nittrouer, 1992; Nittrouer & Miller, 1997).

Nittrouer (1992) compared the “/s/-response” of sixteen adults with that of seventeen 2- to 4-year-old children. Four types of stimuli were given: two hybrid stimulus sets, composed of a *synthetic* /s/- or /ʃ/-like fricative noise and *natural* vocalic portions of the vowel /u/; and two synthetic stimulus sets, composed of a synthetic /s/- or /ʃ/-like fricative noise and synthetic vocalic portions of the vowel /u/. Young children identified a /su/ syllable, with a synthetic noise occurring at less than 3000 kHz and the /u/ vowel from the natural speech syllable /su/, as /s/ more than 60% of the time. In contrast, adults identified the syllable as /ʃ/ more than 60% of the time. This result was interpreted by Nittrouer as indicating that children's perceptual weighting of acoustic cues for fricative consonants depends more on formant transitions in the vowel portion than does the weighting of adults.

Nittrouer and Miller (1997) replicated the /s/ - /ʃ/ labeling task in Nittrouer (1992). Synthetic frication noises were combined with natural vocalic portions from /su, sa, ʃu, ʃa/. The results showed a significant main effect for age: 4-year-olds, 7-year-olds, and adults. The differences among the groups were greater for /s/ transitions than for /ʃ/ transitions. Regardless of vowel context, the phoneme boundary between /s/ and /ʃ/ was highest for adults, less for 7-year-olds, and lowest for 4-year-olds (3135 Hz, 3161 Hz, and 3026Hz, respectively). Furthermore, an interaction between age and transition (i.e., the four vowel-transition contexts) was significant, suggesting developmental trend in perceiving transitions. Taken together, these results for discrimination tasks suggest that the auditory processing of children is different because children weigh acoustic cues differently, and not merely because children need greater differences in each acoustic cue than adults do in order to identify phonemes. Therefore, the results give support to Morrongiello's study, which argued that children weigh formant transitions more than other available acoustic cues when discriminating sounds.

As has been discussed, most of the literature in speech perception has demonstrated that speech perception in young children is significantly different from speech perception in adults. Nevertheless, studies disagree on what accounts for such a difference. One group of studies argues that children and adults use the same acoustic cues, but children need a greater amount of information (e.g., longer duration, with greater formant transitions). The other group of studies argues that children are sensitive to different types of acoustic cues than are adults (e.g., greater sensitivity to formant transitions than to other available acoustic cues). According to literature about the development of speech production in English-speaking children, which will be reviewed in the next section, normally developing children display adult-like accuracy in speech production by ages 6 to 8 years in everyday, casual, connected speech; making no errors in pronunciation of their native language (Dodd, Holm, Hua, & Crosbie, 2003; Smit et al., 1990; Templin, 1957; Wellman et al., 1931). Nonetheless, studies of speech perception in children reveal that speech perception progresses begins to show developmental change earlier but progresses more slowly than speech production.

With regard to the relation between speech perception and production, whether the acquisition of listening abilities is a prerequisite for the acquisition of speaking abilities has been controversial. Some believe that perception is somehow based on the mechanism of speech production (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967); while others believe that perception guides how to implement articulatory gestures (Ladefoged, De Clerk, Lindau, & Papcun, 1972). The former argued that the map of the continuous speech signal into categorical and conceptually discrete (phonemic) units supports the idea that perception is eventually influenced by the phonological system of one's native language. Against to this argument, Ladefoged et al. argued that even non-speech is categorical and a child who cannot listen to a sound also cannot speak it. Results of the present study may contribute to the exploration of the relation between speech perception and production from a different perspective, because most L2 learners learn how to speak L2 sounds when they first hear them (i.e., the beginning point of L2 listening generally coincides with that of L2 speaking), unlike L1 listeners—who perceive speech sounds before they produce their native language. Therefore, it is worth investigating whether L2 learners do or do not show parallel progress in L2 perception (as measured by identification of L2 sounds) and L2 production (as measured by articulation of L2 sounds).

*Speech Production in English-speaking Children*

The vocal tract in infants is smaller and flatter in shape than in adults (Vihman, 1996). This physiological difference restricts the range of articulatory movements and hinders them in making complicated air disturbances in the oral cavity in order to produce target sounds. The descent of the larynx typically starts around 4 months of age and is completed by 3 or 4 years. Articulatory movements and coordination of those movements are required to make speech stably and repetitively as adults do. At birth, infants do not have the fine muscle control to make speech sounds repeatedly and correctly. Due to physiological and neurological limitations, children often produce many phonological processes (i.e., speech errors) until they reach adult-like speech production. It is known that phonological processes such as consonant or syllable deletion, consonant cluster reduction, and syllable reduplication occur until the age of 8 years.

Stress and intonation patterns emerge in children's babbling around 10 months of age, and children typically produce their first words at around 12 or 14 months (Owens, 2005). In English-speaking children, stops such as /p, b, m, n/ are acquired first by age 3 and fricatives such as /v, θ, ð, ʤ/ are acquired as late as age 8 (Dodd et al., 2003; Poole, 1934; Smit et al., 1990; Templin, 1957; Wellman, Case, Mengert, & Bradbury, 1931). Recently, Dodd, Holm, Hua, and Crosbie (2003) recruited 684 children aged between 3 and 6 years old in order to determine norms for phonological development in English-speaking children in the U.K. They analyzed children's spontaneous speech and specified ages when children were 90% correct in producing certain consonants: namely, (a) stops and some fricatives, such as /p, b, t, d, k, g, m, n, ŋ, f, v, s, z, h/, by age 3;5, (b) approximants, such as /w, y, l/ by age 3;11, (c) affricates and one fricative, such as /ʒ, ʧ, ʤ/, by age 4;5, (d) the fricative /ʃ/ by age 5;5, (e) the approximant /r/ by age 6;5, and (f) fricatives such as /θ, ð/ around 6;11 or later. According to many investigators, the English consonants /θ, ð/ are late-acquired consonants, but they should be mastered by the age of 8 at the latest (by the age of 6 years in Wellman et al., 1931; by the age of 7;6 years in Poole, 1934; and by the age of 8 years in Smit et al., 1990).

So far, I have reviewed speech perception and production in English monolingual children. Studies of speech perception showed that infants initially show different perceptual patterns from adults and then later develop adult-like speech perception during early childhood (Nittrouer, 1992; Nittrouer & Miller, 1997; Sussman & Carney, 1989; Sussman, 1993). Nevertheless, it is surprising that not many studies have addressed how the development of speech perception is completed during the school years, up to the point where it becomes adult-like. Furthermore, speech perception in bilingual children has not yet been extensively studied in the current literature. As it is, the present study, which targets speech perception in school-aged bilingual children, will shed light on how speech perception develops when a child is exposed to more than one language.

In order to understand bilingual development of speech perception and production, it is important to consider the similarity of the two phonological systems. The following review of literature addresses how Korean is phonologically and phonetically similar to or different from English, and then discusses which English sounds can be considered similar or new to Korean.

## *Korean Consonants and Vowels*

### *Korean Consonants*

Compared to English, Korean has many stop consonants, including nine oral stops and three nasal stops, and a relatively small number of fricatives, including three fricatives. Korean has two glides, /w/ and /j/, as English does. Korean, however, makes no contrast between /l/ and /r/: rather the two liquids are incorporated as allophones of the same phoneme, alternating /l/ with /r/ in the onset position (S-C. Ahn, 1998; S-B. Cho, 1967). Table 1 is a presentation of the phonemic inventory of Korean consonants (see Appendix B for the phonemic inventory of

English consonants), with the corresponding Korean alphabet symbols printed below each IPA transcription.

In contrast with English which has voiced counterparts for voiceless obstruents, with voiceless /p, t, k, s, ʧ/ and voiced /b, d, g, z, ʤ/ obstruents, Korean is thought by some to lack a voicing contrast among obstruents. Interestingly, Korean obstruents are considered to be distinguished from one another in terms of a laryngeal contrast, resulting in three categories: aspirated /$p^h$, $t^h$, $k^h$, $ʧ^h$/, lenis /p, t, k, s, ʧ/, and fortis /$p^*$, $t^*$, $k^*$, $s^*$, $ʧ^*$/ (T. Cho, S-A. Jun, & Ladefoged, 2002; T. Cho & Keating, 2001; Dart, 1987; J-I. Han, 1996a; M.S. Han & Weitzman, 1970; Hardcastle, 1973; Hirose, C-Y. Lee, & Ushijima, 1974; S-A Jun, 1993; Kagaya, 1974; C-W. Kim, 1965, 1970; M. Kim, 2004; M-R. Kim, 1994; Lisker & Abramson, 1964).

The laryngeal contrast in Korean obstruents is different from English, in which a voice contrast is found in a minimal pair such as *pin* and *bin*. In Korean, for example, /$p^h$aŋ/, /paŋ/, and /$p^*$aŋ/ comprise a triplet, which mean *bang* (onomatopoeia), *room*, and *bread* in Korean, respectively. Examples of triplets of Korean obstruents are presented in Table 2. This unique contrast system for Korean obstruents has been investigated in numerous studies in terms of the following articulatory, acoustic, and aerodynamic characteristics: laryngeal gesture, linguopalatal contact, voice onset time (VOT), fundamental frequency ($F_0$), pitch association, and intraoral pressure ($P_0$).

C-W. Kim (1965) claimed that Korean word-initial stops are all voiceless, contrary to studies by Junker (1955) and H-P. Choi (1954), and that these voiceless stops can all be differentiated from one another in their articulation according to a “tensity” feature of the larynx. The autonomy of the tensity feature was supported by fiberscopic studies which showed three different configurations of the laryngeal gesture (S-A. Jun, Beckman, & H. Lee, 1998; Kagaya, 1974) and by an electromyographic study which showed three different timings of Korean obstruents for suppression of adductor muscle activity in relation to the articulatory release (Hirose, C-Y. Lee, & Ushijima, 1974).

Kagaya (1974) argued that Korean aspirated and fortis stops can be characterized by positive laryngeal gestures performed prior to release of the oral constriction for the obstruents. These laryngeal gestures include positive abduction of the vocal folds when producing aspirated stops and positive adduction of the stiffened vocal folds when producing fortis stops, followed by instant occlusion at the time of the articulatory release (see Keating, Westbury, & Stevens, 1980, for description of the articulatory release). During the production of fortis obstruents, the release occurs just before the onset of voicing, and then prompt glottal relaxation occurs right after the release. During the production of aspirated obstruents, prompt glottal relaxation is made when the vocal folds are maximally opened. In contrast, during the production of lenis obstruents, the glottal opening is more or less continuous and no positive glottal gesture is created (also, cf. S-A. Jun et al., 1998). Hirose, C-Y. Lee, and Ushijima (1974) demonstrated that thyroarytenoid muscle activity increases sharply before the release of fortis stops but not before the release of aspirated and lenis stops.

Linguopalatal contact, i.e., the contact of the tongue against the hard palate, is another articulatory parameter which distinguishes three different types of Korean obstruents. In T. Cho and Keating (2001), electropalatography (EPG) was used to measure peak linguopalatal contact calculated as percentiles, where 1 percentile contacted corresponds to contact with any 1 out of 96 electrodes attached to a custom-fabricated pseudo-palate. Overall, the percentile at peak linguopalatal contact was greatest for fortis stops but smallest for lenis stops, across various prosodic boundaries.

In general, voice onset time (VOT)—the time interval between the release of a stop and the onset of voicing—has been discussed as an important acoustic cue for distinguishing lenis stops from fortis and aspirated stops (T. Cho 1996; M.S. Han & Weitzman, 1970; J-I. Han, 1996a; Hardcastle, 1973; S-A. Jun 1993; C-W. Kim, 1970; M-R. Kim, 1994; Hirose et al., 1974; Lisker & Abramson, 1964). Nevertheless, the aural and perceptual distinction among the three types of obstruents is difficult to make based on the VOT difference alone, due to considerable overlap between lenis and fortis stops. The results of acoustic studies of Korean obstruents show that fundamental frequency ($F_0$) at the vowel onset provides an additional acoustic cue to distinguish lenis from aspirated and fortis consonants (T. Cho, 1996; H. Choi, 2002; J-I. Han, 1996; Hardcastle, 1973; M-R. Kim, 1994; M-R. Kim, Beddor, & Horrocks, 2002). It appears that $F_0$ values are relatively higher at the onset of vowels after fortis obstruents than $F_0$ values at the onset of vowels after aspirated and lenis obstruents. Increased vocal fold tension due to cricothyroid muscle movement usually results in small but statistically significant increases in $F_0$ immediately before and after the closure interval, whereas, $F_0$ measurements at the same points for voiced cognates in English show no significant increases (K.E.A. Silverman, 1987). In addition to VOT and $F_0$, pitch contours may contribute to identification of the three different types of Korean obstruents. Fortis and aspirated obstruents are associated with high tones for the following vowel, but lenis obstruents are associated with low tones (S-A. Jun, 1993; M-R. Kim & Duanmu, 2004).

Figure 1 is an illustration of how the three types of bilabials in a triplet, e.g., /pʰaŋ, paŋ, p*aŋ/, are acoustically different from one another. Figure 1 consists of acoustic waves and spectrograms of each sampled word, e.g., /pʰaŋ, paŋ, p*aŋ/. This sample of a triplet was produced by a female Korean speaker. VOT, $F_0$, and pitch contours are displayed in three annotation tiers. The uppermost tier provides the VOT interval measured in milliseconds; the middle tier indicates the $F_0$ value at the point of the onset of voicing; and the tier at the bottom shows the pitch contours associated with the following vowels. The VOT value for aspirated /pʰ/ is the longest, and the VOT for fortis /p*/ is much shorter than for lenis /p/. The $F_0$ value of aspirated /pʰ/ is about 43 Hz higher than for lenis /p/, but it is only 11 Hz higher than for fortis /p*/. Pitch contours, represented as a blue line on the spectrogram, can be labeled as high-low (HL), low (L), and high-low (HL), respectively, following the Korean Tones and Break Indices (abbreviated as K-ToBI; S-A. Jun, 2000). The pitch contours show that aspirated /pʰ/ and fortis /p*/ are associated with high tones, as most voiceless consonants in word-initial position would be. Nonetheless, unlike in other languages—and perhaps only in Korean—voiceless lenis is associated with a low tone in word initial position (see M-R. Kim & Duanmu, 2004, for the argument that Korean lenis obstruents are underlyingly voiced). Note that the sample in Figure 1 should not be considered representative, as it is taken from just one token of each obstruent produced by only one Korean speaker.

Aerodynamic studies of Korean obstruents show that intraoral pressure ($P_0$) before the release of fortis obstruents is higher and air flow after the release of fortis obstruents is lower than after the release of corresponding lenis consonants (T. Cho et al., 2002; Dart, 1987; N. Han, 1998; Silverman & Jun, 1994). Intraoral pressure before the release of aspirated obstruents is lower than for fortis obstruents, but air flow after the release is the greatest among the three types of obstruents. In summary, fortis obstruents are differentiated from the aspirated and lenis ones in terms of a positive laryngeal gesture, a very short positive VOT, a higher $F_0$, association with a high tone on the ensuing vowel, high intraoral pressure, and low air flow.

*Korean Vowels*

Colloquial Korean has a very symmetric vowel system with three front vowels /i/, /e/, and /ɛ/, two central vowels /ɨ/ and /ʌ/, two back vowels /u/ and /o/, and one low vowel /a/. Until the early twentieth century, Korean had 10 pure monophthongs; however, two rounded front vowels /y, ø/ are now changed into /wi/ and /wɛ/, leaving only 8 monophthongs. The phonemes /j/ and /w/ are considered by Korean linguists to be glide components of diphthongs only, rather than separate consonants (B. Yang, 1996). Unlike English, Korean is considered to have a long-short distinction but not a tense-lax distinction. All eight pure vowels are articulated with either long or short duration but the change from long to short duration is not accompanied by significant spectral or acoustic changes.

In order to discern differences in acoustic quality of vowels within- and across-languages, B. Yang (1996) normalized natural tokens of 8 Korean vowels and 12 English vowels[2],

[2] Peterson and Barney (1952) also suggested formant measures for American English vowels based on perceptual tasks performed by 70 listeners (see Figure 3 on page 177) and on production tasks performed by 76 speakers (see Figure 8 on page 182). The presentation of vowel formants produced by 76 speakers looks different from the one suggested in Yang (1996). This is probably because more than half of the participants in Peterson and Barney's study were from the Middle Atlantic region of the U.S., whereas all of the 20 participants in Yang's study were from the South or Southwest. In addition, unlike in the Yang study, vowel production in the Peterson and Barney study was not normalized for non-linguistic differences such as vocal tract length, fundamental frequency, the ratio of the front cavity to the back cavity, and gender difference. The suggested formats for production of American English vowels are noticeably different in the two studies and warrant further investigation.

eliminating all the linguistically irrelevant differences such as vocal tract length, speech rates, age, emotional state of the speaker, etc. Figure 2 is a display of normalized vowel formants for 12 American-English vowels produced by 30 American-English-speaking females and 8 Korean vowel formants produced by 30 Korean-speaking females from Seoul. The formants are displayed in a Mel scale (Fant, 1973), which shows perceptual distances among vowels.

The present study will focus on three English vowels, /ɪ, ɛ, æ/, combined with 10 consonants, /p, k, s, ʧ, b, g, f, θ, z, ʤ/. In the study of H. Cho, Y. Mo, and S-Y. Yoon (2004), all English vowels /ɪ, ɛ, æ/ produced by native speakers of Korean were merged into /ɛ/, possibly due to the absence of an English /ɪ/-like and /æ/-like vowels in Korean. As seen in Figure 2, English /æ/ and /ɪ/ are absent in the Korean vowel system and acoustically new to Korean. On the other hand, English /ɛ/ is similar to Korean /ɛ/. Korean /e/ and /ɛ/, however, are perceptually very close on the height dimension, i.e., for $F_1$ in mels, as well as on the front-back dimension, i.e., for $F_2$ also in mels. Ingram and Park (1998) showed that modern Korean has only 7 pure monophthongs because the vowels /e/ and /ɛ/ are no longer distinguishable in the speech production of younger generations.

*Familiar versus Unfamiliar English Sounds*

In this document, the term, *familiarity*, will only be used to refer to shared phonemic status in both L1 and L2 sounds as well as acoustic and articulatory similarity. In other words, the grouping of sounds as familiar and unfamiliar sounds is based on similarities such as articulatory and acoustic characteristics of each sound at the *phonemic* level. Familiar sounds refer to English sounds that are present or similar to sounds in the Korean phonemic inventory, such as the English consonants /p, t, k, s, h, ʧ, m, n, ŋ, r, l, w, j/ and the English vowels /i, ɛ, a, o, u/. Unfamiliar sounds, on the other hand, refer to English sounds that are absent from the Korean phonemic inventory or dissimilar to Korean sounds. They include the English consonants /b, d, g, f, v, θ, ð, ʃ, ʤ/ and the English vowels /ɪ, e, æ, ɔ, ʊ/. Familiar and unfamiliar English sounds as compared to Korean are presented in Table 3. To classify sounds as familiar or unfamiliar, first I grouped them at the phonemic level, excluding allophonic variations. For example, the /b, d, g/ sounds exist in Korean, but only as allophonic variations of /p, t, k/ in an intervocalic or voicing context. As such, I consider /b, d, g/ to be *phonemically unfamiliar* sounds in the present study, because Korean does not have phonemes that correspond to English /b, d, g/.

Second, I caution the reader that phonemically familiar sounds do not necessarily share exactly the same acoustic or articulatory characteristics (though they must be similar). For example, Korean /i/ is slightly lower in F1—which is related to height of vowels—and in F2—which is related to frontness of vowels—than its English counterpart (please refer to Figure 2); however, it shares many commonalities in phonological features with English /i/. Therefore, I would consider the English /i/ to be a familiar vowel. Unlike the Korean and English /i/ vowels, Korean /e/ is different from English /e/, even though English /e/ may sound very similar to Korean /e/ to phonetically untrained ears. Korean /e/ is slightly lower and less front than English /e/ (please refer to Figure 2) and occurs only as a monophthong, not as a part of the diphthong /e͞ɪ/. Therefore, particularly due to the monophthong-diphthong difference, I would consider English /e/ to be a phonemically unfamiliar vowel. It should be acknowledged that classifying vowels as familiar or unfamiliar may appear to be subjective to some extent. Nevertheless, any misclassifications should become apparent from subsequent failure to find significant differences in behavioral responses to the two stimulus categories.

Lastly, all familiar sounds do not necessarily have the same degree of similarity, i.e., some familiar sounds in the two languages are more similar and some are less similar. Korean /p, t, k/ are roughly the same as English /p, t, k/, while Korean /ʧ/ is much more front and less fricated than English /ʧ/ (H. Kim, 1999). Nevertheless, I consider Korean /ʧ/ as similar enough to English /ʧ/ to be counted as a familiar sound, not considering phonetic differences. Next, I will review how different theories explain the role of L1 influences and L2 experience in L2 speech perception and production.

## *Speech Perception and Production of a Second Language*

### *First Language Influences and Second Language Experience*

*Child and adult second language learners*. Numerous findings for young infants in cross-linguistic speech perception and production studies support the idea of "the younger, the better" in language learning. These studies report that young children can discriminate nonnative sound pairs better than adults (Best & McRoberts, 2003; Eimas, Siqueland, Jusczyk, & Vigorito, 1971). Also, bilinguals who are exposed to L2 earlier in life produced L2 sounds (or sign language) with higher accuracy than bilinguals who are exposed to L2 later in life.

Many studies have shown that proficiency in L2 is closely related to age of acquisition of L2, which favors a "critical period" hypothesis; however, the exact age range corresponding to the critical period has been controversial. Bilinguals who arrive in an L2-speaking community before the age of 7 generally show native-like speech production (Johnson & Newport, 1987; Ploog, 1984; Yeni-Komshian, Flege, & Liu, 2000). Yeni-Komshian, Flege, and Liu (2000) examined Korean-English bilinguals and found that those who arrived in the United States between the ages of 1 and 5 years produced English close to that of individuals who spoke English as an L1. Conversely, those who arrived between the ages of 12 and 23 years spoke with a heavy foreign accent, showing great variability in L2 proficiency.

An individual who learned L2 between ages 7 and 15 may not reach native-like form, with a negative correlation between age of acquisition and performance (Johnson & Newport, 1987; Ploog, 1984). Mayberry and Eichen (1991) examined the effect of age of acquisition in forty-nine users of American Sign Language (ASL). Participants comprised three groups depending on the age of acquisition of ASL. Sixteen "native" signers were exposed to and learned ASL from infancy, 20 "childhood" signers were exposed to and learned ASL from the ages of 5 to 8 years, and thirteen "adolescence" signers were exposed to and learned ASL from ages of 9 to 13 years. Participants performed two types of tasks: a recall task of 8 long ASL sentences and a second recall task of 14 single-signed digits of ASL. Overall, recall accuracy was greater for native signers than childhood and adolescence signers. Native signers produced more instances of bound morphemes than childhood signers who, in turn, produced more instances than adolescence signers. Also, native signers performed better than childhood and adolescence signers at various linguistic levels of their responses to the sentence recall task, such as morphology, semantic paraphrasing, and syntactic grammaticality. Therefore, the results of this study confirm that age of acquisition affects proficiency in a language, such as sign language.

Different processing in L2 phonology is often attributed to L1 influences. First language (L1) influences are more likely to be strong in adult L2 learners than in child L2 learners, because L1 phonemic categories become more robust with increasing L1 experience. Baker, Trofimovich, Flege, Mack, and Halter (2008) examined child-adult differences in L2 phonological learning. Participants were 16 adult and 16 child bilingual speakers of Korean and English, with a length of residence (LOR) in the U.S. of 6 to 9 months. The mean age of the children was 10;0. Participants were asked to listen to each of the English vowels /i, ɪ, u, ʊ/ in three different CVC templates (e.g., *b_t/k*, *n_t/k*, and *h_d*). They were then forced to classify each vowel as one of 10 Korean vowels, /i, e, ɛ, a, o, u, ɨ, ʌ, y, œ/[3] and provide a goodness-of-fit-rating (from 1 for dissimilarity to 7 for similarity) for each classification.

Results of the classification task showed that significantly fewer tokens of a given English vowel were mapped to one of the 10 Korean vowels by children than in adults. The trend of classification, however, was similar between children and adults. In both of the groups, the English vowels /i, ɪ/ were mapped to the Korean vowel /i/ and the English vowels /u, ʊ/ were mapped to the Korean vowels /u/ or /ɨ, ʌ/. Also, goodness-of-fit-ratings for classifying each stimulus were slightly lower in children than in adults. Results of these two measures suggest that children are somewhat more sensitive than adults to differences between Korean and English. An English production task and discrimination task were also given to all participants. Production accuracy for the lax vowels /ɪ, ʊ/ was higher in children than in adults. Unlike the classification task, perceptual accuracy scores for the /i - ɪ/, /u - ʊ/, and /i - u/ contrasts were not significantly different between children and adults. In other words, the manner of perceiving English sounds differs between children and adults, but the ability to discriminate English tense-lax contrasts does not. In summary, Baker and colleagues found higher accuracy for L2 English

[3] In modern Korean, rounded vowels /y, œ/ have become diphthongized as /wi and wɛ/; however, Korean still has an alphabetic symbol or "a letter of the alphabet" for each vowel.

vowel production and lower goodness-of-fit ratings (for English to Korean vowels) in children than in adults.

To account for child-adult differences in their study, Baker et al. (2008) argued that 10-year-old children are in a *developmental stage of L1 categorization* which gives more ability (or plasticity) to dissimilate L2 English sounds from L1 Korean sounds. Therefore, children in the study produced L2 sounds more native-like and perceived L2 sounds as differing more from L1 sounds than adults did. Nevertheless, discrimination accuracy was not significantly different between children and adults, where both groups had had only minimal exposure to L2 English. The investigators commented that the relative easiness of the discrimination tasks may have obscured discrepancies between children's and adults' perception and that more complicated perceptual tasks are needed for future research. The present study asks participants to identify each stimulus as 1 of 30 English nonsense syllables, produced by 18 different talkers, which includes tasks complicated and difficult enough to discern children's perception from adults. Therefore, I believe the present study heeds the advice of Baker and colleagues. Based on the findings of Bakers and colleagues, I expect KEB children to perceive and produce English sounds more accurately than KEB adults, considering the lesser influence of L1 on L2 perception and production in children.

Nonetheless, some studies argue that there is not sufficient evidence to support an advantage for youth or a critical period hypothesis, arguing that the age of onset of L2 acquisition is not a sufficient measure for predicting one's native-like fluency (Abu-Rabia & Kehat, 2004; Bialystok & Miller, 1999; Birdsong, 1992; Flege, 1987). In a series of studies, Flege has argued that the critical period hypothesis is not applicable to L2 acquisition, considering patterns of language learning in L2 acquisition are different from those in L1 acquisition. In his reply to Flege (1987), Patkowski (1990) argues that the critical period for L2 acquisition should be interpreted as a "sensitive" period during which an L2 learner can learn L2 efficiently and reach native-like proficiency. He also argued that gradual changes in L2 are observable even after the critical period, in accordance with the amount of L2 experience. For such reason, "the sensitive period" for L2 acquisition is more preferred to "the critical period," as defined by Patkowski (1990).

Other factors suggested as affecting one's L2 fluency include length of residence in an L2-speaking community, amount of L2 usage at home or at school, amount of exposure to L2, and even one's personality. Several studies by Flege focused more on L2 learners' experience with L2 such as length of residence (LOR), rather than the age of L2 acquisition alone, to better predict proficiency in L2 production and perception. Flege and McKay (2004) examined discrimination of Canadian English vowel pairs by native Italian speakers who had lived in Canada just for 3 months and compared their performance to native Italian speakers who had lived in Canada for about 42 months and were exposed to L2 English at early or later ages. The result of discrimination performance on L2 English sounds showed a substantial difference between native Italian speakers with a smaller LOR and native Italian speakers with a greater LOR. Nonetheless, no significant difference was found between early L2 learners and late L2 learners, suggesting the age of arrival (AOA) only cannot predict one's L2 proficiency.

*Experienced and inexperienced second language learners.* Cross-linguistic effects are often observed for task performance at various linguistic levels such as syntax, semantics, morphology, phonology, and so on. The degree of cross-linguistic effect seems to vary depending on the age of L2 acquisition or the amount of L2 experience (Gildersleeve-Neumann, Kester, Davis, & Pena, 2008; Goldstein, Fabino, & Iglesias, 2003; Keshavarz & Ingram, 2002; Schnitzer & Krasinski, 1994, 1996). Overall intelligibility of L2 appears to be higher in younger and more experienced (or proficient) L2 learners than older and less experienced ones.

Goldstein and Washington (2001) examined cross-linguistic effects in speech production of Spanish and English single words in twelve 5-year-old Spanish-English bilinguals. Participants were asked to produce 28 Spanish words and 26 English words. Patterns of phonological processes in each language, such as cluster reduction, final consonant deletion, and so on, showed cross-linguistic influences, where production of some English words was influenced by Spanish, or vice versa. In order to extend the study of Goldstein and Washington (2001) and to examine cross-linguistic effects, Goldstein, Fabiano, and Washington (2005) examined a total of 15 bilingual children with a mean age of 5;2, who spoke both Spanish and English. Participants were identified as belonging to three groups: five predominantly Spanish-speaking children (PS), five predominantly English-speaking children (PE), and five children who spoke Spanish and English equally well (i.e. balanced "bilinguals"). Participants were asked to produce single words in both Spanish and English. The words conformed to various syllabic templates (e.g., CV, VC, CCVC, VCC, CCVC, etc.). Percent correct of consonants and syllables, both in Spanish and English was calculated and examined for correlation with language input reported by the participants' parents. Also, the use of phonological processes was examined for any indication of cross-linguistic effects. All measures were compared for bilinguals vs. the PS group and for bilinguals vs. the PE group.

Results showed that the accuracy of consonants and syllables in Spanish was not significantly different between bilinguals and the PS group. Only Spanish stops differed significantly between bilinguals and the PS group (93% vs. 87% correct, respectively). No differences between bilinguals and the PE group were significant. Despite the lack of significant findings for most measures, the authors did interpret some trends as indicative of cross-linguistic effects. Overall, accuracy of Spanish spirants and syllables in bilinguals was higher than in the PS group. Percent correct for English affricates in bilinguals was slightly lower than in the PE group. Percent correct in bilinguals was slightly lower than in the PE group for English VCC syllables (i.e., syllable final clusters), but slightly higher for English CCVC syllables. Correlations between language input and overall accuracy were not significant. The investigators argued that parental report about language input does not seem to be very reliable.

Even though cross-linguistic effects were not statistically significant between the groups (i.e., bilinguals vs. the PS group and bilinguals vs. the PE group); some trends were found in the use of phonological processes. Bilinguals produced Spanish stimuli with more instances of final voicing, but fewer instances of consonant cluster reduction and spirantization than the PS group. Bilinguals produced English stimuli with more instances of final voicing and final consonant deletion, but fewer instances of consonant cluster reduction, fronting, and backing than the PE group. The investigators attributed these observed differences between the groups of bilinguals to cross-linguistic effect.

Even though Goldstein and colleagues found few significant difference between groups, there is a need for studies that tease apart degree of L1 influence as a function of one's L2 experience (or L2 proficiency; Paradis, 2001). More recently, Gildersleeve-Neumann, Kester, Davis, and Pena (2008) investigated cross-linguistic effects among English-speaking monolingual (E), predominantly English-speaking bilingual (PE), and balanced English-Spanish bilingual (ES) children. In their study, 33 children, ages 3;1 to 3;10, produced 65 words as spoken responses to a picture naming task. Grouping of the participants was based on parents and teachers' report on language exposure and usage, speech clinicians' ratings, and assessment of spoken language. Overall, production accuracy for vowels was greater than for consonants (with 86% and 68% accuracy, respectively). Group differences were not significant for overall or vowel accuracy, but were so for consonant accuracy. The E group made significantly fewer consonant errors than the PE group (with 78% and 71% accuracy, respectively) and the ES group (with 55% accuracy), but the PE group performed similarly to the ES group. Thus, the degree of L2 experience may predict production performance for consonants (but not for vowels). It is worth noting that consonant accuracy was divergent among the groups, unlike vowel accuracy, which was similar. This may imply that consonant production is more affected by cross-linguistic influences than vowel production.

A number of studies of bilingualism have examined how an L1 interferes with the production and the perception of an L2, how the process of language learning differs for bilingual adults and monolingual adults, and which sound groups may cause L2 adult learners more difficulty in learning. The most influential frameworks in bilingual research are Flege's Speech Learning Model (SLM; Flege, 1987, 1992; Flege, Takagi, & Mann, 1995) and Best's Perceptual Assimilation Model (PAM; Best, McRoberts, & Sithole, 1988; Best, 1994, 1995).

*Speech Learning Model*

In the 1960s, the Contrastive Analysis Hypothesis (Lado, 1957) was extensively used to provide a theoretical explanation of why some features of an L2 were more difficult to learn than others. Lado (1957) claimed that L2 learners will have more difficulty learning L2 sounds which have no equivalent in the L1 than learning L2 sounds which are similar or equivalent to those in L1. Contrary to the Contrastive Analysis Hypothesis, Flege (1987) showed that English-speaking learners of French can pronounce dissimilar or new French vowels with a more native-like pronunciation. English /u/ is similar—but not identical—to French /u/, as in the French word *tous* "all." The English /u/ has a higher $F_2$ than the French /u/. However, English lacks a high front rounded vowel /y/, as in the French word *tu* "you." In Flege's study, English speaking L2 (i.e., French) learners' production of a French /y/ was compared to that of a French /u/. The results showed that a French /y/ produced by L2 French learners of English-speaking adults was much closer to the French norm than a French /u/ in terms of acoustic characteristics. Flege explained that L2 learners fail to reach the L2 production norm for similar sounds because the learner does not readily notice minor differences between L1 sounds (e.g., English /u/) and target L2 sounds (e.g., French /u/). This finding led to SLM (Flege, 1992; Flege, Takagi, & Mann, 1995; see Flege, 2002 for subsequent refinements), which argues that a new perceptual category is developed for previously unknown sounds over the course of training in L2 and that such a category is more L2 native-like than a perceptual category for similar L2 sounds.

SLM assumes that L2 learners often fail to notice sub-segmental (or non-phonemic) features in an L2 segment and assimilate the segment to the most similar L1 segment. This assumption leads to a prediction that new or dissimilar L2 sounds will be perceived and produced more accurately than similar, but not identical, L2 sounds. At the initial stage of L2 learning, however, new L2 sounds may be more difficult for inexperienced L2 learners to perceive and produce than similar L2 sounds, due to the absence of categories or representations for the new L2 sounds. With the increase in L2 experience, L2 learners may become able to perceive and produce new L2 sounds more accurately than similar L2 sounds due to the development of new categories for the new L2 sounds.

In support of SLM, Flege, Bohn, and Jang (1997) examined the accuracy of English /i, ɪ, ɛ, æ/ production and perception by inexperienced and experienced nonnative English speakers (the latter with a mean of 7.3 years of residence in the U.S.). The nonnative English participants comprised four language groups—German, Spanish, Mandarin, or Korean—with 10 experienced and 10 inexperienced participants per group. The participants were asked to produce a list of words in a carrier phrase, containing synthetic /i, ɪ, ɛ, æ/ vowels. The vowels produced by the participants were then identified by native English speakers, who were asked to indicate which vowel of the four they perceived. Also, acoustic analyses were done on the productions of the four groups. The results of the production task showed that experienced Korean participants produced the English vowel /ɪ/ better but the vowel /i/ worse than inexperienced Korean participants as predicted by SLM. Both experienced and inexperienced Korean participants depended upon a durational difference to distinguish /i/ from /ɪ/, and /ɛ/ from /æ/, in production.

In a perceptual task, participants were asked to discriminate one sound from another along a *beat – bit* (i.e., /i/ vs. /ɪ/) continuum and a *bet – bat* (i.e., /ɛ/ vs. /æ/) continuum, with 11 spectral steps and 3 durational steps. Both Korean groups depended more on durational differences than spectral differences to identify synthetic English /i/-/ɪ/ and /ɛ/-/æ/ contrasts in perception. Recall that Korean does not have a tense-lax vowel distinction, unlike English which distinguishes tense /i/ from lax /ɪ/, tense /e/ from lax /ɛ/, and tense /u/ from lax /ʊ/. Although the between-group difference was not significant, an interaction between experience and vowel continuum was significant, $F_{3, 216} = 3.30$, $p = .02$. Furthermore, experienced Korean participants were slightly better than inexperienced ones at using spectral information to identify the /ɪ/ and /ɛ/ endpoints on the perceptual continua. SLM predicted that L2 experience would facilitate development of a new representation for a new L2 sound, which would help experienced L2 learners produce and perceive it.

To examine how L1 influences perception and production of L2 sounds, Flege and MacKay (2004) examined how L2 English vowels are perceived and categorized into L1 Italian sounds. In their study, nine native Italian speakers performed two different tasks: an oddball task and a classification task with goodness-of-fit-ratings. Performance in native Italian speakers on the oddball task was compared to performance of twelve native English speakers. The stimuli were the Canadian-English vowels /i, ɪ, e, ɛ, æ, ʌ, ɚ, ɒ/ in a CVC context produced by five different talkers. Italian has seven vowels, /i, e, ɛ, a, ɔ, o, u/, fewer than in English.

In the first experiment, participants were asked to identify the odd item in *change trials*, which were composed of two same vowels and one different vowel (e.g., a contrast between /bet/-/bet/-/bit/), and in *no-change trials,* which were composed of three same but audibly different tokens of a vowel (e.g., /bet/-/bet/-/bet/), in terms of talkers, prosody, and so on. Scores for nine vowel contrasts in change and no-change tasks were calculated as a proportion of correct selection in change trials to incorrect selection in no-change trials. Although all scores were above chance, overall, native English speakers performed significantly better than native Italian speakers. Scores for the Italian participants were the lowest for the /ɛ - æ/ contrast and the third lowest for the /ɪ - ɛ/ contrast (out of 9 contrasts). Italian participants reached near native-like scores for the English /ʌ - ɚ/ contrast—where both sounds are lacking in Italian. Scores for the English /æ - ʌ/ contrast—where both sounds are new to Italian—were also high (greater than .85 out of 1.0).

Results addressed so far confirm the predictions of SLM. Contrary to the predictions of SLM, scores for the Italian participants for the English /e - ɛ/ contrast—where both vowels are similar to L1 Italian vowels—were the second highest. Moreover, scores for the English /ɒ - ʌ/ contrast—where both sounds are new to L1 Italian—were the second lowest. This suggests that perceptibility may differ across different vowel contrasts, regardless of similarity or dissimilarity of L2 to L1. These results which conflict with the SLM may be better explained by another cross-linguistic perceptual model, called the Perceptual Assimilation Model (or PAM), developed by Best and her colleagues. Later in this document, I will discuss how Flege and MacKay's varying results may be better predicted by PAM.

In the second experiment of Flege and MacKay's (2004) study, participants were asked to perform a classification and rating task. In this task, an English vowel was presented, and an Italian participant was asked to select which Italian vowel, /i, e, ɛ, a, ɔ, o, u/, it most resembled. Next, the participant was asked to judge the degree of similarity between the given English vowel and the chosen Italian one. The investigators calculated percentages of classification type and average goodness-of-fit ratings. For example, Italian speakers judged English stimulus /i/ to be similar to the Italian vowel /i/ 87% of the time (i.e., the percentage of classification type) and rated the stimulus as very similar to the Italian /i/, with an average rating of 4.2 out of 5 (i.e., the goodness-of-fit rating). Goodness-of-fit ratings for classification of a given English stimulus ranged from 1 which poorly fits to the chosen Italian vowel to 5 for a stimulus which fits well. Results of the second experiment are presented in Table 4.

Results from Flege and MacKay's (2004) second experiment showed that the English vowels /ɒ, ʌ/ were judged to be similar to the Italian vowel /a/ by native Italian speakers. This classification may account for their poor performance on the /ɒ - ʌ/ contrast, assuming that perceptual assimilation of two L2 sounds into one L1 sound makes discrimination of such sounds extremely difficult. Furthermore, if one L2 sound corresponds to one similar L1 sound but the other L2 sounds corresponds to no sound in L1, then discrimination of a contrast between the two should be relatively easier. This hypothesis is supported by the following results. In the first experiment, the average score for the /ʌ - ɚ/ contrast was nearly perfect. In the second experiment, the English vowel /ʌ/ was perceived to be similar to the Italian vowel /a/ 93% of the time, but the English vowel /ɚ/ was not perceived to be similar to any Italian vowel. Finally, two L2 sounds that correspond to two different L1 sounds should be the easiest to discriminate. The average score for the /æ - ʌ/ contrast was the third highest (out of 9 contrasts) in the first experiment. The English vowels /æ/ and /ʌ/ were perceived to be similar to the Italian vowels /ɛ/ and /a/, respectively. Therefore, perceptual similarity between L1 Italian and L2 English seemed to predict performance on discrimination of this L2 contrast.

*Summary*. In Flege (1987), advanced English learners of L2 French produced the French vowel /y/, which is new to L1 English, more native-like than the French vowel /u/, which is similar to the L1 English vowel /u/. Results of this study led to the development of SLM, which argues that a new L2 sound can be produced and possibly perceived correctly more often than an L2 sound similar to one in L1 for experienced L2 learners. In Flege, Bohn, and Jang (1997), Korean participants' accuracy in production and discrimination of English vowels depended on durational differences more so than spectral ones. This was attributed to the fact that Korean vowel distinctions do not rely on spectral effects, but only on durational ones, unlike the tense-lax distinction in English.

Not only does listener experience seem to play a role in bilingual perception, the nature of the vowels to be perceived also seems important. In Flege and MacKay's (2004) study, the /ɪ - ɛ/ and /ɛ - æ/ contrasts in Canadian-English were poorly discriminated by native Italian speakers. Italian has the vowel /ɛ/ but not the vowels /ɪ/ and /æ/. The two different English vowels /ɪ/ and /ɛ/ acoustically overlap with the Italian vowel /e/, and the two English vowels /ɛ/ and /æ/ overlap the Italian vowel /ɛ/. The vowel contrasts with /ɚ/ were discriminated with high accuracy because the English vowel /ɚ/ is dissimilar to any of Italian vowels and distinguishable from other English vowels. The predictions of SLM were consistent with some but not all of the results for the discrimination of Canadian-English vowel contrasts by Italian speakers.

Many studies in support of SLM have focused on perception and production of vowels, but SLM was also tested for perception and production of consonants, as well. SLM predicted the results of Aoyama, Flege, Guion, Akahane-Yamada, and Yamada’s (2004) study, which argued that perception (in a discrimination task) and production of the English /r/ by native Japanese speakers was significantly better than the English /l/, due to similarity of the English /l/

and the Japanese /r/[4]. Therefore, SLM argues that L2 learners have difficulty in using an absent subsegmental (i.e., phonetic) feature in L1, such as spectral information, when discriminating L2 contrasts, even though experienced L2 learners are slightly better at using subsegmental features than inexperienced L2 learners.

*Implications for the present study*. Taken together, in accord with SLM, L2 phones similar, but not identical, to L1 would be expected to be produced and perceived as their L1 counterparts. This yields the prediction that *familiar sounds* such as /p, k, s, ʧ, ɛ/, which are similar to their Korean counterparts, should be more challenging to perceive and produce than *unfamiliar sounds* such as /b, g, f, θ, z, ʤ, ɪ, æ/, which are new or dissimilar to Korean. SLM also argues that perception and production of new sounds (or unfamiliar sounds) may be more challenging to inexperienced L2 learners than experienced L2 learners. This argument yields a prediction that KMs will perform poorer on unfamiliar sounds than familiar ones, unlike experienced bilinguals whom I predict will perform in the opposite direction (i.e., poorer on familiar sounds than unfamiliar ones).

As for the relation of perception to the degree of L2 experience, SLM argues experienced L2 learners are generally better at using subsegmental differences to distinguish between L1 sounds and similar L2 sounds. Therefore, SLM predicts that KMs will perform poorer than KEBs on familiar sounds which require an ability to perceive and produce subsegmental differences. Further, even among KEB children, a child with more L2

---

[4] Takagi (1993) showed that the English /l/, rather than the English /r/, was rated to be closer to the Japanese /r/ by native Japanese speakers. For more details on Japanese /r/, see Best & Strange (1992), Komaki, Akahane-Yamada, & Choi (1999), and Price (1981).

experience—as determined by age of arrival, length of residence in the U.S., and degree of English proficiency—may perceive and produce the familiar sounds better than a child with less L2 experience.

As for the relation between L1 and L2 sounds, performance on L2 sounds seems to vary as a function of the degree of similarity to L1 sounds (Flege & MacKay, 2004). The English familiar vowel /ɛ/ corresponds to the Korean vowel /ɛ/. The English unfamiliar vowel /ɪ/ is likely to be classified as the Korean vowels /i/ or /ɛ/, and the English unfamiliar vowel /æ/ is likely to be represented as the Korean vowel /ɛ/ (H. Cho, Mo, & S-Y. Yoon, 2005). Confusion between /ɛ/ and /æ/—both of which are likely to be classified as /ɛ/—may occur more often than confusion between /ɪ/ and /ɛ/—which includes /ɪ/ that has a chance to be classified as /i/ instead of /ɛ/. I have briefly discussed what the predictions of SLM would be for the present study. It is worth noting that many studies that support SLM were done primarily on production (and perception, recently) of vowels but fewer on those of consonants. Next, I will review another model for L2 learning developed by Best and her colleagues, one that focuses more on perception than production, mostly, of consonants.

*Perceptual Assimilation Model*

In contrast to SLM, which attempts to account for speech production and identification of similar or dissimilar nonnative phones in bilinguals, the Perceptual Assimilation Model (PAM; Best, McRoberts, & Sithole, 1998; Best, 2001) provides a framework for predicting a listener's discrimination of nonnative contrasts mostly in monolinguals. PAM predicts that listeners will tend to assimilate nonnative phones to their native phonological and/or phonetic-articulatory categories whenever possible. According to PAM, the discrimination performance of nonnative contrasts will be better when perceptual assimilation does not occur.

Best, McRoberts, and Sithole (1988) compared the discrimination performance of Zulu click contrasts by 10- to 12-month-old listeners in English-speaking communities, English-speaking adult listeners, and Zulu-speaking adult listeners. English listeners performed best on a pair of voiced palatal /ɟa/[5] and voiceless aspirated palatal /ǂ$^h$a/[6] clicks, but worst on a pair of voiceless unaspirated apical /|a/ and voiceless unaspirated lateral /||a/ clicks. Overall, infants performed better than adults. Interestingly, the overall percentage of correct responses on 18 contrasts across voicing and place was greater than 80% even by English-speaking adults. The findings suggested that Zulu-click sounds are different enough not to assimilate into any English sounds, and that English monolingual adults use the voicing contrasts in English when discriminating the voicing contrast between a voiced palatal /ɟa/ and a voiceless aspirated palatal /ǂ$^h$a/ clicks.

Best, McRoberts, and Goodell (2001) replicated the 1988 experiment to refine their account of perceptual assimilation of nonnative contrasts. In this study, they included the phones in Zulu that are similar to English consonants. Based on articulatory-phonetic characteristics, three contrast pairs of Zulu were selected: (a) voiceless versus voiced lateral fricatives (/ɬ/-/ɮ/), (b) voiceless aspirated versus ejective (glottalized) velar stops (/k$^h$/-/k'/), and (c) plosive versus implosive voiced bilabial stops (/b/-/ɓ/). Twenty-two native English listeners were asked to

---

5 Given IPA symbols are at best approximates to the Zulu click sounds described here. Problems in transcribing Zulu clicks with current IPA symbols alone have been discussed in the literature (Roux, 2007). In Best et al. (1988), clicks are transcribed as /aʗa/ for the voiced palatal click, /ʗ$^h$a/ for the voiceless aspirated palatal click, /ʇa/ for the voiceless unaspirated apical click, and /ʖa/ for the voiceless unaspirated lateral click.

6 The symbol can be used either for palato-alveolar or for voiceless palatal clicks.

identify whether the middle item, X, was the same as the first, A, or the third item, B. This is called a categorical AXB discrimination test, where X is a different physical token than the A or B sound that it matches.

The results of a three way ANOVA showed that the discrimination between the /ɬ/ and /ɮ/ pair was significantly better than the other two pairs even though both /ɬ/ and /ɮ/ are absent in English. They argued that a voicing contrast exists in American-English, and American-English speakers applied the same contrasting feature in distinguishing /ɬ/ from /ɮ/. Zulu /$k^h$/ is roughly identical to English /k/—particularly /k/ in word-initial position or before a vowel—but ejective /k’/ deviates slightly from English /k/. Thus, the /$k^h$/-/k’/ pair of sounds would be considered by English listeners either as a single phone with one good exemplar and a second bad exemplar of the same or a different phone (i.e., two allophones for one or two phonemes) or both as viable options for English prevocalic /k/ (i.e., two variations for one phoneme). In contrast, the Zulu /b/ and /ɓ/ sounds may be perceived as one sound (i.e., a single allophone) through assimilation to each other. The results of the Best et al. study suggest that a pair of sounds which are similar to L1, but with a different type of distinction from L1, would be the most difficult to discriminate. On the other hand, a pair of sounds which are similar to L1 but different enough not to be completely assimilated to each other would only be somewhat challenging to discriminate, and a pair of sounds which are different from L1 but with the same type of distinction as in L1 would be the easiest to discriminate.

Within the framework of PAM, some of results of Flege and MacKay (2004) are better explained. In their study, the /ɚ - ʌ/ and /æ - ʌ/ contrasts were discriminated much more correctly than the /ɒ - ʌ/ contrast, where the four vowels /æ, ɚ, ʌ, ɒ/ are new to L1 Italian. All these vowels were derived either from /b_d/ or /k_d/ contexts and each syllable was equally stressed. The /ɚ - ʌ/ contrast can be easily discriminated because Italian does not have the rhotic (i.e., /r/ coloring) feature, which makes the vowel /ɚ/ very distinctive to Italian listeners. The /æ - ʌ/ contrast can be easily discriminated because Italian has a front-center contrast, as was the case for the /ɬ- ɮ/ contrast in Best et al.'s (2001). The /ɒ - ʌ/ contrast, however, cannot be easily discriminated because Italian lacks a mid-low contrast for center vowels. Italian has only one center vowel /a/; thus, both /ɒ/ and /ʌ/ may go through perceptual assimilation to Italian /a/, as was the case of the /b - ɓ/ contrast in Best et al.'s study of Zulu. Perceptual assimilation of the English vowels /ɒ, ʌ/ to one Italian vowel /a/ was supported by the result from the second experiment of Flege and MacKay (2004). These results were consistent with the claims of PAM, whereas many of the results discussed previously were consistent with the claims of SLM.

From a different point of view, poor performance on discrimination of certain nonnative contrasts may be induced by their inconspicuous nature rather than an L1 influence. If speech perception operates on general auditory operation or depends more on acoustic differences than phonological ones, then some nonnative contrasts should be difficult regardless of one's native language. For example, a Spanish /b/ On the other hand, if speech perception conforms to language-specific operation or depends more on phonological differences than acoustic ones, then some contrasts that are easy for native users to discriminate should be difficult for L2 learners.

To examine whether perceptual assimilation occurs at acoustic or phonological level, Halle and Best (2007) examined cross-linguistic perception of Hebrew consonant clusters by 12 native French speakers, 14 native American-English speakers, and 11 native Hebrew speakers.

Eight consonant clusters /dl, tl, gl, kl, dr, tr, gr, kr/ in Hebrew were selected as stimuli for AXB discrimination and categorization tasks. Each of consonant clusters was generated with one of three Hebrew vowels /i, a, u/ in a CCV template. The /dl/ and /tl/ clusters in word-initial position are phonotactically legal in Hebrew but not in French and English, and the remaining six clusters /gl, kl, dr, tr, gr, kr/ are legal in all three languages. Overall, French and English native speakers performed significantly poorer on discrimination of the /dl - gl/ and /tl - kl/ contrasts than Hebrew native speakers. Percent correct on the AXB discrimination task for French and English speakers ranged from 95% to 98% for the /dr - gr/ and /tr - kr/ contrasts, which are legal in French and English. In comparison, percent correct ranged only from 63% to 77% for the /dl - gl/ and /tl - kl/ contrasts, which are illegal in French and English. Hebrew native speakers discriminated all contrasts with 95% accuracy or above.

Results from the categorization task showed that French and English native speakers perceived the Hebrew clusters /dl/ and /tl/ as /gl/ and /kl/, respectively. French native speakers categorized /dl/ as /gl/ 29% of the time and /tl/ as /kl/ 81% of the time. Similarly, English native speakers categorized /dl/ as /gl/ 39% of the time and /tl/ as /kl/ 86% of the time (Such voicing asymmetry may be attributed to the distribution of the /dl – gl/ and /tl – kl/ contrasts in the world languages (Abdelatty Ali et al., 2001; Tobin, 2002), in which the latter contrast is more frequent. Results of this study suggest that perception of nonnative contrasts is greatly influenced by the effect of one's native phonotactic rules rather than sensitivity to acoustic cues. If nonnative contrasts are perceived at an acoustic level, then the /dl/ and /tl/ clusters, which were poorly perceived by French and English native speakers, should be more difficult even for Hebrew native speakers to discriminate. Also, if so, then the clusters should be easy for English native speakers because the /dl/ and /tl/ clusters exist in word-medial positions in English (e.g., *settle, cradle*) and the clusters must be familiar to English listeners. Nonetheless, English native speakers performed the worst on these unlike Hebrew native speakers who performed nearly perfect on these contrasts. This supports the idea that discrimination of nonnative contrasts is influenced by L1 phonology and seems to depend more on phonological or representational differences between two languages than on acoustic or physiological ease per se.

*Summary*. In Best, McRobers, and Sithole (1988), overall performance on the discrimination task for Zulu click contrasts was better in 10- to 12-month-old English-listening infants than English-listening adults. Interestingly, there was one pair of Zulu click contrasts that was discriminated with high accuracy even by English-listening adults. The investigators argued that a voicing contrast which also exists in English helped English monolinguals to discriminate new Zulu clicks. In Best, McRobers, and Goodell (2001), perceptual assimilation occurred more if two nonnative sounds were similar to one native sound, as was the case for the /b/-/ɓ/ contrast (i.e., plosive and implosive voiced bilabial stops). In Halle and Best (2007), the word-initial /dl/ and /tl/ contrasts in Hebrew, which are phonotactically illegal in French and English, were discriminated poorly by both French and English native listeners but nearly perfect by Hebrew native listeners. Results of this study suggest that discrimination of nonnative contrasts depends on phonological differences rather than acoustic differences and thus is greatly influenced by L1 phonology.

*Implications for the present study*. In the present study, familiar sounds such as /p, k, s, ʧ, ɛ/ are more likely to participate in sound contrasts similar to Korean contrasts than are unfamiliar sounds such as /b, g, f, θ, z, ʤ, ɪ, æ/. KEBs and KMs might be expected to use contrasts of manner and place of articulation similar to Korean ones when discriminating familiar sounds. On the other hand, discrimination of unfamiliar sounds may call for the use of contrasts of manner and place new or dissimilar to L1 Korean ones, such as lax vs. tense manner for vowels or dental vs. alveolar place for consonants. Therefore, PAM predicts that KEBs and KMs may show less confusion for familiar sounds than unfamiliar ones, and therefore, familiar sounds may be less challenging to identify than unfamiliar ones. Unlike SLM, PAM does not make any direct prediction either for production or performance as a function of L2 experience, because PAM is based mostly on perceptual performance on nonnative contrasts by monolinguals' listening to a FL. Nonetheless, relevant inferences to production performance can be drawn from the series of studies by Best and colleagues.

PAM predicts that a pair of nonnative sounds with an unfamiliar contrast will be more difficult to be discriminated than with a familiar contrast. Korean-English bilinguals who speak English as an L2 may be better acquainted with English contrasts than KMs, who speak English as a FL. Thus, KEBs should perform better on an English perceptual task than KMs. Studies of Best and colleagues showed that discrimination of nonnative contrasts is greatly influenced by one's L1 phonology, and such influences diminish discrimination accuracy. If one's amount of L2 experience determines the degree of influence from one's L1 phonology, then participants will show different patterns of sound confusion depending on their L2 experience. In the present study, L2 experience will be operationally defined as AOA, LOR, usage of each language at home or at work, as reported on a survey by parents and adults; and L1 influence will be operationally defined as higher performance for sounds familiar from L1 than unfamiliar ones. On the assumption that the influence of L1 is weaker in bilinguals than in monolinguals, KEBs should perform better on English perceptual task than KMs. Accordingly, KEB children with more L2 experience should perform better than those with less L2 experience.

Even though discrimination of nonnative contrasts appears to occur at the phonological level in adults, infants who are prelinguistic seem to depend more on acoustic differences than phonological ones. Many studies of infants' perception show that infants younger than 12 months discriminate nonnative contrasts with higher accuracy than adults (Best et al, 1987; 2001). This may be because infants are more adept at listening to acoustic differences than adults. On the assumption that older children between the ages 8 and 13 years are still developing toward adult-like perception, KEB children are expected to perceive unfamiliar sounds better than KEB adults, because children are better able to attend to acoustic differences in unfamiliar contrasts.

In summary, SLM argues that establishing a new category for L2 sounds, especially for L2 sounds that are similar but not identical to L1 sounds, is more difficult with increasing L2 experience than L2 sounds that are new to L1. On the other hand, PAM argues that children are tuned to categories for L1 sounds at so early an age that the perceptibility of L2 depends more on the degree of perceptual similarity among L1 and L2 sound features and the familiarity of certain types of phonemic distinctions than on the degree of L2 experience. Both SLM and PAM provide predictions about which L2 sounds would be more difficult for adult L2 learners to produce and perceive; however, SLM focuses particularly on similarity of L2 *phones* to L1 phones and their production and perception in bilinguals. In contrast, PAM focuses on similarity of L2 phone *contrasts* (between one phone and another in the same language) to those in L1, and their perception only. SLM argues that similarity of L2 phones to L1 phones will hinder accurate production and perception; whereas, PAM argues that similarity of L2 contrasts, which distinguish two L2 phones, to L1 contrasts will help accurate perception (and possibly production, with extended interpretation of PAM). Therefore, the two theories make opposite predictions about perception and production of familiar phonemes, for experienced L2 learners. Predictions from the two theories will be compared in detail in Chapter III.

In addition to SLM and PAM, Kuhl's Native Language Magnet model (NLM: Grieser & Kuhl, 1989; Iverson & Kuhl, 1996; Kuhl, 1992) also yields similar predictions. The latter proposes that acoustic prototypes developed early in life interfere with perception of nonnative phones through the perceptual magnet effect. The NLM theory, however, has been challenged by several studies that have shown the magnet-like effect even in the absence of acoustic prototypes and improvement of discrimination in adults with training (Lively, 1993; Lotto, Kluender, & Holt, 1998; Sussman & Lauckner-Morano, 1995).

As a child develops his or her language skills, his or her speaking and listening abilities become tuned to the language to which he or she is extensively exposed. Typically by ages 6 to 8, normally developing children display adult-like speech production in everyday, casual, connected speech; making no errors in pronunciation of their native language (Dodd, Holm, Hua, & Crosbie, 2003; Smit et al., 1990; Templin, 1957; Wellman et al., 1931). Nonetheless, their performance on certain speech perception tasks displays different patterns from adults' until age 13 (Elliott et al., 1986; Nittrouer, 1992; Nittrouer & Miller, 1997; Morrongiello et al., 1984; Sussman, 1993; Sussman & Carney, 1989). Some investigators have argued that children's VOT threshold to discriminate one voiced stop from another is higher than adults' (Elliott et al., 1986; Sussman & Carney, 1989; Sussman, 1993). Others have argued that children weigh available acoustic cues differently than adults (Morrongiello et al., 1984; Nittrouer, 1972; Nittrouer & Miller, 1997). The present study did not aim to determine which theoretical view best explains speech perception in bilingual children, rather, merely to examine whether bilingual children show different L2 perceptual patterns from adults, under the influence of L1.

Literature on speech perception and production of L2 sounds argued that difficulties in listening to or articulating L2 sounds may be predicted by relations between L1 and L2 phonology. In addition, a negative relation has been known to exist between the age of L2 acquisition (i.e., AOA) and proficiency of L2 production (Mayberry & Eichen, 1991; Ploog, 1984; Yeni-Komshian et al., 2000). In general, child L2 learners who are exposed to L2 before age 7 can achieve native-like fluency in L2; however, adult L2 learners typically attain lower L2 proficiency. Further, bilingual children who are exposed to an L2 prior to age 3 are generally considered to be simultaneous bilinguals, whose development of both languages is comparable to that of monolinguals (de Houwer, 1995; McLaughlin, 1978; Oller et al., 1997; Peña et al., 2003). On the other hand, bilingual children who acquire an L2 after age 3 are generally considered to be simultaneous bilinguals whose development of an L1 precedes that of an L2. Kohnert and Bates (2002) showed that successive (Spanish—English) bilingual children who acquire an L2 English between 5 and 7 years initially develop lexical comprehension skills but later, expressive skills in an L2. The investigators also showed that language proficiency of both languages is almost balanced in these children.

Some bilingual studies have argued against the idea of a critical period for language acquisition, showing no significant age effect on L2 proficiency (Bialystok & Miller, 1999; Epstein, Flynn, & Martohardjono, 1996). Difficulty in achieving native-like proficiency may be attributed to several factors other than AOA, such as amount of L2 exposure, environment of L2 usage, motivation to learn L2, and so on. In support of this, Flege and colleagues (1997) found that experienced L2 learners were better than inexperienced ones at perceiving and producing L2 sounds similar to L1 phonemes (i.e., familiar sounds), even though the experienced L2 learners perceived and produced *new* L2 sounds better than the similar ones. Therefore, KEB children with more L2 experience may perceive and produce familiar sounds better than KEB children with less L2 experience.

Considering all the findings which have been discussed so far, children between the ages of 8 and 13 years who were exposed at a young age to two very different languages, such as Korean and English, are expected to have better perception and production of L2 sounds than bilingual adults. It is worthwhile to investigate L2 perception and production in the school-aged population in order to discern which sound features of L2 are salient enough not to experience interference from L1 during L2 perception and production tasks. Unfortunately, there are very few studies—possibly none, to this author's knowledge—which examine speech perception and production of English by Korean-English bilingual children and Korean monolingual children in this age range. As it stands, the present study is intended to fill the gap between studies of speech perception for bilingual adults and infants, and the gap for speech perception in bilingual children.

After considering the previous studies, the present study will explore the issues related to speech perception and production in school-aged bilingual children and adults. The present study aims to investigate (a) the degree of English language (L2) proficiency achieved by younger (viz., school-aged) and older native speakers of Korean, (b) perception and production of English sounds as a second language or as a foreign language in children and adults who speak Korean as their first language; (c) whether KEB children, KEB adults, and KM children perceive and produce English nonsense syllables in a different fashion from one another; (d) whether perceptual and production performance are correlated; and (e) how L2 experience influences perceptual and production performance.

# CHAPTER II

# LANGUAGE SAMPLE ANALYSES

## Rationale, Research Question and Hypothesis

Previous studies of bilingualism have argued that L2 learners are more likely to achieve native-like proficiency when they learn an L2 before the age of 7 years (Johnson & Newport, 1987; Mayberry & Eichen, 1991; Ploog, 1984; Yeni-Komshian et al., 2000). Nonetheless, simultaneous bilingual children, who are exposed to L2 before age 3, achieve native-like fluency in a different way from successive bilingual children, who are exposed to L2 after age 3 (DeHouwer, 1995; McLaughlin, 1978; Oller et al., 1997; Peña et al., 2003). According to these studies, simultaneous bilinguals first learn two languages as a hybrid language and then begin to separate the two different language systems, with increasing L1 and L2 experience. On the other hand, successive bilinguals are thought to learn an L2 as a separate language system from the one learned as L1, from the beginning of language learning (Owens, 2005). Some studies argue that degree of L2 experience, rather than age of acquisition alone, predicts variability of L2 proficiency in successive bilinguals (Baker et al., 2008; Bialystok & Miller, 1999; Flege & MacKay, 2004). Depending on L2 experience, a bilingual individual may speak L2 with higher proficiency than L1, or vice versa.

Research Question: Will younger and more experienced L2 learners exhibit better English proficiency than older and less experienced L2 learners, based on evaluation of English language samples in terms of vocabulary, grammar, and speaking fluency?

Recent sstudies on bilingualism have focused more on language-dominance in bilinguals, classifying them as predominantly L1-speaking bilinguals, predominantly L2-speaking bilinguals, or balanced bilinguals (Goldstein & Washington, 2001; Goldstein et al., 2005; Gildersleeve-Newmann et al., 2008). These speech sound production studies have shown that predominantly L2 English speaking bilinguals of Spanish and English exhibit phonological patterns similar to those of monolingual speakers of L1 English, whereas predominantly L1 Spanish speaking bilinguals exhibit phonological patterns similar to those of monolingual speakers of L1 Spanish, even though some cross-linguistic effects were observed for two bilingual groups. With respect to grammar, Paradis, Nicoladis, and Crago (2007) examined the production of the past tense in 4-year-old French-English bilingual children, where the French past tense forms are less complicated than the English ones. They showed that French-dominant children produced French past tense forms as accurately as French monolingual children, but that English-dominant children did so less accurately. Therefore, if language dominance is an important factor in perception and production of L2 phonology and grammar, then language proficiency and dominance need to be determined with greater precision than has been done in the bilingual literature up to this point.

> Hypothesis: Korean-English bilingual (KEB) children, who were exposed to L2 English earlier, will exhibit greater English proficiency than KEB adults who were exposed to L2 English later. Furthermore, KEB children will exhibit great English proficiency than Korean monolingual (KM) children who have just begun to learn English and have never been exposed to L2 English in naturalistic settings. In the same vein, English-dominant KEB children, who have been exposed to L2 English at earlier age and for a longer period, will exhibit greater English proficiency than Korean-dominant KEB children,

who were first exposed to L2 English at an older age and have been exposed for a shorter period of time.

The need for normative data on bilingual children has been highlighted in recent years, so that children from culturally and linguistically diverse backgrounds will not be over-diagnosed as exhibiting a language disorder in L2, or under-diagnosed as merely exhibiting a language difference when in fact a true language disorder is present, either in L1, L2, or both (Iglesias, Miller, & Nockerts, 2009). Moreover, most of the extant bilingualism literature emphasizes the importance of ascertaining language dominance in bilinguals. Therefore, describing the relation of degree of bilingualism to perception and production proficiency—made possible only by precise measures of bilingualism—should be an important contribution to the extant literature.

## Method

### *Participants*

Initially, participants were (a) 9 Korean-English bilingual (KEB) children between the ages of 8 and 13 years, (b) 8 Korean-English bilingual adults aged 18 or older, and (c) 18 Korean monolingual (KM) children between the ages of 8 and 13 years. These three initial groups will be called the "age-language status" comparison groups. All participants had normal hearing sensitivity and no emotional, psychological or neurological problems, as reported by the participants or by participants' parents. Participants were acquaintances or were recruited through advertisements posted in the community and on school notice boards, which provided information on eligibility for participating in the study. Bilingual participants had lived in the U.S. at least 2 years at the time of participation. Monolingual participants lived in South Korea and had never lived in an English-speaking community at the time of participation. These children were learning English as a foreign language (FL) one hour a week at school; therefore, KM participants in the present study had very limited English speaking skills, but were at least able to read the alphabet.

KEB children were identified as belonging to one of three proficiency subgroups, depending on age of arrival (AOA), length of residence (LOR), and qualitative consideration of a language survey and language sample. English-dominant bilingual children (ED) were exposed to English prior to 3 years of age, with an LOR of at least 5 years, spoke English at an advanced or native level, and spoke English more often and at a more advanced level than Korean. Balanced bilingual children (BAL) were exposed to English between the ages of 3 and 7 years, with an LOR of at least 3 years, and spoke both English and Korean at an advanced or native level. Korean-dominant bilingual children (KD) were exposed to English at the age of 7 years or older, with an LOR of at least 2 years, spoke Korean at an advanced or native level, and spoke Korean more often and at a more advanced level than English. In order to make a reliable and accurate quantitative judgment about the L2 English proficiency of participants in the present study, extensive language sample analyses were conducted. All participants but one (a KEB adult) filled out a Family Language Preference Survey (see Appendix D) and were given two standardized articulation and vocabulary tests. One KEB adult and two KM children did not participated in an oral interview in English.

*Task Procedure*

*Language Survey*. According to a number of studies of bilingualism, L1 interference and L2 experience are closely related to perceptual and production performance (i.e., Flege, Bohn, & Jang, 1997; Flege & MacKay, 2004; Gildersleeve-Neumann et al., 2008; Goldstein, Fabino, &

Iglesias, 2003; Keshavarz & Ingram, 2002; Paradis, 2001; Schnitzer & Krasinski, 1994, 1996). Therefore, a questionnaire—the Family Language Preference Survey (see Appendix D)—was given to adult participants or the parents of the child participants. These surveys were examined for age of participants, AOA, LOR, and language(s) used at home and at school, in order to group KEB children into one of the three proficiency subgroups. Nevertheless, parental or self report about L2 proficiency is not always reliable (cf. Goldstein & Washington, 2001); therefore, qualitative consideration of a language sample was deemed necessary. Language samples in English were collected and were extensively analyzed for vocabulary diversity, grammatical errors, and speaking fluency; using computer software for precise language-sample analysis.

*Articulatory and Vocabulary Tests*. All participants were given a speech and language evaluation, including standardized tests of English articulation, namely the Goldman Fristoe Test of Articulation-2 (GFTA-2), and expressive English vocabulary, namely the Expressive Vocabulary Test (EVT). The GFTA-2 is a systematic means of assessing an individual's articulation of the consonant sounds of standard American English. The test measures presence and type of errors in single words and provides normative scores for monolingual English speakers for different age levels from 2 through 21 years. The EVT is a norm-referenced test of expressive vocabulary and word retrieval intended for assessment of English-speaking children and adults at ages 2;6 (years; months) and older. Each test took about 25 to 35 minutes to administer. I should note that the GFTA-2 and the EVT were not intended to diagnose language disability in English, but rather to confirm the classification of the participants as either bilingual or monolingual, and for the KEB children, their language dominance or balance. The EVT was administered only up through to the level of items indicated for the participant's age group. Had higher items been administered, the child participants might have scored higher.

*Language Sample Analyses*. In addition to the survey and standardized testing, an English language sample was collected from an oral interview. A language sample was collected both in English and Korean. (The Korean sample will not be reported here, but instead retained for the future study). To elicit the English language sample, each bilingual participant was asked to talk about his or her favorite movies, books, or TV shows and to tell a story about them. Such samples provide words and sentences a participant says spontaneously (in a natural setting). The sample length was usually 5 or more minutes. Monolingual participants were asked to answer basic questions typical of daily conversation, such as "What grade are you in; how old are you; how are you; what is your hobby;" and so on. In the present study, samples from the KM children lasted approximately 2 minutes. Having an interview in English with the KM children proved very challenging. Many of their responses were laughter, giggles, and the sentence "I don't know."

Language samples were recorded on a laptop computer (Satellite S205, Toshiba) with a headset microphone (Audio 350 ultimate performance headset, Plantronics) in a quiet lab room at the University of Illinois or in a quiet room at the participant's home. The microphone was at a constant distance from the interviewee (approximately 6 inch from the participant's mouth). The English sample was transcribed and analyzed using Systematic Analysis of Language Transcripts (SALT) software, designed to analyze language samples and compare measures from a sample to a variety of control groups.

At least 50 English utterances were collected from KEB children (*M* = 63, *SD* = 11) and KEB adults (*M* = 61, *SD* = 11), and at least 10 English utterances from KM children (*M* = 17, *SD* = 6). Previous studies have shown that 50 utterances are long enough to provide sufficiently powerful predictive evidence of oral language proficiency in monolingual children (Berman & Slobin, 1994; Miller & Chapman, 1991; Snow, 1983). Kemp and Klee (1997) reported that the majority of speech-language pathologists typically collected 50 utterances from a child, 28% collected less than 50 utterances, and 24% collected more than 50 utterances. Therefore, 50 utterances were considered to be enough to serve the purposes of the present study, to examine English proficiency in each group. Further, SALT provides normative database files with which to compare various measures taken from the participants' language samples.

The following measures are reported in a language sample analysis (LSA) for the three initial age-language status groups (i.e., KEB children, KEB adults, and KM children), and for the three proficiency subgroups (i.e., the ED, BAL, and KD groups), using SALT: length of the sample, as the number of complete and intelligible (C & I) utterances; number of different words (NDW) and of total words (NTW) in approximately 50 C & I utterances for KEBs (or NDW and NTW in all C & I utterances for KMs); mean length of utterance (MLU) in words; and speech rate (in words per minute or WPM). The number of ungrammatical utterances, mazes, and abandoned utterances was also reported; however, those measures had to be manually coded into SALT. Given that the length of utterances (MLU) differs from individual to individual, standard deviations from the normative mean—matched to the speaker's utterance length- and age—are also reported in addition to the absolute values of the measures for the participants.

The SALT software automatically calculates average performance for a relevant sample of children in the database and the statistical significance of any differences between the target child and those in the relevant database, based on the sample length (in C & I utterances) produced by children within 6 months of the target child's age. Target matched data files provide the following measures: MLU in words, WPM, the number of mazes, and the number of abandoned utterances. The subject pool for the selected database was composed of over 350 monolingual, English-speaking children between 3 and 13 years of age at preschools and schools in Madison, Wisconsin. All these measures were subjected to two separate Multivariate Analysis of Variances (MANOVA) for the three age-language status groups and the three proficiency subgroups.

*General Procedure*

Following initial contact with potential participants, either through phone calls or e-mails, a first meeting was arranged. After consent had been given and participants were enrolled, parents of a participating child or an adult participant signed a consent form and were given the Family Language Preference Survey. A perceptual task (the Nonsense Syllable Confusion Matrix task or NSCM) was presented across a number of sessions. The articulation test (i.e., the GFTA-2) and vocabulary test (i.e., the EVT) were typically given at the end of the second or third session of the NSCM task. A single-syllable production task (the Nonsense Syllable Imitation Task or NSIT) took about 10 minutes to administer and was given at the end of the third or fourth session of the NSCM task. A phonological awareness, single- to multi-syllable production task (the Nonword Repetition Task or NRT) took about 10 minutes to administer and was given at the end of the fourth or fifth session of the NSCM task. English and Korean language samples were collected at the last session of the NSCM task. Child participants were frequently asked if they felt comfortable with a task and were willing to continue. Participants completed all of the tests, tasks, and procedures in 5 to 7 experimental sessions on average. This general procedure is summarized in Figure 3.

# Results

## *Language Survey and Articulation and Vocabulary Tests*

### *Comparisons of the Three Age-Language Status Groups*

Table 5 is a summary of the survey results and the two tests for the KEB children, the KEB adults, and the KM children. The average chronological age of the KEB children was 10;3 (years; month; *SD* = 2;1), the average age of the KEB adults was 22;8 (*SD* = 2;1), and the average age of the KM children was 11;0 (*SD* = 0;6). Age of arrival (AOA) was 5;8 (*SD* = 3;7) on average for the 9 KEB children and 18;0 (*SD* = 2;6) for the 7 KEB adults. None of the KM children had lived in an English-speaking community. The length of residence (LOR) of the KEB children was 4;7 (*SD* = 1;7) which is similar to that of KEB adults, 4;9 (*SD* = 0;7). As for perceived language proficiency, some KEB children were reported by their parents to speak English better than Korean but others, to speak Korean better than English. All KEB adults reported themselves to be less proficient in English than in Korean, and all KM children were reported to speak limited English. KEB children were reported to speak English as often as Korean at home but more often at school. KEB adults, most of who lived by themselves in the U.S., answered that they speak only Korean when they are with their family members (over the phone) or their friends, even though they speak English more often than Korean at school. All of the KM children spoke Korean only both at home and at school, although they received an hour-long class in the English language once a week at school.

Raw error scores on the GFTA-2, i.e., the number of incorrect articulations, are presented, instead of standard scores. Raw scores are more informative and better fit the purposes of the present study, because even a small number of articulation errors results in a dramatic drop in the standard score. Performance on the GFTA-2 was significantly different among the three groups, $F_{2, 33} = 20.228$, $p = .000$: the average number of incorrect articulations was smaller in the KEB child group ($M = 3.1$, $SD = 2.2$) than in the KEB adult group ($M = 6.3$, $SD = 1.6$), $p = .000$, and the KM child group ($M = 13.9$, $SD = 5.5$), $p = .000$; and smaller in the KEB adult group than in the KM child group, $p = .000$. The most frequent errors were for production of the consonant /ʤ/ and the coda consonant /z/ in KEB children, the consonants /θ, ð, ʤ/ and the coda consonant /z/ in KEB adults, and the consonants /v, θ, ð, z, ʧ, ʤ/ in KM children.

The values reported for the EVT in Table 5 are standard scores, which have a mean of 100 and a standard deviation of 15. Additionally, performance on the EVT was significantly different among the three groups, $F_{2, 33} = 38.728$, $p = .000$. The average standard score of the KEB children ($M = 80$, $SD = 10$) was lower than that of the KEB adults ($M = 96$, $SD = 12$), $p = .008$, but higher than that of the KM children ($M = 60$, $SD = 7$), $p = .000$. Performance on the EVT was higher for the KEB adults than for the KM children, $p = .000$. Based on the observation of the survey and examination of the GFTA-2 and EVT, it appears that the KEB children—who were exposed to English earlier than the KEB adults but for a similar length of time (LOR was approximately 4 ½ to 5 years)—produced English sounds more accurately, but had a smaller vocabulary size than the KEB adults. (Recall, though, that because KEB children were not administered vocabulary items above their age range, their scores may under-represent their true vocabulary abilities.) The KEB adults—who were exposed to English later but for a longer period than the KM children—produced English sounds more accurately and also had a greater vocabulary size in English than the KM children.

*Comparisons of the Three Proficiency Subgroups*

Table 6 is a summary of data from the Family Language Preference Survey for the three subgroups of KEB children, i.e., the ED, BAL, and KD groups. The average chronological age of the ED children was 8;4 (*SD* = 0;5), the average age of the BAL children was 9;8 (*SD* = 1;6); the average age of the KM children was 12;7 (*SD* = 0;7). AOA was earlier in the ED and BAL groups (*M* = 3;0, *SD* = 1;0 and *M* = 4;0, *SD* = 2;0, respectively) than the KD group (*M* = 10;0, *SD* = 1;0). LOR was also longer in the ED and BAL groups (*M* = 5;6, *SD* = 0;3 and *M* = 5;9, *SD* = 0;2) than the KD group (*M* = 2;7, *SD* = 0;10). As perceived by their parents, children in the ED group were reported to speak English fluently at an advanced or native level, but Korean at an intermediate or advanced level. They also spoke English more frequently than Korean at home. Children in the BAL group were reported to speak both languages fluently at an advanced or native level, and to speak English as often as Korean at home. Children in the KD group were reported to speak Korean at a more advanced level than English, and to speak Korean more frequently than English at home. In general, children in the ED and BAL groups were younger, learned English at earlier age, had lived in the U.S. longer, and spoke English at a more advanced level than the KD group.

Average raw error scores on the GFTA-2 seemed to be slightly higher in the ED and BAL groups (*M* = 2, *SD* = 0 and *M* = 2.3, *SD* = 1.5) than in the KM group (*M* = 5, *SD* = 3). A significance test, however, showed that performance differences among the three proficiency subgroups were not significant, $F_{2,8}$ = 2.147, *p* = .198. Average standard scores on the EVT, however, were higher for the ED and BAL groups (*M* = 86, *SD* = 7 and *M* = 86, *SD* = 8) than for the KD group (*M* = 69, *SD* = 3). Performance on the EVT was significantly different among the three proficiency subgroups, $F_{2,8}$ = 7.268, *p* = .025. According to a Scheffe's posthoc test, the ED group performed similarly on the EVT to the BAL group, and both the ED and BAL groups performed better than the KD group, *p* = 0.43 and *p* = .046, respectively.

*Language Sample Analyses*

*Comparisons of the Three Age-Language Status Groups*

Various measures from the language sample analysis (LAS) in the KEB child, KEB adult, and KM child groups are presented in Table 7 and the related statistical analysis (i.e., MANOVA) is presented in Table 8. The average number of C & I utterances was 63 for the KEB children (*SD* = 11, *R, or range,* = 46 – 75), 61 for the KEB adults (*SD* = 11, *R* = 48 – 74), and only 17 for the KM children (*SD* = 6, *R* = 5 – 25). The MANOVA showed that the average number of C & I utterances was significantly different among the three groups, $F_{2,31}$ = 110.971, *p* = .000, with a large effect size $\eta_p^2$ = .888, *P* = 1.000. A Scheffe's posthoc test showed that the sample length (in C & I utterances) of the KM children significantly shorter than that of the KEB children, *p* = .000, and the KEB adults, *p* = .000. The oral interview in English was typically short with the KM children, because they barely spoke English and also appeared to feel uncomfortable with having a conversation in English.

According to the results of the MANOVA, NDWs and NTWs were significantly different among the groups, $F_{2,31}$ = 143.604, *p* = .000, $\eta_p^2$ = .911, *P* = 1.000, and $F_{2,31}$ = 104.445, *p* = .000,

$\eta_p^2$ = .882, *P* = 1.000, respectively. Scheffe's posthoc tests showed no group differences between the KEB children and the KEB adults for these measures, but significant group difference between the KEB children and the KM children, and between the KEB adults and the KM children. Comparison of NDW and NTW for the KEB children (*M* = 117, *SD* = 12, and *M* = 274, *SD* = 55, respectively) with those for the KM children (*M* = 17, *SD* = 8, and *M* = 30, *SD* = 13, respectively) revealed great differences in English proficiency between the two groups. If KEB children produced 17 C & I utterances, their NDW and NTW are expected to be 32 and 74 on average for English monolingual children (according to the SALT database), which are still greater than NDW and NTW of KM children.

In the LSA, NDW and NTW frequently have been used to measure linguistic development and vocabulary skills of a child and to diagnose the presence of a language impairment (LI; DeThorne, Petril, Schatschneider, & Cutting, 2010; Gavin & Giles, 1996; Miller & Chapman, 1991; Paul, 2007). Paul (2007) argued that a child with LI would produce a smaller NDW and NTW than an age-matched child without LI. DeThorne et al. (2010) showed close relations of NDW and NTW with performance on word-attack measures of reading. Therefore, it can be concluded that the KEB children and the KEB adults in the present study had a greater vocabulary size than the KM children.

In the present study, the number of ungrammatical utterances was manually tallied and reported by counting incorrect word-order, missing elements, unnecessary elements, or incorrect word-inflections. The MANOVA showed a significant group difference for the number of ungrammatical utterances, $F_{2, 31}$ = 18.222, *p* = .000, $\eta_p^2$ = .713, *P* = 1.000. A follow-up Scheffe test showed that the KEB children and adults (*M* = 5.8, *SD* = 3.8 and *M* = 8.8, *SD* = 4.2, respectively) produced more grammatical errors than the KM children (*M* = 0.6, *SD* = 0.7), $p = .002$ and $p = .000$, respectively. The smaller number of utterances in the language samples of the KM children may have resulted in the smaller number of grammatical errors for the KM children.

The average MLU in words was similar between the KEB children (*M* = 5.3, *SD* = 1.5) and the KEB adults (*M* = 6.4, *SD* = 1.3), but it was lowest for the KM children (*M* = 1.5, *SD* = 0.3). The MANOVA showed a significant group difference for MLU in words among the three groups, $F_{2, 31} = 69.187$, $p = .000$, $\eta_p^2 = .832$, $P = 1.000$. A Scheffe test showed that the KEB children and adults had longer MLUs than the KM children, $p = .000$ and $p = .000$, respectively. Standard deviations from the utterance- and age-matched normative means for the English monolingual (EM) children in the SALT database showed that the averaged MLUs in the KEB children and adults were not substantially different from those in the monolingual English profile data (*SD* = 0.02 and *SD* = 0.72, respectively). The average MLU in the KM children, however, was substantially lower than the MLU in the matched profile data (*SD* = -2.34). The MANOVA showed that SD differences were also significant among the three groups, $F_{2, 31} = 37.628$, $p = .000$, $\eta_p^2 = .729$, $P = 1.000$. Its posthoc test showed that the KEB children and adults performed significantly better than the KM children, $p = .000$ and $p = .000$, respectively.

WPM was calculated by dividing the total number of completed words by the elapsed time. The average WPM was similar for the KEB children (*M* = 75, *SD* = 27) and adults (*M* = 72, *SD* = 23), but lowest for the KM children (*M* = 18, *SD* =4). Differences in WPM among the three groups were statistically significant, $F_{2, 31} = 38.840$, $p = .000$, $\eta_p^2 = .902$, $P = 1.000$. According to a Scheffe, the KEB children and adults produced more English words per minute than the KM children did, $p$ = .000 and $p$ = .000, respectively. The average SDs from the normative means suggest that the speech rate of the KEB children ($SD$ = - 0.74) is slightly but not substantially slower than for the EM children in the length- and age-matched profile group in SALT. On the other hand, rates for the KEB adults and KM children ($SD$ = -1.25 and $SD$ = -2.16, respectively) were substantially slower than those for the age-matched profile group. The average SD from the normative mean was also significantly different among the group, $F_{2, 31}$ = 16.174, $p$ = .000, $\eta_p^2$ = .536, $P$ = .999, and a Scheffe showed that the KEB children and adults produced significantly faster than the KM children, $p$ = .000 and $p$ = .014, respectively.

Mazes are portions of utterances that do not contribute to meaning, such as fillers (e.g., "um...") within an utterance (Loban, 1976). For example, a maze is marked in the following sentence with parentheses: "(I um uh I want I want to go) I want to drive to the store." Previous studies have argued that the frequency of mazes provides a reliable diagnosis for a child with LI and also a child from a culturally and linguistically diverse background (Fiesta et al., 2005; Leadholem & Miller, 1992). The MANOVA showed a significant group difference for the number of mazes among the three groups, $F_{2, 31}$ = 12.958, $p$ = .000, $\eta_p^2$ = .481, $P$ = .994. Contrary to what might be expected, the KEB children ($M$ = 27, $SD$ = 22) and adults ($M$ = 18, $SD$ = 6) produced more mazes than the KM children ($M$ = 3, $SD$ = 2), $p$ = .000 and $p$ = .042, respectively.

Some studies have argued that an increase in mazes does not necessarily reflect disfluencies (Starkweather, 1987), rather the number of mazes can increase when delivering complex and abstract thoughts or when speakers are attempting to correct themselves (Leadholem & Miller, 1992; Levelt, 1989). Therefore, the small number of utterances and simple content of thoughts delivered by the KM children may have resulted in the smaller number of mazes in their utterances. SD from the normative mean for the number of mazes were substantially higher in the KEB child group (*SD* = 1.09) than the profile group (i.e., EM children in SALT), but similar in the KEB adult and KM child groups (*SD* = 0.24 and *SD* = -0.36, respectively). A group difference was not significant for the average SD of the number of mazes, $F_{2, 31} = 2.114$, $p = .140$, $\eta_p^2 = .131$, $P = .397$.

Utterances that dropped off before completion of a thought were coded as abandoned utterances. Abandoned utterances were distinguished from mazes. If the discontinued utterance was not followed by a corrected utterance or was never readdressed, then it was classified as an abandoned utterance. For example, an abandoned utterance is marked in the following sentence with parentheses, “(I went fishing with . . .) I hate to go out.” The MANOVA showed that the number of abandoned utterances was significantly different among the groups, $F_{2, 31} = 17.258$, $p = .000$, $\eta_p^2 = .552$, $P = .999$. According to a Scheffe test, the KEB children (*M* = 7.1, *SD* = 4.6) and adults (*M* = 8.8, *SD* = 6.5) produced abandoned utterances more often than the KM children (*M* = 0.1, *SD* = 0.3). Compared to EM children in the matched profile group, the KEB children (*SD* = 2.14) and adults (*SD* = 1.72) produced abandoned utterances substantially more often than the profile group. The average SD of the number of abandoned utterances were also significantly different among the group according to the MANOVA, $F_{2, 31} = 12.488$, $p = .000$, $\eta_p^2 = .471$, $P = .992$. Its follow-up Scheffe test showed that the KEB children and adults produced abandoned utterances significantly more often than the KM children, $p = .000$ and $p = .006$, respectively.

*Comparisons of the Three Proficiency Subgroups*

Descriptive and inferential statistics of the LSA for the three subgroups for English proficiency are presented in Tables 9 and 10. Overall, the BAL group seemed to perform better than the other two groups; however, no significant differences were found for any measures among the three subgroups. WPM was not significantly different among the three subgroups; however, WPM of the KD group (*SD* = - 1.2 from the normative means) was substantially lower than that of the profile group, unlike the ED and BAL groups (*SD* = -0.8 and *SD* = -0.2, respectively). The number of mazes for the KD children was also substantially lower than for the EM children of the profile group (*SD* = 2.53). The numbers of abandoned utterances for all three subgroups (i.e., the ED, BAL, and KD groups) were substantially greater than for the profile group (*SD* = 1.3, *SD* = 3.0, and *SD* = 2.1, respectively).

## Summary and Conclusions

The research question asked if younger or more experienced L2 learners would exhibit better English proficiency than older or less experienced L2 learners, as measured by two standardized tests of English (i.e., GFTA-2 and EVT) and by the language sample analysis (LSA). In the age-language status comparison, it was hypothesized that the KEB children would show better proficiency than the KEB adults and the KM children. The KEB children performed better on the GFTA-2 than the KEB adults, but the KEB children performed worse on the EVT than the KEB adults. The KM children scored lowest among the three groups on both the GFTA-2 and the EVT tests. In the LSA using SALT, KEB children and adults performed significantly better on every measure than KM children. KEBs produced more C & I utterances, had a greater NDW and NTW, had longer MLUs, and had more fluent speech rates (i.e., WPM) than the KM children. Even though KEBs produced more grammatical errors, mazes, and abandoned utterances than the KM children, these do not necessarily reflect lower English proficiency, given that the KM children made only few utterances and those utterances were basically answers to questions rather than storytelling. No measure was significantly different between the KEB children and the KEB adults. Therefore, my hypothesis was only partially consistent with the results of the present study, in that the KEB children showed greater proficiency than the KM children but not the KEB adults.

In the English proficiency comparison of the three subgroups (i.e., the ED, BAL, and KD subgroups), GFTA-2 scores were not significantly different; however, EVT scores were, where the ED and BAL groups scored higher than the KD group. No significant subgroup differences were observed for the LSA measures in the MANOVA. Therefore, the results of the LSA do not confirm the hypothesis that the ED group would show greater proficiency than the BAL and KD groups. Nonetheless, speech rate was significantly slower in the KD group, mazes were more frequent in the KD group, and abandoned utterances were more frequent in all three subgroups (i.e., ED, BAL, and the KD) compared to the EM children in the matched profile data. Considering standard deviations from the normative means, the KD group seemed to have good enough English proficiency to be considered fluent English speakers. For the ED and BAL groups, the LSA showed that these children were not substantially different from the EM children in the matched profile data for many of the measures.

# CHAPTER III

# NONSENSE SYLLABLE CONFUSION MATRIX TASK

## Rationale for a Perceptual Task

Bilingualism is of increasing interest in the field of speech and hearing science because speech-language pathologists (SLPs) report an increasing number of children from culturally and linguistically diverse backgrounds on their caseload (Huang, Hopkins, & Nippold,1997). In order to provide appropriate services to these clients, documented guidelines, such as norm-referenced diagnostic tests, for bilingual individuals should be available to SLPs, so that SLPs can be properly trained in understanding the nature of bilingual language development and disorders. Several studies have argued that bilingual children referred for special education or speech-language therapy are either over- or under-represented due to the lack of knowledge and expertise of referring SLPs, as well as the lack of resources to diagnose bilingual children (Lindsay, Pather, & Strand, 2006; Mennen, Standsfield, & Johnston, 2005). Results of the present study may help us understand the origins of problems in perceiving and producing English as an L2 or FL and contribute to guidelines for the development of linguistic proficiency in English, and determine which English sounds should be the focus of L2 education or clinical treatment.

Early studies of speech perception showed that speech perception is categorical, where listeners are more likely to notice acoustic differences *between* categories than *within* categories (Abramson & Lisker, 1970, 1973). When listeners map the incoming auditory stimulus (i.e., acoustic signal) to phonological forms of words (i.e., phonemes), this abstract representation acts like a magnet that draws acoustic signals near its boundary. For example, a native English listener identifies any speech sound at + 30 ms on a VOT continuum between /p/ and /b/ as either /p/ or /b/, but not as a new phoneme. Association of the physical speech signal with an abstract representation seems to get stronger with an increase in L1 experience (Eimas, 1978; Werker, Gilbert, Humphrey, & Tees, 1981; Werker & Tees, 2002) and possibly also with an increase of L2 experience for L2 sounds. To explore development of speech perception from infancy to adulthood, studies comparing perception of several languages have attracted increasing attention in recent years.

The cross-linguistic perception literature suggests that sensitivity to acoustic differences in nonnative contrasts diminishes with increases in L1 experience, particularly for fragile nonnative contrasts—contrasts that are rare across languages and ones where the two elements are acoustically close to each other (Burnham, 1986). Lack of phonemic categories relevant to L1 may cause infants to focus on acoustic differences, which, in turn, helps them to perform discrimination of nonnative contrasts better than adults (Cohen, Amsel, Redford, & Casasola, 1998; Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Werker, 1995). The present study does not aim to investigate developmental change in speech perception or establishment of perceptual categories. Nevertheless, results of the present study may contribute to our understanding of how perceptual categories develop at the initial stage of L2 learning, by comparing perceptual performance in monolinguals to experienced bilinguals.

As argued in Chapter II, studies of bilingualism should take L1 interference and L2 experience into consideration. Degree of L1 interference seems to relate to age of L2 acquisition and degree of L2 experience. Mayberry and Eichen (1991) showed that age of acquisition of ASL affects fluency of sign language. This holds true for oral languages, too (Johnson &

Newport, 1987). Baker et al. (2008) also suggest that Korean- and English-speaking bilingual children seem to experience less L1 interference than bilingual adults, because they are more sensitive to subtle, phonetic differences between Korean and English than are adults. As mentioned previously, in general, simultaneous bilingual children who learn L2 before age 3 and successive bilingual children who learn L2 before age 7 eventually achieve native-like proficiency (de Houwer, 1995; McLaughlin, 1978; Oller et al., 1997; Peña et al., 2003). Flege and MacKay (2004) argue that LOR (as a proxy for amount of L2 experience) predicts one's L2 proficiency better than AOA alone (as a proxy for age of L2 acquisition). In support of this, several recent studies have shown some trends for a relation between L2 experience and production accuracy (Gildersleeve-Neumann et al., 2008; Goldstein et al., 2005; Goldstein & Washington; 2001). Perhaps the same would be true of perception. To tease apart L2 experience and age of L2 acquisition, the present study includes a monolingual group, as well as two bilingual groups with varying degrees of exposure to English.

Inclusion of a monolingual group allows us to see language-independent (at acoustic level) and language-dependent (at phonological level) effects on perception of consonants and vowels as a FL, as compared with bilinguals, second language learners. The present study asks whether KM children, who represent the least amount of L2 experience, perceive like the bilingual adult group or the bilingual child group. If the KM group were to perform like the KEB adult group, then the amount of L2 experience could be interpreted as a stronger influence than the age of acquisition. Conversely, if the KM group instead performs like the KEB child group, then age of acquisition could be interpreted as a stronger influence than the amount of L2 experience.

Previous studies predict that perceptual performance will be positively impacted by increasing amount of L2 experience but negatively by increasing age of L2 acquisition. Nonetheless, some L2 sounds may be challenging for both experienced and inexperienced L2 learners of any age to perceive, due to difficulties that stem from the acoustic nature of those sounds. Based on observations in the literature about speech perception in native English speakers, voiced consonants, fricatives, and anterior sounds (made near the front of the mouth) are more likely to be identified incorrectly than voiceless consonants (Ali et al., 2000), stops (Phatak & Allen, 2007), and posterior sounds (Hume et al., 1998; Nishi, Lewis, Hoover, Choi & Stelmachowicz, 2010), respectively. Therefore, difficulties that L2 learners might have in perceiving some sounds might stem either from universal, intrinsic acoustic demands or from the influence of L1 Korean phonology, or both. The present study will also examine perceptual performance on each individual sound so to see which English sounds are identified poorly across the groups, possibly due to the intrinsic acoustic nature of those sounds.

## Research Questions and Hypotheses

The present study investigates whether consonants and vowels (sound type), and familiar and unfamiliar sounds (familiarity), are perceived differently in Korean-English bilinguals and Korean monolinguals. SLM predicts that unfamiliar sounds should be performed better than familiar ones, whereas PAM predicts familiar sounds should be performed better than unfamiliar ones. Nonetheless, both studies assume that perception of nonnative *consonants* works in the same way as perception of nonnative *vowels*. Some speech perception studies, however, argue that the auditory mechanism for perceiving consonants differ from that for vowels (Allen, 1996, Caramazza, Chialant, Capasso, & Miceli, 2000; Fletcher, 1929;Owren & Cardillo, 2006). Owren and Cardillo (2006) examined how listeners discriminate talker and word meaning, given consonant-only (the vowel part in the speech stream was replaced with silence) or vowel-only stimuli. And they showed that consonants are more crucial than vowels when delivering linguistic meanings. Caramazza et al. (2000) showed that perception of consonants can be diminished independently of the perception of vowels and argued that consonants and vowels are autonomous and processed at different levels of perception. The present study examined whether perceptual performance differs between consonants and vowels from an L2 or FL; and if so, whether sound type (i.e., consonants or vowels) interacts with the familiarity of sounds in the perception of speech by bilinguals and monolinguals. In addition, perceptual performance of each individual phoneme will be explored in order to find some sounds, either consonants or vowels, that are perceived well (i.e., high-scoring sounds; Phatak & Allen, 2007), but others that are perceived poorly (i.e., low-scoring sounds), regardless of sound type, familiarity of sounds, or participant group. Perceptual performance is also compared across the participant groups in order to examine the influence of L2 experience on perception of English nonsense syllables, where the group comparisons include age group (i.e., KEB children vs. adults), proficiency subgroup (i.e., the ED, BAL, and KD groups), and language status group (i.e., KEB vs. KM children).

*Research Question 1*: *Familiar vs. Unfamiliar Sounds*

The first research question asks if familiar sounds (L2 English sounds that are similar to L1 Korean phonology) are perceived differently from unfamiliar sounds (L2 English sounds that are new to L1 Korean phonology) as addressed by SLM and PAM.

1. Do Korean-English bilinguals and Korean monolinguals perceive familiar English sounds differently from unfamiliar ones?

This question asks which types of English sounds are more challenging for KEBs and KMs to perceive. According to SLM (Flege, 1987; 1992; 2002), the phonetic and/or phonemic difference between an L2 sound and the closest L1 sound helps L2 learners to produce and also perceive such new L2 sounds with native-like fluency. On the other hand, similarity between an L2 sound and the closest L1 sound hinders L2 learners from producing and perceiving such similar L2 sounds. The PAM (Best, McRoberts, & Sithole, 1998; Best, McRoberts, & Goodell, 2001), however, yields a different prediction. According to PAM, a nonnative contrast that is similar to the one in L1 helps L2 learners to discriminate a pair of L2 sounds. On the other hand, a nonnative contrast that is new to L1 phonology is thought to cause a perceptual assimilation of the two sounds in the contrast into a single category, thus it hinders L2 learners from perceiving (and also possibly producing) such new L2 sounds. SLM predicts unfamiliar sounds will be perceived correctly more often than familiar ones, at least by highly experienced listeners; whereas, PAM predicts familiar sounds will be perceived correctly more often regardless of one's L2 experience.

Hypothesis 1a: Korean-English bilinguals will perceive unfamiliar sounds better than familiar ones but performance will be the opposite in Korean monolingual children, according to SLM.

Alternate Hypothesis 1b: Korean-English bilinguals and Korean monolinguals will perceive familiar sounds better than unfamiliar ones, according to PAM.

Predictions from the two theories will be compared in detail in Chapter III. (Please refer also to Appendix C for a comparison of SLM to PAM).

*Research Question 2: Consonants vs. Vowels*

The second research question asks if consonants are perceived differently from vowels by KEBs and KMs.

2. Do Korean-English bilinguals and Korean monolinguals perceive English consonants differently from vowels?

Neither SLM nor PAM makes predictions for perception of consonants separate from that of vowels. A few studies have suggested that consonants are processed neuroperceptually in a different manner from vowels (Boatman, Hall, Goldstein, Lesser, & Gordon, 1997; Caramazza et al., 2000; Romani, Granna, & Semenz, 1996); however, different processing of consonants and vowels has been seldom addressed in the extant literature on L2 learning. Romani et al. (1996) showed that monolingual speakers with aphasia made more production errors on vowels than on consonants, and Owren and Cardillo (2006) argued that intelligibility of word stimuli was higher for consonant-only words than for vowel-only words. Cole, Yan, Mak, Fanty, & Bailey (1996), however, showed the opposite, where the presence of vowels helped English-speaking listeners recognize words twice as often as the presence of consonants.

The present study investigates whether perception of English consonants differs from the perception of English vowels in Korean-English bilinguals and Korean monolinguals, as evaluated by a Nonsense Syllable Confusion Matrix (NSCM) task. In recent years, Johnson and her colleagues (2007) examined confusion matrices for English nonsense syllables and observed that some English-speaking children with reading disabilities perceived consonants better than vowels. I hypothesize that the perception of consonants is better than the perception of vowels, as found by Johnson et al. Of the speech perception studies just reviewed, the experimental design in that study is most similar to the NSCM task used in the present study.

Hypothesis 2: Korean-English bilinguals and Korean monolinguals will perceive consonants better than vowels.

*Research Question 3: Group Comparisons*

Perceptual performance was compared between groups with regard to age (i.e., KEB children vs. adults), proficiency (i.e., the ED, BAL, and KD groups), and language status (i.e., KEB vs. KM children).

3. Does perceptual performance differ due to age (i.e., bilingual children vs. bilingual adults), language status (i.e., bilingual children vs. monolingual children), and proficiency (i.e., comparisons of English-dominant, balanced, and Korean-dominant bilingual children)?

Studies of speech perception in children, reviewed previously, showed that children appear to have different perceptual strategies on certain tasks from adults (with respect to difference limens: Elliott *et al*., 1986; Sussman, 1993; and a shift with age in the weighting of formant transitions: Morrongiello *et al*., 1984; Nittrouer, 1992). It was generally shown that young infants discriminate nonnative contrasts correctly more often than adults (Best & McRobers, 2003; Eimans *et al*., 1971). Nevertheless, there is a possibility that KEB adults would be better at perception, if children do have different ways of perceiving sounds than adults on certain tasks, as suggested by studies of speech perception in children; and if such differences negatively influence performance. These studies taken together raise the question of whether KEB children will show different patterns of perceptual sound confusion from adults, not only for L1 sounds but for L2 sounds as well.

In addition, the literature in L2 acquisition suggests that the proficiency of L2 is closely related to the age of L2 acquisition and the amount of exposure to an L2 (i.e., L2 experience). When L2 learners are experienced and develop L2 phonological system, they are better at noticing subtle differences between L1 and L2 sounds (Baker et al., 2008; Flege & MacKay, 2004). Therefore, the KEB children with more L2 experience (e.g., earlier AOA, longer LOR, and so on) are expected to perform better in perception and production than children with less L2 experience and adults who are exposed to an L2 later in life. In the present study, KMs who are just beginning to learn English as a FL are the least experienced with L2; therefore, overall perception of L2 English sounds is expected to be the lowest for them.

Hypothesis 3: Overall, bilingual children will perceive L2 English sounds better than bilingual adults and Korean monolingual children; bilingual children will perceive L2 English sounds better than Korean monolingual children; and English-dominant and balanced bilingual children will perceive L2 English sounds better than Korean-dominant bilingual children.

*Research Question 4: Sound Confusion*

The present study examined which consonants and vowels are well or poorly identified, regardless of a listener's age, language status, or proficiency.

4. Are some English sounds identified at a high or low rate, regardless of a listener's L2 background, as examined by perceptual confusion of English consonants and vowels in Korean-English bilingual children and adults and Korean monolingual children?

Previous studies showed that some sounds are more confusing than others: voiced more than voiceless consonants (Abdelatty Ali et al., 2000), fricatives more than stops (Phatak & Allen, 2007), and anterior more than posterior sounds (Hume et al., 1998; Nish et al., 2010). Findings of these studies lead to a possibility that some English sounds are poorly identified due to their intrinsic perceptual difficulty, regardless of an individual's L1.

> Hypothesis 4: Some English sounds are almost always perceived correctly (i.e., they are acoustically "salient") and others are almost always confused with other sounds (i.e., they have weak acoustic saliency), regardless of one's language background.

The purpose of the present study is not to propose which sounds are acoustically salient so that all listeners can identify them well (i.e., not to identify sounds that are "universally" well perceived), but to propose the idea that perceptual performance in bilinguals should take universality into account. "Saliency" might also be operationally defined as the frequency of the sound in English (cf. Dewey (1923) and Fletcher,(1925). Here I adopt the term "saliency" merely as a convenient, descriptive synonym for a high rate of correct perception (or production) across all participant groups. For example, in the present study a salient sound refers to a high-scoring sound in identification on the NSCM task, regardless of one's L2 experience or familiarity. To determine universally salient sounds, many studies of listeners from various language backgrounds would be needed. To my knowledge, very few such studies exist. However, in the long run, results of the present study may contribute to literature that identifies which English sounds are acoustically salient to speakers of many of the world's languages. Presumably these high scoring sounds would require little effort for L2 learners to perceive.

# Method

*Participants*

Nine KEB children, 8 KEB adults, and 18 KM children participated in five to seven sessions of the NSCM task. As mentioned in Chapter II, the KM children had some, though minimal, exposure to English in a foreign language class taken once a week. One KEB child was excluded from the analysis of the NSCM task because this child was the very first participant and performed the first three sessions of the NSCM task with a different method. Later, he performed two sessions of the NSCM task with the method of the present study; however, his data was excluded for the sake of reliability. One KEB adult was excluded from this analysis of the NSCM task and the NRT as well, because he withdrew from the experiment and did not complete the whole task. Data collected from two KM children were excluded from the analyses of the NSCM task because they performed at a much lower level than the other participants, at only 25% correct perception on average (i.e., these two KM children were outliers). Therefore, task results from the remaining 8 KEB children, 7 KEB adults, and 16 KM children were subjected to descriptive and inferential statistics.

*Stimuli*

To lessen the memory load required for the speech perception task, limited sets of consonants and vowels were selected as stimuli: namely, those where the most confusion is expected. Only four consonants /p, k, s, ʧ/ of thirteen possible familiar consonants /p, t, k, s, h, ʧ, m, n, ŋ, r, l, w, j/ and six consonants /b, g, f, θ, z, ʤ/ of nine possible unfamiliar consonants /b, d, g, f, v, θ, ð, ʃ, z, ʤ/ were selected to represent familiar and unfamiliar consonants. As for vowels, one /ɛ/ vowel of five possible familiar vowels /i, ɛ, a, o, u/ and two vowels /ɪ, æ/ of five possible unfamiliar vowels /ɪ, e, æ, ɔ, ʊ/ were selected to represent familiar and unfamiliar vowels. The 13 English sounds were chosen based on similarity to L1 sounds phonemically or phonetically, or both. Familiar sounds /p, k, s, ʧ, ɛ/ in English are phonemically similar to the /p, k, s, ʧ, ɛ/ in Korean, even though the Korean /ʧ/ is more aspirate and front than its English counterpart. Among unfamiliar sounds, English /b, g/ are phonetically similar (or familiar) to Korean [b, g]; however, the latter occur only as allophones of /p, k/ in intervocalic context and thus are categorized as unfamiliar ones (for detail, see the "Familiar versus Unfamiliar English Sounds" of literature review in Chapter I of this document). The 10 consonants and the 3 vowels chosen for the task stimuli are shown in Table 11.

All consonants were combined with all vowels in a CV template, i.e., a syllable with an onset consonant and a following vowel. This yielded 30 CV stimuli: /pɪ, pɛ, pæ/, /bɪ, bɛ, bæ/, /kɪ, kɛ, kæ/, /gɪ, gɛ, gæ/, /fɪ, fɛ, fæ/, /θɪ, θɛ, θæ/, /sɪ, sɛ, sæ/, /zɪ, zɛ, zæ/, /ʧɪ, ʧɛ, ʧæ/, and /ʤɪ, ʤɛ, ʤæ/. Stimuli for the NSCM task were drawn from the Linguistic Data Consortium at the University of Pennsylvania (http://www.ldc.upenn.edu), which are professionally recorded by 20 native speakers of English and English bilinguals without foreign accents. Therefore, there were 20 different sound clips available for each stimulus sound. In the pilot study, any sound clip that is perceived as deviant or different from the label for the clip, as judged by a panel of listeners, was excluded. For example, if a syllable labeled /ʧɛ/ is perceived as either different (e.g., /ʧɪ/) or deviant (e.g., /ʦɛ/) from the label by more than one listener on the panel, the sound clip was not used as a stimulus. The panel consisted of three native speakers of English and the author of the present study, who speaks both Korean and English. All members of the panel had previous training in reading phonetic transcription and listening to phonetic aspects of sounds.

*Task Procedure*

Each syllable was played over headphones (Audio 350 ultimate performance headset, Plantronics) connected to a laptop computer (Satellite S205, Toshiba or ThinkPad A21, IBM), with the volume adjusted by the participant to a comfortable listening level. For each stimulus presented, participants were asked to click on one of 30 syllables displayed on the computer screen, to identify what they think they heard. Response options appeared on the screen in English print (e.g., pi, pe, pa) with a word tag under the button as a cue, as shown in Figure 4. A participant was also asked to indicate whether he or she is confident about his or her response by clicking the symbol "?" for uncertainty and the symbol "!" for certainty. The participant was allowed to repeat a sound clip as many as 5 times.

Each experimental session were run in a quiet experimental room for about an hour. Each session was composed of three or four 10-minute listening blocks with approximately 100 listening trials per block and three intervening 5-minute play breaks. A short practice session was given at the first session or when it is necessary. A total of about 1500 trials for the 30 syllables (30 CV syllables x 3 or 4 presentations of each talker x 17 to 20 different talkers) were collected from each participant. This gave approximately 150 trials for each of the 10 consonants and 500 trials for each of the 3 vowels for each participant. To complete this task, five to seven hour-long sessions were required for each participant.

## Results

Percent correct was calculated for each phoneme, and percent correct was entered as the dependant variable into descriptive and inferential statistical analyses. Percent correct on the NSCM task was normalized via an arcsine transformation prior to testing for statistical significance. In order to determine whether familiar and unfamiliar consonants and vowels are more challenging for participants to identify, familiarity (i.e., familiar vs. unfamiliar) and sound type (i.e., consonants vs. vowels) were chosen as within-group factors in several repeated measures ANOVAs. Between-group factors were (a) age (i.e., the KEB children vs. adults) and (b) language status (i.e., the KEB vs. KM children), and proficiency (i.e., the ED, BAL, and KD subgroups). Effect size for statistical significance was estimated using partial eta squared, $\eta_p^2$. Several planned comparisons were conducted to follow up any significant interactions of familiarity x sound type (i.e., familiar and unfamiliar consonants and vowels). Paired *t*-tests were planned because they examine within-group interactions and there were more than two groups.

### *Familiar vs. Unfamiliar Consonants and Vowels*

#### *Comparisons of the Three Age-Language Status Groups*

Table 12 is a display of descriptive statistics for perception on the NSCM task by the KEB children—including three proficiency subgroups—the KEB adults, and the KM children. Overall, the 10 consonants (i.e., /p, b, k, g, f, θ, s, z, ʧ, ʤ/) were identified better (*M* = 74% correct, *SD* = 12) than the 3 vowels (i.e., /ɪ, ɛ, æ/; *M* = 60, *SD* = 12). The 5 familiar sounds (i.e., /p, k, s, ʧ, ɛ/) were identified better (*M* = 73, *SD* = 16) than the 8 unfamiliar sounds (i.e., /b, g, f, θ, z, ʤ, ɪ, æ/; *M* = 63, *SD* = 12). The KEB children and adult groups perceived the 13 English phonemes correctly (*M* = 78, *SD* = 3; *M* = 72, *SD* = 4, respectively) more often than the KM child group (*M* = 59, *SD* = 9).

A two-way repeated measures ANOVA was conducted to compare the three age-language status groups (G), with sound type (S) at two levels (i.e., consonants vs. vowels) and familiarity (F) at two levels (i.e., familiar vs. unfamiliar sounds). Results of this ANOVA are presented in Table 13. The main effect for group was significant, $F_{2, 30}$ = 15.835, *p* = .000; with a large effect size, $\eta_p^2$ = .531, and observed power, *P* = .999. A Scheffe's posthoc test showed that the KEB children (*M* = 78, *SD* = 3) performed significantly better than the KM children (*M* = 59, *SD* =9), *p* = .000, and also the KEB adults (*M* = 72, *SD* = 4) performed significantly better than the KM children, *p* = .007. No significant group interactions were found.

Perceptual performance on consonants (*M* = 74, *SD* = 12) was significantly better than vowels (*M* = 60, *SD* = 12), $F_{2, 30}$ = 140.986, *p* = .000; with a large effect size, $\eta_p^2$ = .834, and a great observed power, *P* = 1.000. Perceptual performance on familiar sounds (*M* = 73, *SD* = 16) was significantly better than unfamiliar sounds (*M* = 63, *SD* = 12), *F*2, 30 = 16.443, *p* = .000, with a moderate effect size, $\eta_p^2$ = .370, and observed power, *P* = .975. To ensure that each of the individual comparisons was truly significant, a Bonferroni correction was also conducted, which confirmed that the significance of the main effects for S (*p* = .000) and F (*p* = .000) was unlikely to have occurred by chance.

In the repeated measures ANOVA, the first-order interaction of S x F was also significant, $F_{2, 30}$ = 51.621, *p* = .000; with a large effect size, $\eta_p^2$ = .648, and observed power, *P* = 1.000. For comparisons among the four sound categories (i.e., familiar and unfamiliar consonants and vowels; FC, UC, FV, and UV, respectively), a set of paired *t*-tests was conducted for a follow-up of the main analysis and for each of the three age-language status groups. Table 14 is a brief summary of the results from several paired *t*-tests. The paired *t*-tests showed that FC (*M* = 88, *SD* = 12) was perceived correctly significantly more often than UC (*M* = 64, *SD* =13), $t_{30}$ = 14.804, *p* = .000; however, FV (*M* = 57, *SD* = 22) was perceived similarly to UV (*M*

= 62, *SD* = 14), $t_{30}$ = -1.310, *p* = .200. Similar patterns—FC better than UC and FV similarly to UV—were observed within each group (see Table 12). In summary, all participants taken together correctly perceived (a) consonants significantly better than vowels, (b) familiar sounds significantly more often than unfamiliar ones, and (c) familiar consonants significantly more often than unfamiliar ones, but (d) familiar and unfamiliar vowels similarly.

### *Select Group Comparisons*

In this section, an age effect were examined (i.e., the KEB child vs. adult groups) separately from the comparison of the three age-language status groups (i.e., the KEB child, KEB adult, and KM child groups). A group effect of proficiency (i.e., the ED, BAL, and KD groups) was also examined within the KEB group.

#### *Comparisons of the Two Age Groups*

The three age-language status group comparisons showed no significant group difference between the KEB children and KEB adults. In order to avoid a confounding of age and language status, an age comparison was conducted in a separate ANOVA. Table 15 is a display of results of a two-way repeated measures of ANOVA, with sound type and familiarity as within-group factors and age as a between-group factor. The main effect for age became significant in this age-only comparison, $F_{1, 14}$ = 6.952, *p* = .021, with a moderate effect size, $\eta_p^2$ = .348, and a moderate observed power, *P* = .684. The age effect probably failed to reach significance in the previous three group comparison because the KM child group contributed to greater variability among the participants and thus a larger error term. Overall, KEB children (*M* = 78, *SD* = 3) performed better on the NSCM task than KEB adults (*M* = 72, *SD* =4).

The remainder of the results for the age ANOVA conforms to those of the previous 3-group ANOVA: The main effects for S and F were significant for the age comparison. Consonants were identified significantly better than vowels, $F_{1, 14} = 69.360$, $p = .000$, $\eta_p^2 = .842$, $P = 1.000$ ($M = 83$, $SD = 5$ for consonants; $M = 69$, $SD = 9$ for vowels). Familiar sounds ($M = 81$, $SD = 6$) were perceived correctly more often than unfamiliar sounds ($M = 72$, $SD = 7$), $F_{1, 14} = 28.877$, $p = .000$, with a large effect size, $\eta_p^2 = .690$, and a large observed power, $P = .999$. The first-order interaction of S x F was also significant, $F_{1, 14} = 33.160$, $p = .000$, $\eta_p^2 = .718$, $P = 1.000$. A paired *t*-test was conducted as a follow-up posthoc analysis, which revealed that FC ($M = 94$, $SD = 3$) was perceived correctly more often than UC ($M = 75$, $SD = 6$), $t_{14} = 16.316$, $p = .000$. FV ($M = 68$, $SD = 12$), but perceived similarly to UV ($M = 69$, $SD = 11$), $t_{14} = -.415$, $p = .684$. No group interaction or the second-order interaction of S x F x G was statistically significant.

*Comparisons of the Three Proficiency Subgroups*

Perception on the NSCM task by S, F, and S x F was compared among the three English-proficiency subgroups of KEB children as shown in Table 16. The repeated measures ANOVA showed a significant main effect for proficiency, $F_{2, 7} = 11.484$, $p = .014$. A Scheffe's posthoc test, following this repeated measures ANOVA, showed that the ED group performed significantly better ($M = 81$, $SD = 3$) than the KD group ($M = 74$, $SD = 2$), $p = .014$, but other group differences were not significant. The interaction of S x G was also significant, $F_{2, 7} = 7.429$, $p = .032$. A posthoc test via one-way ANOVA showed perception of vowels was significantly different among the three subgroups, $F_{2, 7} = 26.249$, $p = .002$; however, perception of consonants was not. The three subgroups performed similarly for consonants ($M = 86$, $SD = 1$

for the ED group; *M* = 81, *SD* = 4 for the BAL group; *M* = 86, *SD* = 4 for the KD group). A Scheffe's posthoc test, following up this ANOVA, showed the ED group perceived vowels correctly (*M* = 80, *SD* = 4) more often than the KD group (*M* = 63, *SD* = 1), *p* = .003. Also, the BAL group perceived vowels correctly (*M* = 79, *SD* = 2) more often than the KD group, *p* = .009. Interactions of S x G and S x F x G were not significant.

Main effects for both of the within-group factors were also significant: sound type, $F_{1,7}$ = 51.657, *p* = .001, and familiarity, $F_{1,7}$ = 9.539, *p* = .027. As in the repeated measures ANOVA for the age groups, the repeated measures ANOVA for the proficiency subgroups showed that consonants (*M* = 85, *SD* = 3) were perceived correctly more often than vowels (*M* = 73, *SD* = 9). Again, familiar sounds (*M* = 83, *SD* = 5) were perceived correctly more often than unfamiliar ones (*M* = 76, *SD* = 6). The first-order interaction of S x F was also significant, $F_{2,7}$ = 17.169, *p* = .009. In order to follow-up this significant interaction, a paired *t*-test was conducted as a posthoc test on the four cell-means. The result of the posthoc test showed that the children perceived FC (*M* = 94, *SD* = 3) significantly better than UC (*M* = 78, *SD* = 4), $t_7$ = 20.220, *p* = .000, but FV (*M* = 72, *SD* = 10) similarly to UV (*M* = 74, *SD* = 12).

### *Perceptual Performance for Individual Sounds*

Participants showed similar patterns of consonant and vowel perception. Table 17 and Figure 5 are displays of the percent correct for the 10 consonants and the 3 vowels perceived by the KEB children, the KEB adults, and the KM children. Overall, perceptual performance was generally better for familiar consonants /p, k, s, ʧ/ than unfamiliar ones /b, g, f, θ, z, ʤ/ across the three age-language status groups. All participants taken together performed poorest on the identification of /f/ (*M* = 34, *SD* = 15) and /θ/ (*M* = 56, *SD* = 26), but best on the identification of

/p/ (*M* = 93, *SD* = 14), /k/ (*M* = 92, *SD* = 15), and /g/ (*M* = 95, *SD* = 10). The five consonants /b, f, θ, z, ʤ/ seemed to be identified poorly across the groups and the three consonants /p, k, g/ seemed to be identified at a high rate across the groups. Even though general patterns were observed, percent correct for some consonants diverged among the groups. A MANOVA was conducted for percent correct for each consonant and vowel for the three age-language status groups to determine which phonemes were perceived differently among the groups. Its results are displayed in Table 18.

*Group Differences for Each Phoneme*

The MANOVA showed that the group difference in perceptual performance on the NSCM task was significant for the stimuli /ʧ, b, θ, z, ʤ, ɛ, æ/ but not for the stimuli /p, k, s, g, f, ɪ/. (See Figure 6.) The consonant /ʧ/ was the only familiar consonant that showed a group difference, $F_{2, 30} = 5.825$, $p = .008$. According to a Scheffe post-hoc test, following up this MANOVA, the KEB children (*M* = 96, *SD* = 3) perceived /ʧ/ correctly more often than the KM children (*M* =77, *SD* = 20), $p = .028$; the KEB adults (*M* = 95, *SD* = 2) perceived it correctly more often than the KM children, $p = .040$; but the bilingual children and adults perceived it at a similar rate.

Among unfamiliar consonants, percent correct for /g, f/ was similar among the groups, where /g/ was perceived at a high rate but /f/ at a low rate across the groups. Percent correct for /b/ was significantly different among the groups, $F_{2, 30} = 3.904$, $p = .032$; however, the Scheffe was not significant among the three groups. This is possibly due to large standard deviation in the KEB and KM children. In general, the KEB children and adults (*M* = 85, *SD* = 16; *M* = 89, *SD* = 8, respectively) seemed to perceive /b/ correctly more often than the KM children (*M* = 69, *SD* = 23).

Percent correct for /θ/ was significantly different among the groups, $F_{2, 30}$ = 17.464, *p* = .000. According to the Scheffe, the KEB children (*M* = 75, *SD* = 10) identified it correctly more often than the KEB adults (*M* = 61, *SD* = 19), *p* = .006, and the KM children (*M* = 33, *SD* = 20), *p* = .000. Percent correct for /z/ was significantly different among the groups, $F_{2, 30}$ = 30.207, *p* = .000. The Scheffe showed that the KEB children (*M* = 89, *SD* = 5) identified /z/ correctly more often than the KEB adults (*M* = 70, *SD* = 15), *p* = .001, and the KM children (*M* = 43, *SD* = 16), *p* = .000. Percent correct for /ʤ/ differed significantly among the groups, $F_{2, 30}$ = 24.398, *p* = .000, where the KEB children (*M* = 88, *SD* = 6) identified /ʤ/ correctly more often than the KEB adults (*M* = 79, *SD* = 11), *p* = .001, and the KM children (*M* = 55, *SD* = 14), *p* = .000.

Percent correct for the stimulus /ɛ/ was significantly different among the groups, $F_{2, 30}$ = 5.113, *p* = .013. According to a Scheffe's posthoc, the KEB children (*M* = 72, *SD* = 10) identified /ɛ/ similar to the KEB adults (*M* = 64, *SD* = 13), but significantly better than the KM children (*M* = 46, *SD* = 24), *p* = .020. Percent correct for the stimulus /ɪ/ was low, but not significantly different among the groups. Percent correct for the stimulus /æ/ was also significantly different among the groups, $F_{2, 30}$ = 3.692, *p* = .038. A Scheffe showed that the KEB children (*M* = 78, *SD* = 11) identified /æ/ similar to the KEB adults (*M* = 71, *SD* = 15), but significantly better than the KM children (*M* = 59, *SD* = 20), *p* = .049.

In summary, two patterns were observed in the MANOVA: the consonant /ʧ/ and the vowels /ɛ, æ/ were perceived correctly more often by the KEB children and adults than the KM children; on the other hand, the consonants /θ, z, ʤ/ were perceived correctly more often by the KEB children than the KEB adults and KM children.

*Perceptual Confusions for Each Phoneme*

Perceptual confusion, possibly due to L1 influence, may be inferred from any low rates of percent correct. Therefore, sound confusions for the seven phonemes that were identified differently among the groups and two phonemes that were identified poorly were examined: they are /ʧ, b, θ, z, ʤ, ɛ, æ/ and /f, ɪ/, respectively. Table 19 is a presentation of perceptual confusions for these six consonants and all three vowels (The confusion matrices for all consonants and vowels are presented in Appendix E). Both the KEB children and adults identified /ʧ/ at high rates ( 96% and 95%, respectively) and occasionally misidentified it as its voiced counterpart /ʤ/ and /k/. Unlike the bilinguals, the KM children misidentified it as not only /ʤ, k/ but also /θ, s, z/. Participants misidentified /b/ as /f/ 6% of the time or as /θ/ 7% of the time. The bilinguals showed a /b/ → /θ/ confusion more often than a /b/ → /f/ confusion; whereas the monolingual children showed the opposite pattern. For the stimulus /f/, all groups showed poor identification, misidentifying it as /b/ 18% of the time or /θ/ 33% of the time. Again, the bilinguals showed a /f/ → /θ/ confusion more often than /f/ → /b/; whereas, the monolinguals showed an opposite pattern. All participants misidentified /θ/ either as /f/ 18% of the time or /s/ 8% of the time. The stimulus /z/ was confused with its voiced counterpart /s/ 12% of the time or with /ʤ/ 17% of the time. A voicing confusion of /z/ → /s/ was less frequent than a place-manner confusion of /z/ → /ʤ/ in the KEB adults and the KM children. For the stimulus /ʤ/, all participants confused it as /z/ 13% of the time or /ʧ/ 5% of the time.

There were only a few unidirectional confusions: for example, /p/ → /b/ and /g/ → /k/ confusions never occurred; whereas, a /b/ → /p/ and /k/ → /g/ confusions occurred 4% and 2% of the time, respectively (see Appendix E for sound confusions that are not presented in Table 19). Most of the confusions were bidirectional but were not always symmetrical. For example, a /z/ → /s/ confusion occurred 12% of the time, but a /s/ → /z/ confusion occurred only 2% of the time. Perceptual confusions for the following consonants were bidirectional: the consonants /b/ and /f/, the consonants /θ/and /f/, and the consonants /z/ and /ʤ/ were mutually confused with each other. Overall, a /b/ ↔ /f/ confusion seemed to occur more often in the KM children (/b/ → /f/ 14% of the time and /f/ → /b/ 22% of the time) than in the KEB adults (/b/ → /f/ 3% of the time and /f/ → /b/ 19% of the time) and the KEB children (/b/ → /f/ 2% of the time and /f/ → /b/ 15% of the time). Similar to these patterns, a /f/ ↔ /θ/ confusion occurred more often in the KM children (/f/ → /θ/ 18% of the time and /θ/ → /f/ 24% of the time) than in the KEB adults (/f/ → /θ/ 19% of the time and /θ/ → /f/ 19% of the time) and the KEB children (/f/ → /θ/ 47% of the time and /θ/ → /f/ 12% of the time). A /z/ ↔ /ʤ/ confusion occurred more often in the KM children (/z/ → /ʤ/ 29% of the time and /ʤ/ → /z/ 25% of the time) and the KEB adults (/z/ → /ʤ/ 19% of the time and /ʤ/ → /z/ 13% of the time) than in the KEB children (/z/ → /ʤ/ 2% of the time and /ʤ/ → /z/ 2% of the time).

Vowel confusions showed the same patterns across the three groups. The familiar vowel /ɛ/ was confused mostly with /æ/, and the unfamiliar vowels /ɪ/ and /æ/ were confused mostly with /ɛ/ across the three groups. Vowel confusions were also bidirectional and asymmetric. A /ɛ/ → /ɪ/ confusion occurred only 7% of the time, but a /ɪ/ → /ɛ/ confusion occurred 32% of the time. A /ɪ/ ↔ /æ/ confusion occurred least frequently (/ɪ/ → /æ/ 8% of the time and /æ/ →/ɪ/ 3%

of the time). A /ɛ/ ↔ /æ/ confusion occurred most frequently across the three groups (/ɛ/ → /æ/ 33% of the time and /æ/ →/ɛ/ 27% of the time). In general, taking both consonants and vowels into consideration, the KEB children showed less dispersion of confusion errors than the KEB adults and the KM children.

## Summary and Conclusions

The first research question asked if consonants are perceived differently from vowels. A set of repeated measures ANOVAs showed that consonants were perceived significantly better than vowels in all participants taken together (i.e., the three age-language status group comparison), in the KEB child and adult groups (i.e., the age comparison), and in all proficiency subgroups (i.e., the three proficiency subgroup comparison, among the KEB children). The percent correct for consonants was 85%, 81%, and 66%, but the percent correct for vowels was 73%, 63%, and 52% for the KEB children, the KEB adults, and the KM children, respectively. This finding was predicted based on the previous studies (Johnson et al., 2007; Romani et al., 1996).

The second research question asked if familiar sounds are perceived differently from unfamiliar sounds. Repeated measures ANOVAs showed that familiar sounds were perceived correctly more often than unfamiliar ones by all age-language status groups and all proficiency subgroups. These results are consistent with the predictions of PAM. Interactions were evident, however, in that participants perceived FC correctly more often than UC, but FV similarly to UV. Although results of the present study generally confirm the proposed hypothesis favoring familiar sounds, neither SLM nor PAM takes into consideration these interactions of familiarity with sound type (i.e., consonants and vowels) in predicting perceptual difficulties of L2 sounds.

The third research question asked if perceptual performance differs among the groups depending on age, language status, and L2 proficiency. Significant group effects were found for the three age-language status group comparison, age comparison, and proficiency comparison. When comparing the KEB children, the KEB adults, and the KM children, significant differences were found in perceptual performance between the KEB groups and the KM child group. To avoid confounding group effects, a select group comparison was conducted for age, in which the KEB children performed significantly better than the KEB adults. In the proficiency comparison, the ED subgroup performed significantly better than the KD subgroup. Therefore, results of the present study are consistent with the proposed hypothesis that younger or more experienced listeners will outperform older or less experienced ones.

The fourth research question asked if some English sounds are more difficult to correctly identify than others, regardless of a listener's L2 background. According to the MANOVA, perceptual performance for /b, θ, z, ʧ, ʤ/ and /ɛ, æ/ was significantly different among the groups. Two of the seven phonemes, /ʧ, ɛ/, are familiar sounds, but five of the seven phonemes, /b, θ, z, ʤ, æ/, are unfamiliar sounds. On the other hand, the remaining consonants /p, k, g, f, s/ and the vowel /ɪ/ were perceived in a similar way across the groups (at either a high or low rate across the groups). These results suggest that not only L1 influences the perception of L2 sounds, but also that the nature of individual sounds themselves also plays an important role in identifying L2 phonemes.

# CHAPTER IV

# NONSENSE SYLLABLE IMITATION TASK

## Rationale for a Production Task

Studies of phonological development (i.e., speech production) in English-speaking children have shown that the stops and nasals such as /p, b, t, d, k, g, m, n/ are typically acquired early in life, but fricatives such as /θ, ð/ and liquids are typically acquired later in life: closer to the age of 8 years (Dodd et al., 2003, Poole, 1934; Smit et al., 1990; Wellman et al., 1931). Phonological development in English-speaking *bilingual* children may not follow this general observation, but may diverge depending on one's L1 phonology. For instance, the Spanish consonant /j/ is acquired by the age of 3 years in Spanish-monolingual children (Paulson, 1991); whereas, the English consonant /j/ is acquired later—near the age of 4 years—in English-monolingual children (Dodd et al., 2003). Furthermore, children who are exposed to more than one language may show different patterns of phonological development from monolinguals. For example, Spanish fricatives (or spirants) were produced more accurately by Spanish-English bilingual children than by their Spanish monolingual peers (Goldstein & Washington, 2001).

In order to explain how one's L1 phonology interferes with the production of L2 sounds, SLM argued that L2 sounds similar to L1 sounds are produced less accurately than L2 sounds new to L1, especially by experienced L2 learners. This suggests that the interference from L1 phonology is stronger for similar L2 sounds but less strong for new L2 sounds. In the earlier literature review of this document, PAM was introduced mainly to account for results from perceptual tasks; however, a few recent studies showed that its extended interpretation also works on production of consonants (Halle & Best, 2007; Krebs-Lazendic & Best, 2007).

The present study examines which types of English sounds are produced well enough to avoid Korean interference in monolingual children who speak English as a foreign language (i.e., KM children), as well as bilingual children and adults who speak English as a second language (i.e., KEB children and adults). Phonological development of English sounds in KMs may differ at least from the ED and BAL subgroups of the KEB children, given that KMs began to learn English after the age of 7 years (Johnson & Newport, 1987; Mayberry & Eichen, 1991). Alternatively, if some English sounds lack universal ease of articulation (or ease of perception of their acoustic results), then all children (whether KMs or KEBs) may master these sounds late and produce them less accurately than other English sounds. For example, the consonants /t, k, s, m, n/ appear in English, Spanish, Cantonese, and Arabic. Among these consonants, the consonants /t, m/ were acquired as early as by 3 years old in English-, Spanish-, Cantonese-, and Arabic-speaking children; however, the consonant /s/ was acquired later in all children speaking these languages (Amayre & Dyson, 1998; Paulson, 1991; Smit et al, 1990; So & Dodd, 1995). Results of the present study may show which English sounds are more susceptible to interference from L1 phonology, and this should help us develop speech therapy or educational curricula for children who struggle with speaking English.

In addition, results from the NSIT will be compared to results from the NSCM task in order to see whether perceptual performance is related to production performance. In other words, the present study will examine L2 English sounds to see if sounds that are poorly identified are also poorly produced by the KEBs and KMs. This will allow us to see how independent of or dependent on perceptual performance production performance is, in speakers who are the most likely to experience L1 interference, namely, those with the *least L2 experience* (the KM group) and those with the *most L1 experience* (the ED group). Comparisons between familiar and unfamiliar sounds and between perceptual and production performance should help to determine whether poor performance on the NSIT is caused by L1 interference, by perceptual difficulties, or by motoric difficulties that are intrinsic to the sound.

## Research Questions and Hypotheses

The present study examines which sound types are more challenging for KEBs and KMs to produce on the Nonsense Syllable Imitation Task (NSIT) and also whether production performance can be explained by SLM and PAM. Four research questions were asked in the present study.

### *Research Question 1: Familiar vs. Unfamiliar Sounds*

Determining whether familiar or unfamiliar sounds are more challenging should help direct language educators, including speech-language pathologists, to sounds that should be addressed with English second-language or foreign-language learners who have difficulty speaking and listening to English. To answer this question, rate of correct production of familiar sounds was compared to that of unfamiliar ones.

1. Do Korean-English bilinguals and Korean monolinguals produce familiar English sounds differently from unfamiliar ones?

SLM argues that unfamiliar sounds are produced more native-like than familiar sounds by experienced L2 learners, because they often fail to notice subsegmental differences between L1 and similar (familiar) L2 sounds. Therefore, SLM predicts that production of unfamiliar sounds in KEBs should be closer to the native-like norm for English than are production of familiar sounds. Nonetheless, this pattern should be the opposite in KMs; that is, KMs should produce unfamiliar sounds less correctly than familiar ones, due to their lack of experience with L2. PAM, however, does not make explicit predictions about L2 production. Nevertheless, an extended interpretation of PAM predicts that *production* of familiar nonnative sounds (which include similar features to L1 phonology) will be produced more accurately than unfamiliar nonnative sounds.

Hypothesis 1a: Korean-English bilinguals (KEBs) will produce unfamiliar sounds more accurately than familiar ones, but this will be the opposite in Korean monolinguals (KMs), as predicted by SLM.

Alternate Hypothesis 1b: Both KEBs and KMs will produce familiar sounds more accurately than unfamiliar ones, as predicted by PAM.

*Research Question 2: Consonants vs. Vowels*

The second question asks if production performance differs between consonants and vowels.

2. Do Korean-English bilinguals and Korean monolinguals produce English consonants differently from vowels?

Neither SLM nor PAM provides predictions about whether production performance for consonants differs from that for vowels. Movement of the tongue would seem to be more continuous for vowel production than for consonant production. Consonants would seem to be more discrete in nature. Production of the three vowels chosen for the present study might be more confusing than consonant production, because all three vowels are considered lax in some distinctive feature systems (Lowe, 2010) and probably they are very close in articulatory and perceptual space (i.e., they would all be considered front vowels, articulated in the palatal region of the mouth). Yet, it is not sure whether production for consonants would be better than that for vowels.

Hypothesis 2: Korean-English Bilinguals and Korean monolinguals will produce English consonants differently from vowels, at least in immediate imitation.

*Research Question 3: Group Comparisons*

Results on the NSIT were compared among the groups with regard to age, proficiency, and language status.

3. Does production performance differ due to age (i.e., bilingual children vs. adults), language status (i.e., bilingual vs. monolingual children), and L2 proficiency (i.e., comparisons of English-dominant, balanced, and Korean-dominant bilingual children)?

Based on previous studies, early or experienced L2 learners are more likely to succeed in achieving native-like production with L2 English sounds than late or inexperienced L2 learners. Therefore, in the present study bilingual children were expected to produce L2 sounds correctly at a higher rate than bilingual adults and KM children, and KD children were expected to perform the poorest among the three L2 proficiency subgroups.

Hypothesis 3: The younger and more highly experienced L2 learners are, the better they will perform production tasks for L2 speech sounds, such as the Nonsense Syllable Imitation Task.

*Research Question 4: Correlations between Perception and Production*

The fourth question asked if participants who demonstrate less perceptual confusion in English would have better production of English phonemes. SLM was developed based on both production and perception tasks; however, PAM was developed based on perception tasks, alone.

L2 learning models suggested by SLM and PAM seem to assume that L2 perceptual learning works similarly to L2 production learning.  By answering to this question, the present study will examine whether perceptual learning mechanisms for L2 differ from production learning mechanisms.

4. Is perceptual confusion of speech sounds related to rate of correct speech sound production, as measured by simple imitation of nonsense syllables?

To answer this question, Pearson-product moment correlations were calculated between performance on the NSCM task and on the NSIT for overall performance, and performance for consonants, vowels, familiar sounds, and unfamiliar sounds.

It has been well documented that young L2 learners generally reach native-like ability for speech production, whereas older L2 learners generally continue to have heavy foreign accents (Johnson & Newport, 1987; Mayberry & Eichen, 1991; Ploog, 1984; Yeni-Komshian, Flege, & Liu, 2000).  Doubtless, one's early linguistic experiences have a lasting effect on speech production and perception; however, the accuracy of L2 perception seems to grow even in older L2 learners while the accuracy of L2 production does not.  McClasky, Pisoni, and Carrell (1983) and Pisoni, Lively, and Logan (1994) showed that native English listeners can identify and discriminate nonnative voicing contrasts without any special training and their discrimination accuracy greatly increases with a very short period of training.  Furthermore, many studies of children's speech perception have shown that perceptual ability continues to expand throughout the life-span (Nittrouer, 1993; Sussman, 1992; Nishi et al., 2010) unlike production development which is typically completed by age 8 or so (Dodd, Holm, Hua, & Crosbie, 2003).

Findings from various studies imply that early linguistic experience has an impact more on L2 production than on L2 perception. In other words, young L2 learners are more likely to achieve native-like L2 production than older L2 learners with same amount of L2 experience; however, older L2 learners may achieve native-like L2 perception with increasing L2 experience, just as young L2 learners do. Therefore, I hypothesize that the correlation between perceptual performance and production performance should be observed in young L2 learners but not in older L2 learners.

Hypothesis 4: Child participants in the present study (i.e., Korean-English bilingual children and Korean monolingual children) may reveal a close correlation between perceptual and production performance; whereas, adult participants (i.e., Korean-English bilingual adults) may not.

The present study will examine whether articulatory patterns in KEBs and KMs are similar to perceptual patterns. Results of the present study will determine whether an account of L2 learning should separate perception from production. Such knowledge might help us develop better educational methodologies for teaching L2 English, where methods for teaching listening might differ from those for teaching articulation. Furthermore, answers to this question may shed light on our understanding of whether L2 production learning is perceptually-based or not.

## Method

### *Participants*

KEBs and KMs who participated in the NSCM task were also asked to perform the NSIT, i.e., 9 KEB children, 8 KEB adults, and 18 KM children (for further detail about participants, see the "Participants" section in Chapter II).

*Stimuli*

The same 30 stimuli from the NSCM task served as stimuli for the NSIT. In other words, the stimuli were presented in a CV template that consisted of one of the four familiar consonants /p, k, s, ʧ/ or six unfamiliar ones /b, g, f, θ, z, ʤ/ with the one familiar vowel /ɛ/, or the two unfamiliar ones /ɪ, æ/.

*Task Procedure*

At the end of the third or fourth experimental sessions for the NSCM task, each of the 30 stimuli was presented one at a time in random order. For each stimulus presentation, the participants were asked to listen to a syllable over headphones while they simultaneously saw a visual cue on the computer screen (i.e. the English print symbol used in the NSCM task as the button for that stimulus. See Figure 4). While listening to and looking at the stimulus, the participant imitated the stimulus two times in succession. His or her oral responses were audio recorded on a laptop computer (Satellite S205, Toshiba) via Audacity, free software that allows recording and editing of speech sounds, using a headset microphone (Plantronics, Audio 350 Ultimate Performance Headset). The headset microphone was used to keep a constant distance from the speaker's mouth to the microphone (of approximately 2 inches), to allow retrieval of high quality spontaneous productions of speech sounds.

The 30 syllables, each produced twice by the participant, were transcribed in the International Phonetic Alphabet (IPA) and judged on the correctness of pronunciation by each of two transcribers—the author and a native English speaker—trained in phonetic transcription. Both transcribers had a list of the sequence of target sounds to which they were listening and judged whether the syllable was correctly produced; and, if not, what syllable was produced.

The author made the final decision about any disagreements between the two transcribers. In order to improve methodological reliability in the present study, a single evaluator calculated percent correct for consonants and vowels on the NSIT. Participants who had already performed the NSIT before the new evaluator were reexamined and recalculated. The new evaluator was blind to the evaluation made by the author or the previous evaluator. The author served as a second evaluator, for determining interjudge reliability throughout the study.

## Results

Percent correct for each sound (10 consonants and 3 vowels) was calculated and normalized via arcsine transformation prior to statistical analyses. Several two-way repeated measures ANOVAs and posthoc tests were conducted to determine statistical significance of production differences within and between groups. Production performance on the NSIT was compared among the KEB children, the KEB adults, and the KM children (i.e., the three age-language status groups), between the KEB child vs. adult groups (i.e., age), between the KEB vs. KM child groups (i.e., language status), and among the ED, BAL, and KD groups (i.e., the three proficiency subgroups, or proficiency).

### *Familiar vs. Unfamiliar Consonants and Vowels*

#### *Comparisons of the Three Age-Language Status Groups*

Percent correct on the NSIT for familiar and unfamiliar consonants and vowels is displayed in Table 20. Overall, consonants ($M$ = 86, $SD$ = 9) were produced correctly at higher rates than the three vowels ($M$ = 76, $SD$ = 14), and familiar sounds ($M$ = 90, $SD$ = 7) were produced correctly at higher rates than unfamiliar ones ($M$ = 79, $SD$ = 16). The rate of correct articulation of the 13 English phonemes was higher for the KEB children ($M$ = 91, $SD$ = 7) than for the KM children ($M$ = 79, $SD$ = 11).

Table 21 is a display of a two-way repeated measures ANOVA with two within-group factors, sound type (consonants vs. vowels) and familiarity (familiar vs. unfamiliar sounds), and one between-group factor for the three age-language status groups. The between-group effect was significant, $F_{2, 34} = 4.325$, $p = .022$, with a weak effect size, $\eta_p^2 = .213$, but large observed power, $P = .710$. A Scheffe test showed that the KEB children ($M$ = 91, $SD$ = 7) performed significantly better than the KM children ($M$ = 79, $SD$ =11), $p = .026$. The main effects for S and F were also significant, where consonants ($M$ = 86, $SD$ = 9) were produced more accurately than vowels ($M$ = 76, $SD$ =14), $F_{2, 34} = 16.902$, $p = .000$, with a moderate effect size $\eta_p^2 = .346$, and a great observed power $P = .979$, and familiar sounds ($M$ = 90, $SD$ = 7) were produced more accurately than unfamiliar ones ($M$ = 79, $SD$ = 16), $F_{2, 34} = 23.590$, $p = .000$, $\eta_p^2 = .424$, and $P = .997$. The first order interaction of S x F was not significant.

The group effect showed interactions with familiarity, $F_{2, 34} = 10.818$, $p = .000$. A MANOVA showed that unfamiliar sounds were produced significantly better in the KEB child group ($M$ = 91, $SD$ = 9) than in the KM child group ($M$ = 68, $SD$ = 16), $F_{2, 34} = 8.859$, $p = .001$. The second order interaction of S x F x G was also significant, $F_{2, 34} = 21.228$, $p = .000$, $\eta_p^2 = .570$, and $P = 1.000$. According to results of a follow-up MANOVA and its Scheffe's posthoc test, familiar consonants, $F_{2, 34} = 12.258$, $p = .000$, were produced more accurately in the KEB child group than in the KEB adult group, $p = .000$, and in the KM child group, $p = .001$. Unfamiliar consonants, $F_{2, 34} = 4.007$, $p = .028$, were produced more accurately in the KEB child group than in the KM child group, $p = .029$. Familiar vowels, $F_{2, 34} = 4.120$, $p = .026$, were produced less accurately in the KEB child group than in the KM child group, $p = .026$. Unfamiliar vowels, $F_{2, 34} = 11.913$, $p = .000$, were produced more accurately in the KEB child group than in the KEB adult group, $p = .014$, and the KM child group, $p = .000$.

The interaction of S x F was not significant in the repeated measures ANOVA; however, paired *t*-tests for each of the three age-language status groups, as presented in Table 22, indicate that the lack of significance for the S x F interaction in the main analysis is due to different patterns of production among the three groups. The KEB children produced familiar consonants ($M = 99$, $SD = 3$) more accurately than unfamiliar ones ($M = 91$, $SD = 6$), $t_8 = 2.742$, $p = .025$, but familiar vowels ($M = 79$, $SD = 14$) less accurately than unfamiliar ones ($M = 91$, $SD = 14$), $t_8 = -3.337$, $p = .010$. The KEB adults produced familiar consonants ($M = 88$, $SD = 7$) similarly to unfamiliar consonants ($M = 86$, $SD = 10$), $t_7 = .337$, $p = .746$, but familiar vowels ($M = 91$, $SD = 13$) more accurately than unfamiliar ones ($M = 71$, $SD = 8$), $t_7 = 5.128$, $p = .001$. The KM children produced familiar consonants ($M = 90$, $SD = 6$) more accurately than unfamiliar ones ($M = 77$, $SD = 15$), $t_{17} = 3.511$, $p = .003$ and also familiar vowels ($M = 92$, $SD = 13$) more accurately than unfamiliar ones ($M = 59$, $SD = 8$), $t_{17} = 6.806$, $p = .000$. Different patterns of an S x F interaction across the groups appeared to have caused insignificant interaction in the main repeated measures ANOVA; therefore, paired *t*-tests were conducted to see whether familiar or unfamiliar consonants and vowels were produced more accurately within each group.

*Select Group Comparisons*

*Comparisons of the Two Age Groups*

A two-way repeated measures ANOVA was conducted with Sound type and familiarity as within-group factors and Age as a between-group factor. Its results are presented in Table 23.

Production performance was significantly different between the KEB children and adults, $F_{1, 16}$ = 8.957, $p$ = .009, $\eta_p^2$ = .374, and $P$ = .799. Overall, the KEB children produced English nonsense syllables ($M$ = 91, $SD$ = 7) more accurately than the KEB adults ($M$ = 82, $SD$ = 7). The main effect for S was significant, $F_{1, 16}$ = 11.635, $p$ = .004, $\eta_p^2$ = .437, and $P$ = .890, where consonants were produced ($M$ = 90, $SD$ = 7) more accurately than vowels ($M$ = 82, $SD$ = 11). The main effect for F was also significant, $F_{1, 16}$ = 5.409, $p$ = .034, $\eta_p^2$ = .265, and $P$ = .585, where familiar sounds ($M$ = 89, $SD$ = 7) were produced more accurately than unfamiliar ones ($M$ = 85, $SD$ = 10).

A group interaction of S x G was not significant. Both the KEB children and adults produced consonants ($M$ = 94, $SD$ = 4 and $M$ = 87, $SD$ = 7, respectively) more accurately than vowels ($M$ = 87, $SD$ = 13 and $M$ = 77, $SD$ = 8, respectively). A group interaction of F x G, however, was significant, $F_{1, 16}$ = 5.409, $p$ = .034, $\eta_p^2$ = .265, and $P$ = .585. The KEB children produced familiar sounds ($M$ = 89, $SD$ = 6) correctly at similar rates to unfamiliar ones ($M$ = 91, $SD$ = 9); on the other hand, the KEB adults produced familiar sounds ($M$ = 89, $SD$ = 9) more accurately than unfamiliar ones ($M$ = 78, $SD$ = 8). A MANOVA as a posthoc test showed that production performance on unfamiliar sounds contributed to a significant effect for the F x G interaction, $F_{1, 16}$ = 14.216, $p$ = .002.

A first-order interaction of S x F was not significant; however, a second-order group interaction of S x F x G was significant, $F_{1, 16}$ = 28.415, $p$ = .000, with a large effect size $\eta_p^2$ = .654, and $P$ = .999. A MANOVA showed that the KEB children produced familiar consonants ($M$ = 99, $SD$ = 3), $F_{1, 16}$ = 21.986, $p$ = .000, and unfamiliar vowels ($M$ = 91, $SD$ = 14), $F_{1, 16}$ = 17.830, $p$ = .001, more accurately than the KEB adults did ($M$ = 88, $SD$ = 7 and $M$ = 71, $SD$ = 8, respectively).

*Comparisons of the Three Proficiency Subgroups*

Production performance on the NSIT was compared across the three proficiency subgroups of the KEB children (i.e., the ED, BAL, and KD groups), depending on their linguistic proficiency in English. Descriptive statistics of production performance on the NSIT for the three subgroups is presented in Table 24. A repeated measures ANOVA with English proficiency as the between-group factor are presented in Table 25. Production performance on the NSIT was not significantly different among the three subgroups: the ED, BAL, and KD groups. Nevertheless, the main effect for S was significant, $F_{2, 8} = 16.425$, $p = .008$, with a large effect size $\eta_p^2 = .720$, and $P = .902$, where consonants ($M = 94$, $SD = 4$) were produced more accurately than vowels ($M = 87$, $SD = 13$). The main effect for F was not significant. No group interactions were significant: neither S x G nor F x G. An interaction of S x F was significant, $F_{2, 8} = 14.675$, $p = .009$, $\eta_p^2 = .710$, $P = .888$. A paired *t*-test showed that familiar consonants ($M = 99$, $SD = 3$) were produced more accurately than unfamiliar ones ($M = 91$, $SD = 6$), $t_8 = 2.742$, $p = .025$, and that familiar vowels ($M = 79$, $SD = 14$) were produced *less* accurately than unfamiliar ones ($M = 91$, $SD = 14$), $t_8 = -3.337$, $p = .010$ (see also Table 22).

*Correlations between Perceptual and Production Performance*

Pearson product-moment correlation coefficients were calculated to see the relation between perceptual performance on the NSCM task and production performance on the NSIT. Pearson correlation coefficients were calculated for nine different sound groups: (a) all sounds, (b) consonants, (c) vowels, (d) familiar sounds, (e) unfamiliar sounds (f) familiar consonants, (g) unfamiliar consonants, (h) familiar vowels, and (i) unfamiliar vowels. Table 26 is a display of nine coefficients for all participants and for the KEB children, the KEB adults, and the KM children. Two different correlations are presented in the table, group controlled and group uncontrolled. Group controlled correlations allow us to determine the correlation after removing the third variable that may be related to the correlations between the two variables, i.e., group effect, based on the performance of each group, whereas group uncontrolled correlations are calculated based on the performance of all the participants taken together as a whole. In most cases, a partial correlation is smaller than the original correlation without the controlled effect. When the group effect was controlled, correlations between perception and production become insignificant.

The following significant correlations were observed: for consonants, $r = .37$, $p = .044$; for unfamiliar sounds, $r = .40$, $p = .027$; for unfamiliar consonants, $r = .38$, $p = .036$; and for unfamiliar vowels, $r = -.30$, $p = .013$. Though significant, these correlations are modest at best. The correlation coefficient was positive for familiar consonants, whereas it was negative for unfamiliar vowels. This suggests that an individual who identified unfamiliar consonants with high rates also produced those consonants correctly at high rates; on the other hand, an individual who identified unfamiliar vowels at high rates produced those vowels correctly at low rates.

## Summary and Conclusions

The first research question asked if rate of correct production of familiar sounds differs from that of unfamiliar ones. Rate of correct production of familiar sounds was significantly different from that of unfamiliar ones in the three age-language status group comparison and age group comparison, but not in the proficiency subgroup comparison. Effect sizes were moderate. In general, Alternate Hypothesis 1b, the extended interpretation of PAM, predicted the results of comparison among the three age-language status groups. Nevertheless, the proficiency subgroups of the KEB children showed no difference of production performance between familiar and unfamiliar sounds; thus, the proficiency comparison does not confirm the hypothesis.

The second research question asked if the rate of correct production of consonants in English nonsense syllables differs from that of vowels in the KEB child group, the KEB adult group, and the KM child group. The rate of correct production for consonants was generally higher than that for the three vowels in the three age-language status group comparison, in the age group comparison, and in the proficiency subgroup comparison, with moderate and large effect sizes. Among the three age-language status groups, the rate of correct production of consonants was 94%, 87%, and 83%, and that of vowels was 87%, 77%, and 70% in the KEB child, KEB adult, and KM child groups, respectively. Among the three proficiency subgroups of the KEB children, the rate of correct production of consonants was 93%, 95%, and 96%, and that of vowels was 91%, 94%, and 78% in the ED, BAL, and KD groups, respectively. These results are consistent with the proposed hypothesis, that production of L2 consonants is less challenging than that of the three L2 vowels I selected as stimuli. It should be noted, though, that absolute differences on the NSIT were generally small, with all F and S means above 70%.

The third research question asked if production performance on the NSIT differs among the groups depending on age, language status, and proficiency. Repeated measures ANOVAs showed that overall performance on the NSIT was significantly higher for KEB children than for KEB adults and the KM children. Among the three subgroups of the KEB children, no group differences were observed.

The fourth research question asked if perceptual performance on the NSCM task is closely related to production performance on the NSIT. Pearson product-moment correlation coefficients indicated that rate of identification is significantly and positively correlated with rate of correct articulation for consonants, $r = .37$, unfamiliar sounds, $r = .40$, and unfamiliar consonants, $r = .38$. Note that these correlations are modest. Interestingly, identification rate for unfamiliar vowels is significantly and negatively correlated with rate of correct articulation, $r = -.30$. The results from the present study do not confirm the hypothesis that there would be a close relation between perceptual and production performance for KEB children, but not KEB adults.

# CHAPTER V

# NONWORD REPETITION TASK

## Rationale, Research Questions and Hypotheses

A nonword repetition task (NRT) was chosen to assess the participants' phonological awareness of English as well as their speech production performance for pronunciation of longer words. Previous studies on bilingualism have shown that perception and production of L2 sounds are greatly influenced by L1 phonology. For example, Halle and Best (2007) showed that the /dl/ and /tl/ clusters in word-initial position are poorly identified by the English and French speakers, because those clusters are phonotactically illegal in word-initial position in English and French. Phonological awareness in one's native language (L1) would be similar or even equal for experienced and inexperienced L2 learners. Nonetheless, phonological awareness in L2 may differ among L2 learners, depending on AOA, LOR, the speaking environment for and usage of L2, and so on. These factors were found to influence L2 vowels, particularly unfamiliar ones, in the ERP pilot study for perception and production. Therefore, the present study asked whether phonological awareness in L2 English is also correlated with perceptual and production performance for L2 sounds. With the addition of the KM group, these correlations may be stronger in the dissertation study than in the ERP study. With the least amount of knowledge of L2 vowels, this group should be the least likely to transfer any L2 perception or production knowledge to L2-like word pronunciation.

Many studies have pointed out that knowledge-dependent, norm-referenced tests are inherently biased against test takers from minority backgrounds (Adler & Birdsong, 1983; Bishop, North & Donlan, 1996; Campbell, Dollaghan, Needleman, and Janosky, 1997;

Dollaghan & Campbell, 1998; Munson, Edwards, & Beckman, 2005; Seymour, 1992; Taylor & Payne, 1983; Terrel & Terrel, 1983; Weismer, Laures, Jeng, Kent, & Kent, 2000). Because of increasingly substantiated claims that test stimuli, methods, and concepts are derived from white, middle class school settings and the mainstream culture of English-speaking communities (Campbell, Dollaghan, Needleman, and Janosky, 1997; Van Kleeck, 1994; Wyatt, 1995), Dollaghan and Campbell (1998) developed the Nonword Repetition Task (NRT). The task is purported to minimize the test takers' need for familiarity with culture-dependent factors and to maximize scoring accuracy and reliability of testing.

Research Question: Is perceptual and production performance on English nonsense syllables related to phonological awareness in L2, as measured by repetition of nonsense words on the Nonword Repetition Task that resemble the English language?

This question asked if participants who demonstrate less perceptual confusion in English would have stronger phonological awareness in English than participants with more perceptual confusions in English. The present study will examine whether correlations between performance on the NSCM task and phonological awareness in L2 English are significant in the three age-language status and proficiency subgroups.

Hypothesis: L2 or FL learners whose phonological awareness is greater, as measured by the Nonword Repetition Task (NRT) for English, will perceive and produce English nonsense syllables better, as measured by the Nonsense Syllable Confusion Matrix (NSCM) Task and the Nonsense Syllable Imitation Task (NSIT), respectively.

# Method

*Participants*

KEB and KM individuals who participate in the previous tasks were asked to perform the NRT (for further detail about participants, please see the "Participants" section in Chapter II) at the end of experimental sessions for the NSCM task. Data collected from the very first subject was excluded from the analysis for the BAL group because of the poor quality of recording, and data from the other participant in the BAL group were lost due to a computer problem. Therefore, participants for the NRT were 7 KEB children, 7 KEB adults, and 18 KM children.

*Stimuli*

Stimuli for the NRT are nonsense words that increase in length from 1 to 4 syllables. A set of 16 stimuli conforms to English phonotactic constraints (e.g., /ʧɪnɔɪtaʊb/) and none of the 16 nonwords corresponds to any real English word. Stimuli for the English NRT in the present study were taken from Dollaghan and Campbell's (1998) study, but re-recorded to improve the sound quality. The 16 stimuli for the English NRT, presented in Table 27, were recorded by a native English female speaker and each syllable was produced with equal stress.

*Task Procedure*

The NRT was given at the end of experimental sessions for the NSCM task. Each of the 16 stimuli was presented to the participant once over headphones, proceeding from 1- to 4-syllable nonwords. The participant was asked to repeat after each stimulus. All spoken responses were recorded by the same headset used in previous tasks and digitized on the same laptop computer for later phonetic transcription and evaluation.

Performance of each child was converted to numerical values, by computing the percent of phonemes produced correctly (i.e., Percentage of Phonemes Correct; PPC) for each nonword length (i.e., 1- to 4- syllables and for the task overall; Shriberg, Austin, Lewis, McSweeny & Wilson, 1997). Substituted and deleted phonemes were counted as incorrect productions, whereas distortions (as marked by diacritics) were counted as correct productions. Counting distortions as correct follows the original scoring procedure devised by Shriberg and Kwiatkowski (1982) for their Percent of Consonants Correct metric.

## Results

PPC for each nonword length was normalized via arcsine transformation and subjected to three separate repeated measures ANOVAs, with length as a within-group factor at 5 levels (i.e., 1- to 4-syllable lengths and total length). A linguistic proficiency analysis was not conducted because there was only one remaining participant in the BAL group. Using percent correct values from the experimental tasks (i.e., the NSCM task, the NSIT, and the NRT), Partial Pearson Product-Moment Correlations were computed to see if performance on the NRT relates significantly to performance on the NSCM task or the NSIT. These analyses may lead to better understanding of the relations among speech perception, speech production, and phonological awareness.

### *Comparisons of the Three Age-Language Status Groups*

Means for PPC for each nonword length and for total length (i.e., TOT-PPC) are presented in Table 28. Overall, the KEB children ($M = 86$, $SD = 4$) and the KEB adults ($M = 89$, $SD = 4$) performed better for each syllable length than the KM children ($M = 74$, $SD = 7$). The KEB adults appeared to perform the best of all three groups for each syllable length and for total length, except for 3-syllable nonwords. The KM children performed the worst of all three groups for each syllable length and for total length. For the 1-syllabe nonwords, the KEB children (*M* = 83, *SD* = 10) and the KEB adults (*M* = 83, *SD* = 5) performed similarly but better than the KM children (*M* = 70, *SD* = 8). The same appeared to be true for the 2-syllable nonwords: the KEB children (*M* = 88, *SD* = 5) and KEB adults (*M* = 93, *SD* = 5) performed better than the KM children (*M* = 78, *SD* = 9). PPCs for the 3-syllable length were similar between the KEB children (*M* = 91, *SD* = 5) and the KEB adults (*M* = 89, *SD* = 3) but higher than the KM children (*M* = 80, *SD* = 9). For the 4-syllable length, the KEB children (*M* = 81, *SD* = 6) performed less well than the KEB adults (*M* = 88, *SD* = 8), and the KM children performed worst (*M* = 67, *SD* = 12).

An arcsine transformation was performed on the percent correct data (i.e., PPCs). The first repeated measures ANOVA was conducted with length (L) as a within-group factor (5 levels: 1- to 4- syllable length and total length) among the three age-language status groups. The statistical results are presented in Table 29. The ANOVA revealed significant main effects for group, $F_{2, 31} = 25.082$, $p = .000$, with a large effect size $\eta_p^2 = .634$, $P = 1.000$, and length, $F_{4, 31} = 10.482$, $p = .000$, with a small effect size $\eta_p^2 = .228$, $P = 1.000$. A Scheffe's post hoc test for between-group comparisons showed that the KEB children (*M* = 86, *SD* = 4) and the KEB adults (*M* = 89, *SD* = 4) performed better than the KM children (*M* = 74, *SD* = 7), $p = .000$ and $p = .000$, respectively. Within-group comparisons were tested for significance using a Sidak's posthoc test. The Sidak procedure is a follow-up test that corrects for multiple comparisons. The Sidak test showed that the 1-syllable length (*M* = 75, *SD* = 9) was performed worse than the 2- (*M* = 84, *SD* = 10) and 3-syllable (*M* = 84, *SD* = 9) lengths, $p = .007$ and $p = .004$, respectively, and that the 2- and 3-syllable lengths were performed better than the 4-syllable length (*M* = 75, *SD* = 14), $p = .042$ and $p = .001$, respectively. In the repeated measures ANOVA, the first order interaction of length x group was not significant, $F_{2, 31} = 1.617$, $p = .127$, $\eta_p^2 = .100$, $P = .075$.

*Correlations among Perception, Production, and Phonological Awareness*

Pearson Product-Moment correlation coefficients and partial correlations were computed to see if phonological awareness, as measured by the NRT, was associated with perceptual performance on the NSCM task and production performance on the NSIT. Overall performance on the NRT (i.e., TOT-PPC) was compared to perceptual and production performance for familiar and unfamiliar consonants and vowels, and all sounds. Coefficients for the linguistic proficiency subgroups were not computed due to the insufficient number of participants for each proficiency subgroup.

The correlation coefficients for performance on the NSCM task and the NRT are displayed in Table 32. A correlation was conducted with or without group effect at once. Correlations between perceptual performance and phonological awareness seem to be affected by group only in two instances. When participants were grouped as the KEB child, KEB adult, and KM child groups (i.e., group controlled), perceptual performances for unfamiliar vowels and unfamiliar sounds were significantly correlated with overall performance on the NRT, $r = .47$, $p = .013$ and $r = .39$, $p = .046$, respectively.

When groups were not observed and instead, all participants were combined (i.e., group uncontrolled), overall perceptual performance on the NSCM task was positively correlated with overall performance on the NRT, $r = .53$, $p = .004$. In other words, an individual who performed well on the NSCM task, regardless of his or her age or language status, performed well on the

NRT. No correlation of perceptual performance for familiar consonants, familiar vowels, or familiar sounds with overall performance on the NRT was significant. Nonetheless, perceptual performance for unfamiliar sounds was significantly correlated with phonological awareness, as measured by the NRT. TOT-PPC on the NRT was correlated with unfamiliar consonants, $r = .57$, $p = .002$, unfamiliar vowels, $r = .54$, $p = .003$, and unfamiliar sounds, $r = .63$, $p = .000$. Furthermore, perceptual performance for consonants and vowels was significantly correlated with overall performance on the NRT, $r = .51$, $p = .006$ and $r = .54$, $p = .003$, respectively.

Partial Pearson Product-Moment correlation coefficients were also computed between production performance on the NSIT and phonological awareness, as measured by the NRT. Those coefficients are presented in Table 33. Correlations between phonological awareness and production performance on the NSIT were not significant for any sound type, familiarity level, or group; however, correlations for unfamiliar consonants and unfamiliar sounds were marginally significant, $r = .35$, $p = .051$, and $r = .34$, $p = .061$, respectively.

## Summary and Conclusions

The research question asked if L2 learners with greater phonological awareness perceive and produce English nonsense syllables correctly at higher rates. Two correlation analyses were conducted: one between perceptual performance on the NSCM task and phonological awareness on the NRT (i.e., with the overall TOT-PPC score) and the other between production performance on the NSIT and phonological awareness on the NRT. Partial correlations when group was controlled showed that performance on the NSCM task is significantly correlated with overall performance on the NRT for unfamiliar vowels, $r = .47$, and unfamiliar sounds, $r = .39$. In other words, the group that identified unfamiliar sounds well, particularly unfamiliar vowels, also performed the NRT with high PPCs. When group was uncontrolled, more significant correlations were found between NSIT scores and TOT-PPC: unfamiliar consonants, $r = .57$, unfamiliar vowels, $r = .54$, unfamiliar sounds, $r = .63$, consonants , $r = .51$, vowels, $r = .54$, and overall performance, $r = .53$. In general, L2 learners who performed well on the NRT also performed well on the NSCM, as predicted by the hypothesis.

Significant correlations were not found between production performance on the NSIT and overall performance on the NRT. The rate of correct articulation for unfamiliar consonants and unfamiliar sounds on the NSIT, however, showed marginally significant correlation with overall performance on the NRT. In general, these results are not consistent with my prediction that L2 learners who perform well on the NRT would perform well on the NSIT. The lack of significant correlations may come from the fact that every participant produced nonsense syllables well on the NSIT. With regard to correlations, my general impression is that overall performance on the NRT is related more closely to perceptual performance, as measured by nonsense syllable identification, than production performance, as measured by nonsense syllable imitation. Consequently, individuals with high phonological awareness in English, as measured by the NRT, seem to perceive unfamiliar English sounds well.

# CHAPTER VI

# DISCUSSION

The purpose of the present study was to examine (a) perception and production of English sounds as a second language (L2) or as a foreign language (FL) in children and adults who speak Korean as their first language (L1); (b) whether Korean-English bilingual (KEB) children, KEB adults, and Korean monolingual (KM) children perceive and produce English nonsense syllables in a different fashion from one another (i.e., how L2 experience influences perceptual and production performance); (c) whether perceptual and production performance are correlated; and (d) how best to account for perception and production, given the present findings. First, I will discuss L2 experience and L1 influence on performance of the tasks in relation to participants' proficiency, as measured by the English language sample analysis (LSA). Second, I will discuss how two models, which have been widely examined in the literature, failed to predict perceptual and production performance of L2 or FL English sounds in the present study, and what is suggested by results of the present study. Lastly, I will argue that separation of L2 learning mechanisms for consonants and vowels best explains and predicts perception and production of English nonsense syllables from an L2 or FL in children and adults.

## Influence of L2 Experience and L2 Proficiency

Numerous findings from cross-linguistic speech studies have shown that perceptual and production performance of nonnative speech sounds are closely related to the age of the participant when he or she first began learning the L2 (or age of arrival, AOA) and the amount of exposure to L2 (often determined by the length of residence, LOR, in a new country; Baker et al, 2008; Best & McRoberts, 2003; Eimas et al, 1971; Johnson & Newport, 1987; Mayberry &

Eichen, 1991; Ploog, 1984; Yeni-Komshian et al. 2000). Nonetheless, considering that the KEB children in the present study are school-aged and that school-aged children may still be developing adult-like perception by age 13 (Elliott *et al*., 1986; Morrongiello *et al*., 1984; Nittrouer, 1992; Sussman, 1993), different processing of speech perception in children might negatively affect perceptual performance and result in poorer performance on the nonsense syllable confusion matrix (NSCM) task for KEB children than for adults.

Unlike the preceding prediction, the KEB children, who were 13 years old or younger, perceived correctly more often on the NSCM task than the KEB adults, as expected from the AOA difference (despite comparable LOR), and also better than the KM children, as expected from the LOR difference (despite comparable current age). Moreover, the ED subgroup perceived English nonsense syllables correctly more often than the KD subgroup (with a later AOA and shorter LOR) and the KEB adults perceived correctly more often than the KM children (with less L2 experience, i.e., no LOR). Therefore, L2 learners with a younger AOA or a longer LOR appear to perceive English nonsense syllables better than L2 learners with an older AOA or a shorter LOR.

The finding that the KEB adults perceived L2 sounds correctly more often than the younger KM children can be interpreted in either of two ways (or perhaps both): (a) the KM children are old enough that they have lost the perceptual flexibility to perform better than adults or (b) the perceptual task in the present study is not designed to detect disadvantages of children's perceptual processing over that of adults'. Either way, it does appear that L2 experience (as represented by AOA or LOR) is more influential than one's current age or level of perceptual development for school-aged children to identify L2 or FL nonsense syllables.

On the nonsense syllable imitation task (NSIT) the KEB children produced L2 English sounds correctly more often than the KEB adults and the KM children, as expected. No group differences, however, were found across the three proficiency subgroups, where the 10 consonants were produced correctly more often than the 3 vowels, but familiar sounds were produced similarly to unfamiliar ones. There are two possible interpretations of the absence of a group difference or group interactions: (a) the task—repeating only a simple and open syllable (i.e., a consonant and vowel structure)—itself is not challenging enough to discern the ability of producing spontaneous speech in English, (b) the English proficiency of the participants in the three proficiency subgroups was not different enough to be reflected in performance in a spoken imitation task, or (c) parental or self-reporting is not always reliable (cf. Goldstein & Washington, 2001; Goldstein, et al 2005). Therefore, a spontaneous speech sample was collected from an oral interview, and a language sample analysis (LSA) was conducted using Systematic Analysis of Language Transcripts (SALT) in order to examine English proficiency of the participants in more detail.

In summary, the comparisons of the three age-language status groups both on the NSCM task and the NSIT showed that the KEB children perceived and produced correctly more often than the KEB adults and the KM children. In the LSA, the KM children performed significantly poorer than the KEB children and the KEB adults on the following measures: the length of Complete and Intelligible (C & I) utterances, number of different and total words (NDW; NTW), mean length of utterance (MLU), and words per minute (WPM). This was consistent with the prediction that the KM children would have less proficiency; however, measures of the LSA were not significantly different between the KEB children and adults, and among the three proficiency subgroups.

Even though most of the LSA measures showed no group difference among the two age groups and the three proficiency subgroups, to my ears, the English proficiency of the KEB adults seemed to be lower than that of the KEB children; and the proficiency of the KD subgroup seemed to be lower than that of the ED and BAL subgroups. Perceived English proficiency seemed to be greatly influenced by articulatory accuracy in words and sentences, as measured by the GFTA-2 and speech rates in spontaneous speech during the oral interview, as measured by WPM. Indeed, significant differences were found for these group comparisons on the articulation test that was given: GFTA-2. Besides, when comparing these measures with the normative means (for English monolingual children), automatically provided by SALT, speech rates of the KEB adult and KD groups were significantly slower: more than 1 or 2 standard deviations below the mean of the respective monolingual profile group. Statistical tests for correlations of the LSA measures with overall performance on the NSCM task and the NSIT were not conducted in the present study; however, a future study may examine these relationships.

## Speech Learning Model and Perceptual Assimilation Model

The Speech Learning Model (SLM; Flege, 1987, 1992; Flege, Takagi, & Mann, 1995; Flege, 2002) predicts that unfamiliar sounds should be performed better than familiar ones, whereas the Perceptual Assimilation Model (PAM; Best, McRoberts, & Sithole, 1998; Best, 2001) predicts familiar sounds should be perceived and produced correctly more often than unfamiliar ones. Both SLM and PAM assume that (a) perception and production of nonnative *consonants* work in the same way as those of nonnative *vowels*, (b) *perception* of nonnative sounds work in the same way as *production* of nonnative sounds, and (c) no interaction between sound type (i.e., consonants vs. vowels) and familiarity (i.e., familiar vs. unfamiliar sounds) would be expected. Nevertheless, findings from the NSCM task and the NSIT indicated that the assumptions of the two theories are not sufficient to predict L2 performance.

On the NSCM task, consonants were perceived differently from vowels, where familiar consonants (FC) were identified significantly better than unfamiliar ones (UC), but familiar and unfamiliar vowels (FV and UV, respectively) were identified similarly by all the groups of participants. Also, production performance on the NSIT for consonants showed different patterns than for vowels. The KEB and KM children produced FC correctly more often than UC, but the KEB adults produced FC correctly no better than UC. Interestingly, the KEB children produced UV correctly more often than FV (a prediction made for experienced listeners by SLM), in contrast with the KEB adults and the KM children who produced FV correctly more often than UV. Correlations between perceptual performance on the NSCM task and production performance on the NSIT were limited to unfamiliar sounds, including UC and UV, yet such correlations were weak.

Findings from the three experimental perception and production tasks (viz., the NSCM task, NSIT, and Nonword Repetition Task or NRT) suggest that (a) consonants are perceived and produced in a different fashion from vowels, (b) familiarity interacts with sound type, and (c) perceptual performance does not consistently predict production performance for L2 sounds. One major finding for both perception and production was that familiar sounds were generally perceived and produced correctly more often than unfamiliar ones, in keeping with the predictions of PAM. Nonetheless, neither SLM nor PAM can provide consistent and coherent predictions or explanations for all the findings of the present study, because both theories assume that perceptual and production performance for consonants will be the same as for vowels. Further, these two theories do not differentiate between perceptual performance and production performance. Given the lack of power of these two prevailing theories, the present study strongly points to the need for a new, more comprehensive theory of bilingual speech perception and production. Consequently, I will discuss findings from the present study to confirm the validity of two arguments: one is the separation between perceptual learning and production learning, and the other is the separation between consonant learning and vowel learning.

## Familiarity Interplay for Perception and Production of L2 Consonants and Vowels

### *Relations between Perceptual and Production Performance*

Two different correlations were reported in the present study: group controlled and group uncontrolled. Overall, a group controlled correlation—which measured the association between two variables (i.e., perceptual and production performance) with a control for other variables (i.e., a group effect)—revealed less significant coefficients, compared to the group uncontrolled ones. This suggests that group effects, such as age, language status, and proficiency, do not strengthen the correlations between perceptual and production performance. Some significant coefficients in the group uncontrolled correlations show that an individual, not as a group member, who performed better on the NSCM task generally performed better on the NSIT and the NRT.

Some investigators believe that successful acquisition of L2 phonemes requires accurate perception of phonemic contrasts (Brown, 2002); however, findings in the present study suggest that certain L2 phonemes that are correctly perceived may not be correctly produced, and vice versa. In general, an individual who scored higher on the NSCM task also scored higher on the NSIT; however, correlations were statistically significant only for some sets of sounds and even when significant, weak at best. Positive correlations were found for consonants, unfamiliar consonants, and unfamiliar sounds, but a negative correlation was found for unfamiliar vowels. In other words, an individual who identified unfamiliar consonants well tended to produce them well; however, an individual who identified unfamiliar vowels well tended to produce them poorly, or vice versa. A negative correlation between the NSCM task and the NSIT for UV seems to be attributable to the KEB children who perceived FV correctly at a similar rate to UV, but produced FV correctly *less* often than UV. These findings corroborate the idea that consonants act differently from vowels, as previously argued, and indicate that the correlation between perceptual and production performance is weak and positive for some sounds, yet negative for others.

Perceptual performance on the NSCM task showed moderate positive correlations with phonological awareness, as measured by the NRT. Interestingly, no correlation was found between production performance on the NSIT and the overall performance on the NRT. It is hard to believe that phonological awareness is unrelated to the articulatory ability. The NSIT is to repeat open-syllable, mono-syllabic nonwords (i.e., single syllables with a CV structure), whereas, the NRT is to repeat closed-syllable, mono- or multi-syllabic nonwords (i.e., one- to four-syllable sequences with a CVC syllabic structure). Therefore, the lack of correlation between the two production tasks may stem from the fact that the NSIT has smaller latencies between the stimulus and the response than the NRT (because the listener must process multiple syllables in the NRT before responding). As a result, the immediacy of imitation in the NSIT may contribute to the relatively high performance across the groups and to its lack discernment of the true level of articulatory ability of the various groups who participated in the present study (Porter & Castellanos, 1980).

These observations give a general impression that perceptual performance does not consistently predict production performance, where a correlation can be positive for some L2 sounds but negative for others. Furthermore, phonological awareness, as measured by the NRT, appears to be more closely related to the perceptual part of L2 learning rather than to the mere ability to immediately imitate L2 syllables. Even though individuals who perceived L2 sounds at a high rate also produced those at a high rate, patterns of perceptual performance are not the same as those of production performance. When reflecting my own experience of living in a community of foreign students in the U.S., the differences between L2 listening ability and L2 articulatory ability becomes apparent. I sometimes encounter L2 learners whose listening skill is as good as a native English listener's, but whose production skill is much less than a native English speaker's. I speculate that production skill stops developing at some point and perhaps becomes fossilized, while perceptual skill continues to develop along with one's linguistic experience. This leads to a prediction that one's L2 experience or L2 proficiency would influence perceptual learning more than production learning.

Nevertheless, we would not want to conclude that perceptual experience does not contribute to production learning. Onishi, Chambers, and Fisher (2002) showed that adult L2 speakers of English can acquire some phonotactic constraints which do not exist in their native language even with only a brief auditory experience. Indeed, findings of the present study showed that the NSCM task had significant, moderate-correlations with phonological awareness on the NRT. Not only linguistic experience but also non-linguistic experience such as musical training also may have an effect on language development. Tallal and Gaab (2006) argued that general auditory processing deficits in children with language-learning impairments are closely related to phonological deficits in these children. Many studies of auditory experience reveal that (a) there is a positive relation between auditory experience and language development, and (b) auditory experience keeps developing or at least accumulating throughout the life span (Kraus, Skoe, Parbery-Clark, & Ashley, 2009; Lynch, Eilers, Oller, & Urbano, 1990; Saffran & Griepentrog, 2001; Tallal & Gaab, 2006). According to Musacchia, Sams, Skoe, and Kraus (2007), musicians who had more experience with musical stimuli had greater brainstem responses to speech and music stimuli and were better at encoding fundamental frequency than non-musicians. Further, these musicians showed some training effects on pitch coding, which may be vital to understanding pragmatic meanings of speech. These findings lead to the point that linguistic or even non-linguistic auditory experience may enhance plasticity for perceiving nonnative speech sounds. Therefore, the KEB adults may be able to develop perceptual learning of L2 sounds concomitantly with increasing L2 experience.

### *High-scoring vs. Low-scoring Sounds*

The present study addressed the possibility of having some universal properties in speech sounds: some English consonants are easy or difficult not only for nonnative listeners but also for native listeners of that language to correctly identify. In a study of perceptual confusion in English monolinguals done by Phatak and Allen (2007), consonants were categorized into one of the three groups: consonants with low-scores, intermediate-scores, or high-scores. I propose that some sounds are easy to identify even for nonnative listeners, but other sounds are hard, even for native listeners. Previous studies of speech perception in monolinguals have suggested that some English sounds are more confusing than others. For example, /θ/ and /f/ often have been observed to be confused with each other and to cause high error rates for identification, even by English monolingual adults. In general, voiced consonants are more likely to be confused than voiceless ones (Abdelatty Ali et al., 2001); fricatives are more likely to be confused than stops (Miller & Nicely, 1955; Phatak & Allen, 2007); obstruent consonants (i.e., stops, fricatives, and affricates) are more likely to be confused than sonorant consonants (i.e., nasals, glides, and liquids; Miller & Nicely, 1955; Phatak & Allen, 2007); and anterior sounds (made near the front of the mouth) are more likely to be confused than posterior ones (Benki, 2001; Nishi et al, 2010). If some sounds are more confusing to native ears, then those also may be just as confusing to L2 or FL learners.

In contrast, with respect to perception, some English consonants and vowels were perceived at similar rates (i.e., percent correct) across the groups in the present study, regardless of sound type or familiarity. Across the groups, /p, g/ were the two consonants that were identified the best; /k, ʧ/ were generally identified well; /b, z, ʤ/ were generally identified poorly; and /f, θ/ were the two consonants that were identified the worst. These findings are consistent with those of previous studies in the literature (Abdelatty Ali et al., 2001; Benki, 2001; Miller & Nicely, 1955; Nishi et al, 2010; Phatak & Allen, 2007), in that fricatives /f, θ, z, ʤ/ were generally identified less correctly than stops /p, k, g/; anterior fricatives /f, θ/ were generally identified more poorly than more posterior fricatives and affricates /s, z, ʤ /; and voiced consonants /b, z, ʤ/ were generally identified more poorly than voiceless ones /p, s, ʧ/.

Therefore, general tendencies found for English monolingual listeners in the literature were also observed in bilinguals and Korean monolinguals in the present study.

With respect to production, as previously mentioned, stops such as /p, b, k, g/ are generally acquired earlier than other sounds; then some fricatives such as /s, z, f/ are acquired by age 3;5; the familiar consonant /ʧ/ is typically acquired by age 4;5; but the unfamiliar consonant /θ/ is typically acquired by age 6;11, much later than the other consonants (Dodd, Holm, Hua, & Crosbie, 2003; Smit et al., 1990; Templin, 1957; Wellman et al., 1931). (The consonant /θ/ would not seem to require a difficult articulatory gesture, in that the tongue is merely placed between the teeth. The poor audibility of the resulting frication may, however, be intrinsically difficult for a talker to monitor.) In support of the concept of intrinsic production difficulty, in the present study, unfamiliar consonants /b, g, z/ were produced correctly at a high rate, but the familiar consonant /ʧ/ was produced correctly at a relatively lower rate across the groups. Therefore, even though the child groups generally produced familiar consonants better than unfamiliar ones, sometimes this L1 influence was countered by the intrinsic nature of certain sounds.

Based on what has been studied in monolinguals and the findings of the present study, there seem to be some sounds which are easy or difficult to perceive and produce regardless of one's age, language status, or linguistic proficiency or one's native language. High-scoring sounds such as /p, k, g/ may have some universal perceptual or production features that are so prominent, clear, and easy to capture in the speech stream that any listener can identify or imitate them correctly at a high rate, regardless of his or her L2 proficiency and possibly L1 consonant inventory. On the other hand, the English consonants /f, θ/ are perceptually confusing and hard to identify, not only for L2 listeners, but also for L1 listeners. For convenience of reference, I will use the term "salient" to describe high-scoring sounds that are perceived and produced correctly at a high rate across the groups in contrast with low-scoring sounds. In production, /p, b, k, g/ are salient and thus produced correctly at relatively high rates, whereas /ʧ, ʤ, θ/ are not salient and thus produced correctly at relatively low rates. Therefore, the construct of "saliency" (or high-scoring and low-scoring sounds) should be incorporated into a theory that accounts for an interaction of sound type and familiarity, and for L2 perception separately from L2 production.

*Sound Type: Consonants vs. Vowels in Perception*

As mentioned earlier, performance for familiar sounds was different from performance for unfamiliar sounds, depending on sound type (consonants or vowels). Also, perceptual patterns differed from production patterns. Therefore, a theory which accounts for an interaction of sound type and familiarity and for L2 perception separately from L2 production should be developed. A hypothetical model or heuristic that predicts perception of nonnative consonants and vowels is needed to explain the current findings for all age, language status, and proficiency groups: including the findings that (a) consonants were perceived correctly more often than vowels, (b) *familiar sounds* were perceived correctly more often than unfamiliar ones, (c) *familiar consonants* were perceived correctly more often than unfamiliar ones, whereas *familiar and unfamiliar vowels* were perceived equally well (i.e., there was an interaction of sound type and familiarity), and (d) such perceptual patterns as in (a) to (c) were similar across the groups (i.e., there were no group interactions for perception). Table 32 is a presentation of the observed findings from the NSCM task (i.e., perception) and the NSIT (i.e., production) in the present study.

In the present study, consonants were perceived correctly more often than vowels across the groups. There can be two different explanations for this. First, listeners in the present study may have may have paid more attention to consonant distinctions than vowel distinctions when identifying a syllable because there were 10 consonants but only 3 vowels to discriminate from one another. Furthermore, confusions among these 3 vowels were likely compounded because they were close in perceptual space and could all be considered lax. Even so, a recent study by Carreira, Gillon-Dowens, Bergara, and Perea (2009) suggests that consonants themselves are more attention-grabbing more than vowels. These investigators showed that there was a greater delay in recognizing words presented visually when consonants, rather than vowels, were omitted from the spelling. They argued that the brain distinguishes consonants from vowels for visual-word recognition. In the practice of communication in our daily lives, consonants may carry the content of a message, whereas vowels may carry subsidiary information such as the gender or mood of the speaker, delivered by suprasegmental features. Therefore, consonants would seem to be more critical than vowels in perceiving words or messages carried by sequences of speech sounds. This may hold true as well for bilinguals or monolinguals perceiving L2 or FL sounds, respectively.

The second possible explanation is that consonants are easier for L2 or FL learners to correctly identify than vowels, because of acoustic differences. Early acoustic studies of speech sounds showed that consonants involve rapid spectral changes while vowels show relatively stable spectral patterns (Liberman, Delattre, Cooper, & Gerstman, 1954). Therefore, categorical perception can be more readily demonstrated for consonants than for vowels. Such a difference may have caused less confusion for consonants than vowels on the identification task in the present study. Given that vowels carry subsidiary information about gender, dialect, and melodic or voice tone, individual target vowels may vary phonetically more than individual target consonants, which may in turn distract a listener from identifying signature acoustic spectral patterns. This should not necessarily lead to the conclusion that perception takes advantage of consonant noises more than vowel resonances, however: Although some studies have reported that consonants were perceived correctly more often than vowels (Fu & Shannon, 1998; Johnson et al, 2007), others have reported the opposite pattern (Boothroyd, 1984; Dubno, Dirks, & Langhofer, 1982).

Whether better performance on consonants than vowels is attributed to the greater number of consonant stimuli, the attention-grabbing characteristics of consonants, or the advantage of discerning signature consonantal acoustic traits, these observations suggest that the perception of consonants is processed independently of and differently from the perception of vowels (or at least three English lax vowels), at some linguistic or physiological level of speech. Some recent studies argue that consonants and vowels may have an independent status in language processing (Caramazza et al., 2000; Kolinsky, Lidji, Peretz, Besson, & Morais, 2009; Poeppel, 2003; Zatorre et al, 2002).

Kolinsky, Lidji, Peretz, Besson, and Morais (2009) examined processing interactions between consonants and vowels sung on two-note intervals with lyrics of bisyllabic nonwords (e.g., one pair of /daty/ and /daky/ in F and G notes, respectively for each syllable; and the other pair in F and C notes). The investigators argued that the phonological dimension seems to be integrated more closely with the melodic dimension for vowels than consonants. Indeed, early literature showed that there is a right ear and left hemisphere advantage for linguistic elements, particularly for consonants, but a left ear and right hemisphere advantage for non-linguistic elements such as prosody carried on vowels (Kimura, 1961, 1964; Spellacy & Blumstein, 1970). Assuming that perception of consonants is independent of perception of vowels, the familiarity effect (i.e., L1 influence) may also affect perception of consonants in a different way from that of vowels. Next, I will discuss the familiarity effect (i.e., L1 influence) on sound type in perceptual performance of L2 or FL sounds.

*Sound Type by Familiarity in Perception*

Familiar consonants proved less confusing for L2 learners than unfamiliar ones, as predicted by PAM. Familiar consonants contrast with other L1 phonemes, on a number of acoustic dimensions or features. Assuming these features or contrasts are similar in L1 and L2, they may have helped L2 listeners identify familiar consonants better than unfamiliar ones. On the other hand, perception of vowels appeared to be less influenced by similarity or differences between L1 and L2 (i.e., familiarity) in the present study.

The influence of L1 Korean or perceptual performance on the NSCM task can be observed in more detail when looking at the patterns of sound confusions. A voicing confusion between /p, k/ and /b, g/ occurred less often than a voicing confusion between /s, ʧ/ and /z, ʤ/. This is probably because Korean has the voiced counterparts [b, g] as allophones of the voiceless stops /p, k/ (thus making /b, g/ partially familiar). This is not true for the voiceless fricatives /s, ʧ/. In another example of L1 influence, the KEB adults and the KM children showed more place and manner confusions than the KEB children, such as the bidirectional /z/ ↔ /ʤ/ (place and manner) and /f/ ↔ /b/ (manner) confusions (where both members of the pairs are unfamiliar consonants). This is probably because Korean lacks many of the place- and manner-of-articulation contrasts found among obstruents (i.e., stops, fricative, and affricates) in English. It is interesting that the /f/ ↔ /b/ confusion was more frequent than /f/ ↔ /p/ confusion. One may argue that absence of the voiced counterpart /v/ may have forced the listeners to shift manner and choose /b/ for the stimulus /f/. Nevertheless, it is still uncertain why /b/ but not /p/ was selected more often. Perhaps this occurred from another L1 influence, namely that the English sounds /f/ and /v/ in English loanwords in Korean are often spelled and produced as an aspirated Korean /$p^h$/ and a lenis /p/ (with its allophone [b]), respectively. In yet a third example of an L1 influence, the KEB adults and KM children showed more frequent occurrence of /ɪ/ → /ɛ/ and /ɛ/ → /æ/ confusions than the KEB children. Because all three vowels in the present study were lax and acoustically close, they may have been highly confusable. Therefore, I speculate that some of the patterns of sound confusion in the present study may be due to the influence of L1 Korean, which lacks some voiced phonemes or voiced allophonic variants and some place- or manner-of-articulation features among consonants, and a tense-lax distinction among vowels.

*Sound Type: Consonants vs. Vowels in Production*

Consonants were produced more accurately than vowels across the groups. With respect to physiological aspects of speech production, learning how to articulate L2 vowels may be more difficult than learning to produce L2 consonants. The tongue positions for producing different vowels would seem to be less discrete than those for consonants. In addition, place of articulation would seem to be less varied along the vocal tract for vowels than consonants. Further, only the tongue, jaw, and lip shape among all other articulators are prominent in vowel production, unlike consonants which require involvement of other articulators such as the teeth, palate, and lip contact. Consonants, therefore, would seem to offer more tactile or kinesthetic feedback than vowels, which would seem to rely more on proprioceptive feedback (for a discussion of feedback mechanisms, see Kent, 1998). Taken together, these considerations may make it difficult for L2 or FL learners to build internal models of vowels (cf. DeThorne, Johnson, Walder, & Mahurin-Smith, 2009) to guide motor production, when L2 learners try to speak English on their own or when language teachers try to give explicit and clear instructions about how to produce vowels accurately.

Another possible reason for better performance of consonants over vowels is that L2 learners may have focused on production of consonants more than vowels, based on the assumption that consonants are more crucial to delivering meanings in continuous speech. For example, one participant in the present study produced the English word "pin" as /pin/ not as /pɪn/, but another produced it as /bɪn/. In this situation, the chance of successful communication with a native English listener would seem to be higher for the former speaker than the latter. Owren and Cardillo (2006) examined the role of consonants and vowels in discerning a talker's identity and the meaning of words. In their study, participants identified word meaning more accurately with consonant-only stimuli, but the talker more accurately with vowel-only stimuli. The stimuli for the production task in the present study, however, were all nonsense syllables, which involve no encoding of meaning. A talker's identity was irrelevant, but phonetic information was important, perhaps leading my participants to concentrate on precisely articulating the consonants, as they might in meaningful speech.

In Caramazza et al. (2000), two Italian-speaking aphasic patients showed complementary performance in producing consonants and vowels. One patient made errors in producing vowels much more often than consonants, whereas the other did the opposite. The investigators also examined whether the sonority feature contributes to different error rates between consonants and vowels, but this was not the case: The second patient, who made errors mostly on consonants, produced liquids (i.e., /l, r/, which are sonorant) as poorly as obstruents (i.e., stops, fricatives, and affricates, which are not sonorant). The investigators concluded that consonants and vowels are autonomous dimensions of speech. Next, I will discuss interplay of sound type and familiarity in production.

### *Sound Type by Familiarity in Production*

On the NSIT, a participant group effect interacted with sound type by familiarity. Rates of correct articulation in the KEB and KM children were significantly higher for FC than for UC; however, articulatory accuracy of the KEB adults was not significantly different between FC and UC. Rates of correct articulation for vowels, however, tell a different story from production of consonants and also from perception of vowels. The KEB children (who were younger and more experienced L2 learners) produced the familiar vowel /ɛ/ *worse* than unfamiliar ones, and the KEB adults (who were older L2 learners) and the KM children (who were less experienced L2 learners) produced the familiar vowel *better* than unfamiliar ones. Experienced L2 learners appear to take advantage of *unfamiliarity* for vowels; however, less experienced L2 learners appear to do the opposite, namely, to take advantage of *familiarity* for vowels. This is another intriguing aspect of L1 influence on L2 learning. Perhaps some threshold of L2 production proficiency must be achieved before an L2 learner can attend to new phonetic aspects of vowel production. This appeared to be the case in Flege's studies of SLM (1987, 1992, 2002), where the advantage for unfamiliar vowels was seen in experienced L2 speakers.

Towards an Interplay Model of Processing of L2 Consonants and Vowels

Overall, findings of the present study suggest that nonnative consonants are perceived and produced in a different way from nonnative vowels, perceptual performance on the NSCM task does not always predict production performance on the NSIT or NRT, and L1 influence (e.g., familiarity) and L2 experience (e.g., group effect) affects perceptual and production performance to some extent, but not for all sounds. Neither SLM nor PAM consistently predicts or explains the present findings. I propose that sound type (i.e., consonants vs. vowels) interacts with familiarity (i.e., familiar vs. unfamiliar sounds), with some weighting provided by universal perceptual and production properties of individual sounds.

In general, L2 or FL learners take advantages of "familiarity" in perceiving and producing nonnative sounds. A brief listening experience or even non-linguistic training effect on speech perception indicate that auditory experience helps develop phonological intuitions and better ability to perceive and produce speech sounds. Therefore, perception of L1 Korean sounds that are similar to L2 or FL English sounds may have helped the participants in the present study identify and imitate familiar sounds correctly more often than unfamiliar sounds. Familiarity is by definition language-dependent; thus, degree of L2 experience (such as age of acquisition or level of proficiency) should play a role in learning new or different acoustic or articulatory features of many unfamiliar L2 sounds.

The present study showed that the degree of L2 experience does interact with familiarity and sound type. On the NSCM task, interaction of L2 experience was not observed because all groups perceived FC correctly more often than UC but FV as often as UV. On the NSIT, the KEB children who were the most experienced L2 learners produced UV correctly more often than FV, but the KEB adults and the KM children did the opposite. This pattern, however, was not observed for production of consonants. As mentioned previously, consonants involve shorter durations of spectral change than vowels. If perceptual or articulatory saliency is related to short spectral changes, then vowels might be expected to be less salient than consonants. Therefore, it seems that these two effects (i.e., familiarity and saliency) together influence the learning of nonnative speech sounds, and that advantages of familiarity and saliency are different for consonants and vowels. Nevertheless, from their high or low rate of identification, some sounds or sound groups appeared to be processed in a language-independent or universal way (either well or poorly perceived or produced).

Regardless of familiarity or L2 experience, some high-scoring sounds were perceived correctly at a high rate and they were /p, k, g/. The consonant /g/ is an unfamiliar sound, unlike the other two familiar sounds. Interestingly, this unfamiliar sound was produced correctly 95% of the time, the highest rate. High-scoring sounds were apparent among the unfamiliar sounds, perhaps because they are more "attention-grabbing" (Carreira et al., 2009) when they are new and unfamiliar. As an example, Flege (1987, 1992, 2002) argued that a new sound is much easier to learn than a sound similar to one in L1, because L2 learners can readily tell the differences between it and sounds already familiar from L1. Likewise, a new sound that is very prominent and salient among unfamiliar sounds will be more easily heard or produced than one already known from L1. In the production of consonants, as measured by GFTA-2, the KEB children and adults made no errors in producing the consonant /f/. In perception, however, /f/ was identified at the lowest rate in these groups. Therefore, a perception-based natural class seemed to exist, possibly with some dimensions of acoustic properties that hinder the perception of that sound, different from a production-based natural class. Future studies may explore which sounds constitute a perception-based natural class.

If one accepts that saliency interacts with familiarity and sound type, then the following findings from the present study and previous literature can be explained: (a) why the unfamiliar consonant /g/ was perceived well (viz., because it is salient); (b) why the unfamiliar consonants /b, g/ were produced more accurately than the familiar consonants /s/ and /ʧ/ (viz., because /b, g/ productions have salient acoustic results); and (c) why the unfamiliar consonants /f, θ/ were perceived poorly by bilinguals and monolinguals (viz., because these two sounds are not acoustically salient). Of course, this is a bit of circular reasoning: What is needed is some independently determined acoustic or articulatory explanation for what makes some sounds salient or high scoring, and others low-scoring. My only point is that some sounds actually seem to be processed better than others, due to their intrinsic nature rather than the speech sound inventory of the native language. Including familiarity of sounds in the interaction also appears to be a strong and viable way to predict which L2 consonants and vowels generally can be learned better (e.g., for perception, consonants familiar from the L1 inventory; or for production, vowels familiar from the L1 inventory, for less experienced participants).

In conclusion, perception and production of L2 sounds appear to entail an interaction of sound type, familiarity, and L2 experience, where perceptual learning is related to yet independent from production learning. The separation of consonants from vowels, the separation of perceptual learning from production learning of L2 or FL sounds, and the interaction of these effects with each other, and with familiarity and saliency, may predict (a) which L2 consonants and vowels generally can be learned better (i.e., consonants will be learned better than vowels;

and familiar sounds will be learned better than unfamiliar ones), (b) how L2 experience affects learning of L2 or FL consonants and vowels (i.e., experienced L2 learners will produce unfamiliar vowels better than familiar vowels; however, this will be the opposite in inexperienced L2 learners), (c) that some nonnative sounds can be learned without great effort regardless of resemblance to one's L1 inventory of phonemes (i.e., high-scoring sounds in either perception or production, which are highly likely to be early-acquired sounds), and (d) that auditory experience with some nonnative sounds may aid their production (i.e., perceptual learning may assist the production of unfamiliar sounds, as seen in the small correlations between the NSCM task and the NSIT).

## Limitations and Future Study

The present study used two prevailing models to predict perception and production of English nonsense syllables in Korean-English bilingual children and adults, and in Korean monolinguals. These models, SLM and PAM, fail to account in a comprehensive way for many of the findings of the present study. I am claiming that sound type interacts with familiarity and saliency, and that any model of bilingual speech-sound processing should address perception separately from production. Nevertheless, the concept of interaction that I propose among the three factors requires more elaboration. Also, despite the separation of perception and production, any L2 process model should explain the weak positive correlation between perception and production of unfamiliar consonants, and the weak negative one between perception and production of unfamiliar vowels. Future studies may include examination of consonants produced correctly in spontaneous speech (collected from the oral interviews).

In particular, the concept of saliency needs to be developed further, building on inferences from studies of acoustic and articulatory aspects of phonemes, and perceptual and production performance across the world's languages. To define such universal saliency, future studies could extend the present method to speakers of different native languages. Universal saliency may be closely related to the frequency of occurrence of sounds in the world's language. In this sense, the concept of universal saliency is similar to Lindblom's (1992, 2000) notion of "core" segments that occur in many of the world's language, as opposed to "exotic" segments that occur only in languages with large phonemic repertoires (Vihman, 2002). Extension of my current findings to other L1 speakers would provide conceptual rigor, because saliency should include aspects of the speech signal that all human beings can physically detect, independent of language differences (i.e., familiarity). Investigating many languages may contribute to defining perceptually based natural sound classes as well.

In the present study, only nine KEB children participated and there were only three children for each of the three linguistic proficiency subgroups. However, I believe the data from the NSCM task are reliable even with this small number of participants because many trials—ranging from 1500 trials to 2400 trials for the 30 syllables—were collected from each participant during 5 to 6 hourly experimental sessions. Nonetheless, it would be valuable to explore further the apparent advantages of consonants over vowels, and vowel differences related to familiarity and different degrees of L2 experience. Future research should include more KEB children for each of the ED, BAL, and KD subgroups, to determine whether performance on the NSIT and L2 proficiency (as evaluated by the LSA) might truly be different among the three proficiency subgroups. My informal observation during the oral interviews that the KEB adults pronounced

English words less accurately than the KEB children, and also that the KD children did so less accurately than the ED and BAL children, could be substantiated by future quantitative analysis of the Percent of Consonants Correct (PCC) and Percent of Vowels Correct (PVC) in the English language samples I collected (cf., Shriberg et al., 1997).

Even though the LSA showed no group difference among the three English proficiency subgroups, the results of the LSA provide valuable information about the acquisition of English in a bilingual population. Recent studies in speech-language pathology highlight the need for language assessment measures for bilingual speakers that are comparable to the ones for monolinguals. To develop such language assessment measures, a vast amount of descriptive, normative data on bilingual speakers should be collected. SALT includes an extensive database for Spanish-English speaking children, but databases for bilingual speakers of languages other than Spanish are still lacking. There has been an argument that English standardized tests are not sufficient to identify children with language impairments, even in English monolingual populations (Dollaghan & Campbell, 1998; Plante & Vance, 1994).

LSA potentially overcomes many of the limitations that standardized tests have and is a valid indicator of expressive language performance in children. Findings from the present study should contribute to the building up of a rich normative database for Korean-English speaking bilingual children. Further, determining whether familiar or unfamiliar consonants and vowels contribute to difficulty in learning English as an L2 should direct language educators (possibly including speech-language pathologists) to sounds that should be addressed in educational and clinical treatment programs with Korean-English bilinguals or Korean monolinguals just beginning to learn some English.

# TABLES

Table 1

*Korean Consonant Inventory*

| | Place | | | | | |
|---|---|---|---|---|---|---|
| Manner | Bilabial | Alveolar | Alveo-palatal | Palatal | Velar | Glottal |
| Stop | $p^h$, p, $p^*$<br>ㅍ, ㅂ, ㅃ | $t^h$, t, $t^*$<br>ㅌ, ㄷ, ㄸ | | | $k^h$, k, $k^*$<br>ㅋ, ㄱ, ㄲ | |
| Affricate | | | $ʧ^h$, ʧ, $ʧ^*$<br>ㅊ, ㅈ, ㅉ | | | |
| Fricative | | s, $s^*$<br>ㅅ, ㅆ | | | | h<br>ㅎ |
| Nasal | m<br>ㅁ | n<br>ㄴ | | | ŋ<br>ㅇ | |
| Approximant | w<br>ㅚ,ㅙ,ㅟ,ㅞ | l ~ r<br>ㄹ | | j<br>ㅑ,ㅕ ㅛ,ㅠ | | |

*Note*. The diacritic "$^h$" refers to aspiration; the diacritic "*" refers to fortis sounds; and the symbol "~" refers to alternation, where [l] may substitute for [r] in onset position and vice versa. The Korean alphabet symbol for each sound is given below the IPA symbol.

Table 2

*Triplets of Korean Obstruents*

| | Aspirated | Lenis | Fortis |
|---|---|---|---|
| Bilabial stops | /pʰaŋ/ *bang* 팡 | /paŋ/ *room* 방 | /p*aŋ/ *bread* 빵 |
| Alveolar stops | /tʰam/ *greed* 탐 | /tam/ *fence* 담 | /t*am/ *sweat* 땀 |
| Velar stops | /kʰɨm/ *big* 큼 | /kɨm/ *gold* 금 | /k*ɨm/ *extinguish* 끔 |
| Alveolar fricative | N/A | /sam/ *three* 삼 | /s*am/ *pack* 쌈 |
| Palatoalveolar affricate | /ʧʰaŋ/ *spear* 창 | /ʧaŋ/ *market* 장 | /ʧ*aŋ/ *boss* 짱 |

*Note*. There are three different types of Korean stops and affricates, but only two different types of Korean alveolar fricatives. The Korean alphabet symbols are given to the right of each word.

Table 3

*Familiar and Unfamiliar English Consonants and Vowels Compared to Korean*

| | Familiar | Unfamiliar |
|---|---|---|
| Consonants | p, t, k, s, h, ʧ, m, n, ŋ, r, l, w, j | b, d, g, f, v, θ, ð, ʃ, ʤ |
| Vowels | i, ɛ, a, o, u | ɪ, e, æ, ɔ, ʊ |

Table 4

*Percentages of Classification Types and Goodness-of-fit Ratings for a Given English Stimulus to an Italian Vowel*

| English Stimuli | Italian Vowels | | | | | |
|---|---|---|---|---|---|---|
| | **i** | **e** | **ɛ** | **a** | **ɔ** | **o** |
| ɪ | 65 (2.9) | 35 (4.0) | - | - | - | - |
| ɛ | - | 47 (3.6) | 53 (3.8) | - | - | - |
| æ | - | 10 (3.2) | 75 (3.8) | 15 (2.6) | - | - |
| ɒ | - | - | - | 47 (3.4) | 33 (4.1) | 20 (3.7) |
| ʌ | - | - | - | 93 (3.7) | - | - |
| ɚ | 18 (1.8) | 63 (1.6) | 15 (1.6) | - | - | - |

*Note*. Adapted from “Perceiving vowels in a second language” by Flege and MacKay, 2004, *Studies in Second Language Acquisition,* 26, p. 12. Percentages of classification smaller than 2% are not presented. Goodness-of-fit ratings are given in parentheses, ranging from 1 (poor fit) to 5 (good fit).

Table 5

*Characteristics of Korean-English Bilingual (KEB) Children, KEB Adults, and Korean Monolingual (KM) Children*

| | Korean-English Bilingual Children[a] | Korean-English Bilingual Adults[b] | Korean Monolingual Children[c] |
|---|---|---|---|
| Age | 10;3 (2;1) | 22;8 (2;1) | 11;0 (0;6) |
| Age of Arrival | 5;8 (3;7) | 18;0 (2;6) | N/A |
| Length of Residence | 4;7 (1;7) | 4;9 (0;7) | N/A |
| Perceived Language Proficiency | N/A | Korean > English | Korean > English |
| Language spoken at home | Korean = English | Korean | Korean |
| Language spoken at school or work | English > Korean | English > Korean | Korean |
| GFTA-2 | 3.1 (2.2) | 6.3 (1.6) | 13.9 (5.5) |
| EVT | 80 (10) | 96 (12) | 60 (7) |

*Note.* The means and standard deviation of ages, ages of arrival (AOA), and lengths of residence (LOR) in the U.S. are given in "years; months." GFTA-2 refers to the Goldman Fristoe Test of Articulation-2, and the given value is the number of sound errors on average for the group. EVT refers to the Expressive Vocabulary Test, and the given value is the average standard score. KEB and KM children were 8 to 13 years old; KEB adults were 21 to 33 years old. $N = 34$; [a]n = 9; [b]n = 7; [c]n = 18.

Table 6

*Characteristics of the Three Proficiency Subgroups in Korean-English Bilingual (KEB) Children*

| | Korean-English Bilingual Children | | |
|---|---|---|---|
| | English-dominant (ED)[a] | Balanced (BAL)[b] | Korean-dominant (KD)[c] |
| Age | 8;4 (0;5) | 9;8 (1;6) | 12;7 (0;7) |
| Age of Arrival | 3;0 (1;0) | 4;0 (2;0) | 10;0 (1;0) |
| Length of Residence | 5;6 (0;3) | 5;9 (0;2) | 2;7 (0;10) |
| Perceived Language Proficiency | English > Korean | English ≥ Korean | Korean > English |
| Language spoken at home | English > Korean | Korean = English | Korean > English |
| Language spoken at school or work | English > Korean | English > Korean | English > Korean |
| GFTA-2 | 2<br>(2, 2, 2) | 2.3<br>(1, 2, 4) | 5<br>(2, 5, 8) |
| EVT | 86<br>(78, 88, 92) | 86<br>(77, 88, 92) | 69<br>(66, 68, 72) |

*Note*. Ages, ages of arrival (AOA), and lengths of residence (LOR) are given in "years; months." GFTA-2 refers to the Goldman Fristoe Test of Articulation-2, and EVT refers to the Expressive Vocabulary Test. KEB children were 8 to 13 years old. $N = 9$; [a]n = 3; [b]n = 3; [c]n = 3.

Table 7

*English Language Sample Analyses for the Three Age-Language Status Groups*

| | Korean-English Bilingual Children[a] | Korean-English Bilingual Adults[b] | Korean Monolingual Children[c] |
|---|---|---|---|
| Complete & Intelligible Utterances | 63 (11) | 61 (11) | 17 (6) |
| Different Words in 50 Utterances | 117 (12) | 137 (37) | 17 (8) |
| Total Words in 50 Utterances | 274 (55) | 334 (99) | 30 (13) |
| Grammatical Errors in C & I utterances | 5.8 (3.8) | 8.8 (4.2) | 0.6 (0.7) |
| Mean Length of Utterance | 5.3 (1.5) | 6.4 (1.3) | 1.5 (0.3) |
| Standard Deviation from the normative means | 0.02 | 0.72 | -2.34†† |
| Words per Minute | 75 (27) | 72 (23) | 18 (4) |
| Standard Deviation from the normative means | -0.74 | -1.25† | -2.16†† |
| Mazes | 27 (22) | 18 (6) | 3 (2) |
| Standard Deviation from the normative means | 1.09† | 0.24 | -0.36 |
| Abandoned Utterances | 7.1 (4.6) | 8.8 (6.5) | 0.1 (0.3) |
| Standard Deviation from the normative means | 2.14†† | 1.72† | -0.73 |

*Note.* Analyses were done using *Systematic Analyses of Language Transcripts* (SALT). Korean-English bilingual (KEB) and Korean monolingual children were 8 to 13 years old; KEB adults were 21 to 33 years old. The symbol " † " means 1 SD and " †† " means 2 SDs from the mean. *N* = 32; [a]n = 9; [b]n = 7; [c]n = 16.

Table 8

*Multivariate Analysis of Variance for English Language Sample Measures in the Three Age-Language Status Groups*

| | *df* | *F* | *p* |
|---|---|---|---|
| Complete & Intelligible Utterances | 2 | 110.971** | .000 |
| Different Words in 50 Utterances | 2 | 143.604** | .000 |
| Total Words in 50 Utterances | 2 | 104.445** | .000 |
| Grammatical Errors in C & I utterances | 2 | 18.222** | .000 |
| Mean Length of Utterance | 2 | 69.187** | .000 |
| Standard Deviation from the normative means | 2 | 37.628** | .000 |
| Words per Minute | 2 | 38.840** | .000 |
| Standard Deviation from the normative means | 2 | 16.174** | .000 |
| Mazes | 2 | 12.958** | .000 |
| Standard Deviation from the normative means | 2 | 2.114 | .140 |
| Abandoned Utterances | 2 | 17.258** | .000 |
| Standard Deviation from the normative means | 2 | 12.488** | .000 |

*Note.* Analyses were done using *Systematic Analyses of Language Transcripts* (SALT). Korean-English bilingual (KEB) and Korean monolingual children were 8 to 13 years old; KEB adults were 21 to 33 years old. *N* = 32; [a]n = 9; [b]n = 7; [c]n = 16.

Table 9

*English Language Sample Analyses for the Three Proficiency Subgroups of Korean-English Bilingual (KEB) Children*

| | English-dominant[a] | Balanced[b] | Korean-dominant[c] |
|---|---|---|---|
| Complete & Intelligible Utterances | 62 (14) | 67 (1) | 59 (14) |
| Different Words in 50 Utterances | 119 (12) | 124 (10) | 110 (13) |
| Total Words in 50 Utterances | 268 (40) | 302 (54) | 252 (76) |
| Grammatical Errors in C & I utterances | 1.7 (1.1) | 4.7 (5.5) | 9.7 (3.5) |
| Mean Length of Utterance | 5.3 (1.1) | 6.1 (1.2) | 5.4 (2.1) |
| Standard Deviation from the normative means | -0.2 | 0.4 | -0.03 |
| Words per Minute | 68 (24) | 88 (34) | 68 (26) |
| Standard Deviation from the normative means | -0.8 | -0.2 | -1.2† |
| Mazes | 23 (19) | 29 (28) | 30 (25) |
| Standard Deviation from the normative means | -0.1 | 0.8 | 2.53†† |
| Abandoned Utterances | 5.3 (4.1) | 8.7 (6.7) | 7.3 (4.0) |
| Standard Deviation from the normative means | 1.3† | 3.0††† | 2.1†† |

*Note.* Analyses were done using *Systematic Analyses of Language Transcripts* (SALT). KEB children were 8 to 13 years old. The symbol " † " means 1 SD and " †† " means 2 SDs from the mean.[a]n = 3; [b]n = 3; [c]n = 3.

Table 10

*Multivariate Analysis of Variance for English Language Sample Measures in the Three Proficiency Subgroups of Korean-English Bilingual (KEB) Children*

| | *df* | *F* | *p* |
|---|---|---|---|
| Complete & Intelligible Utterances | 2 | .333 | .729 |
| Different Words in 50 Utterances | 2 | .968 | .432 |
| Total Words in 50 Utterances | 2 | .563 | .597 |
| Grammatical Errors in C & I utterances | 2 | 3.031 | .123 |
| Mean Length of Utterance | 2 | .027 | .974 |
| Standard Deviation from the normative means | 2 | .099 | .907 |
| Words per Minute | 2 | .497 | .631 |
| Standard Deviation from the normative means | 2 | .716 | .526 |
| Mazes | 2 | .076 | .927 |
| Standard Deviation from the normative means | 2 | .587 | .585 |
| Abandoned Utterances | 2 | .325 | .735 |
| Standard Deviation from the normative means | 2 | .363 | .710 |

*Note.* Analyses were done using *Systematic Analyses of Language Transcripts* (SALT). KEB children were 8 to 13 years old. [a]n = 3; [b]n = 3; [c]n = 3.

Table 11

*Familiar and Unfamiliar English Consonants and Vowels Stimuli*

| | Familiar | Unfamiliar |
|---|---|---|
| Consonants | p, k, s, ʧ | b, g, f, θ, z, ʤ |
| Vowels | ɛ | ɪ, æ |

Table 12

*The NSCM Task: Average Percent Correct for Familiar and Unfamiliar Consonants and Vowels by Korean-English Bilingual (KEB) Children—including the English-dominant (ED), Balanced (BAL), and Korean-dominant (KD) KEB child groups—KEB Adults, and Korean Monolingual (KM) Children.*

| | Consonants | Vowels | Sound Type Combined |
|---|---|---|---|
| | | KEB Children[a] | |
| Familiar | 94 (3) | 72 (10) | 83 (5) |
| Unfamiliar | 78 (4) | 74 (12) | 76 (6) |
| Familiarity Combined | 85 (3) | 73 (9) | 78 (3) |
| | | KEB Adults[b] | |
| Familiar | 93 (4) | 64 (13) | 79 (6) |
| Unfamiliar | 72 (7) | 64 (8) | 68 (7) |
| Familiarity Combined | 81 (5) | 64 (5) | 72 (4) |
| | | KM Children[c] | |
| Familiar | 83 (15) | 46 (24) | 65 (18) |
| Unfamiliar | 54 (10) | 55 (13) | 55 (8) |
| Familiarity Combined | 66 (11) | 52 (10) | 59 (9) |
| | | All Participants | |
| Familiar | 88 (12) | 57 (22) | 73 (16) |
| Unfamiliar | 64 (13) | 62 (14) | 63 (12) |
| Familiarity Combined | 74 (12) | 60 (12) | 67 (11) |

Table 12 (continued)

| | Consonants | Vowels | Sound Type Combined |
|---|---|---|---|
| | | ED Children[d] | |
| Familiar | 95 (2) | 76 (5) | 85 (2) |
| Unfamiliar | 81 (1) | 83 (8) | 82 (5) |
| Familiarity Combined | 86 (1) | 80 (4) | 81 (4) |
| | | BAL Children[e] | |
| Familiar | 93 (5) | 77 (1) | 85 (8) |
| Unfamiliar | 73 (4) | 79 (8) | 76 (2) |
| Familiarity Combined | 81 (4) | 79 (2) | 80 (1) |
| | | KD Children[f] | |
| Familiar | 94 (3) | 65 (12) | 79 (5) |
| Unfamiliar | 79 (4) | 62 (6) | 70 (5) |
| Familiarity Combined | 86 (4) | 63 (1) | 74 (2) |

*Note*. The given value is the mean percent correct with standard deviation in "( )." KEB and KM children were 8 to 13 years old; KEB adults were 21 to 33 years old. $N$ = 31; [a]n = 8; [b]n = 7; [c]n = 16; [d]n = 3; [e]n = 2; [f]n = 3.

Table 13

*The NSCM Task: A Repeated Measures ANOVA with Sound Type and Familiarity as Within-group Factors in Korean-English Bilingual (KEB) Children, KEB Adult, and Korean Monolingual (KM) Children*

| | *df* | *F* | *p* |
|---|---|---|---|
| | | Between-group Factor | |
| Group (G) | 2 | 15.835** | .000 |
| | | Within-group Factors | |
| Sound Type (S) | 2 | 140.986** | .000 |
| S x G | 2 | .737 | .488 |
| Familiarity (F) | 2 | 16.443** | .000 |
| F x G | 2 | .152 | .860 |
| S x F | 2 | 51.621** | .000 |
| S x F x G | 2 | 1.587 | .222 |

*Note*. KEB and KM children were 8 to 13 years old; KEB adults were 21 to 33 years old. $N = 31$ ($n = 8$ for KEB children, $n = 7$ for KEB adults, $n = 16$ for KM children).

*$p < .05$. **$p < .01$.

Table 14

*The NSCM Task: Paired t-test for Familiar and Unfamiliar Consonants and Vowels in Korean-English Bilingual (KEB) Children, KEB Adults, and Korean Monolingual (KM) Children.*

| Group | Perceptual Performance | *t* | *p* |
|---|---|---|---|
| | | Consonants | |
| All participants | Familiar > Unfamiliar | $t_{30}$ = 14.804** | *p* = .000 |
| KEB children | Familiar > Unfamiliar | $t_{7}$ = 20.220** | *p* = .000 |
| KEB adults | Familiar > Unfamiliar | $t_{6}$ = 9.768** | *p* = .000 |
| KM children | Familiar > Unfamiliar | $t_{15}$ = 9.729** | *p* = .000 |
| | | Vowels | |
| All participants | Familiar ≈ Unfamiliar | $t_{30}$ = - 1.310 | *p* = .200 |
| KEB children | Familiar ≈ Unfamiliar | $t_{7}$ = - .629 | *p* = .549 |
| KEB adults | Familiar ≈ Unfamiliar | $t_{6}$ = .035 | *p* = .974 |
| KM children | Familiar ≈ Unfamiliar | $t_{15}$ = - 1.256 | *p* = .228 |

*Note.* $N$ = 31 ($n$ = 8 for KEB children, $n$ = 7 for KEB adults, $n$ = 16 for KM children).

*$p < .05$. **$p < .01$.

Table 15

*The NSCM Task for Age: A Repeated Measures ANOVA with Sound Type and Familiarity as Within-group Factors in Korean-English Bilingual (KEB) Children and Adults*

| | *df* | *F* | *p* |
|---|---|---|---|
| | | Between-group Factor | |
| Group (G) | 1 | 6.952* | .021 |
| | | Within-group Factors | |
| Sound Type (S) | 1 | 69.360** | .000 |
| S x G | 1 | 1.306 | .274 |
| Familiarity (F) | 1 | 28.877** | .000 |
| F x G | 1 | .940 | .350 |
| S x F | 1 | 33.160** | .000 |
| S x F x G | 1 | .016 | .901 |

*Note.* KEB children were 8 to 13 years old; KEB adults were 21 to 33 years old. $N = 15$ ($n = 8$ for KEB children, $n = 7$ for KEB adults). $^*p < .05$. $^{**}p < .01$.

Table 16

*The NSCM Task for Proficiency: A Repeated Measures ANOVA with Sound Type and Familiarity as Within-group Factors in English-dominant (ED), Balanced (BAL), and Korean-dominant (KD) Korean-English Bilingual Children*

| | *df* | *F* | *p* |
|---|---|---|---|
| | | Between-group Factor | |
| Group (G) | 2 | 11.484* | .014 |
| | | Within-group Factors | |
| Sound Type (S) | 1 | 51.657** | .001 |
| S x G | 2 | 7.429* | .032 |
| Familiarity (F) | 1 | 9.539* | .027 |
| F x G | 2 | .293 | .758 |
| S x F | 1 | 17.169** | .009 |
| S x F x G | 2 | .232 | .801 |

*Note*. $N = 8$ ($n = 3$ for ED, $n = 2$ BAL, and $n = 3$ for KD). *$p < .05$. **$p < .01$.

Table 17

*The NSCM: Percent Correct for the 10 Consonants and 3 Vowels in Korean-English Bilingual (KEB) children, KEB adults, and Korean Monolingual (KM) children*

| Phonemes | Korean-English Bilingual Children[a] | Korean-English Bilingual Adults[b] | Korean Monolingual Children[c] | All Participants |
|---|---|---|---|---|
| | | Consonants | | |
| Familiar | | | | |
| p | 97 (4) | 97 (4) | 86 (17) | 93 (14) |
| k | 96 (5) | 97 (3) | 84 (19) | 92 (15) |
| s | 88 (9) | 85 (14) | 85 (20) | 86 (16) |
| ʧ | 96 (3) | 95 (2) | 77 (20) | 89 (17) |
| Unfamiliar | | | | |
| b | 85 (16) | 89 (8) | 69 (23) | 81 (21) |
| g | 97 (2) | 97 (3) | 90 (13) | 95 (10) |
| f | 31 (19) | 37 (12) | 34 (14) | 34 (15) |
| θ | 75 (10) | 61(19) | 33 (20) | 56 (26) |
| z | 89 (5) | 70 (15) | 43 (16) | 67 (24) |
| ʤ | 88 (6) | 79 (11) | 55 (14) | 74 (19) |
| | | Vowels | | |
| Familiar | | | | |
| ɛ | 72 (10) | 64 (13) | 46 (24) | 60 (22) |
| Unfamiliar | | | | |
| ɪ | 70 (21) | 56 (17) | 51 (23) | 59 (22) |
| æ | 78 (11) | 71 (15) | 59 (20) | 70 (18) |

*Note*. The given value is the mean percent correct with standard deviation in "( )." KEB and KM children were 8 to 13 years old; KEB adults were 21 to 33 years old. $N = 31$; [a]n = 8; [b]n = 7; [c]n = 16.

Table 18

*The NSCM: MANOVA for the 10 Consonants and 3 Vowels in Korean-English Bilingual (KEB) children, KEB adults, and Korean Monolingual (KM) children*

| Phonemes | *df* | *F* | *p* |
|---|---|---|---|
| | | Consonants | |
| Familiar | | | |
| p | 2 | 2.808 | .077 |
| k | 2 | 2.888 | .072 |
| s | 2 | .077 | .926 |
| ʧ | 2 | 5.825** | .008 |
| Unfamiliar | | | |
| b | 2 | 3.904* | .032 |
| g | 2 | 2.198 | .130 |
| f | 2 | .226 | .799 |
| θ | 2 | 17.464** | .000 |
| z | 2 | 30.207** | .000 |
| ʤ | 2 | 24.398** | .000 |
| | | Vowels | |
| Familiar | | | |
| ɛ | 2 | 5.113* | .013 |
| Unfamiliar | | | |
| ɪ | 2 | 2.196 | .130 |
| æ | 2 | 3.692* | .038 |

*Note*. KEB and KM children were 8 to 13 years old; KEB adults were 21 to 33 years old. $N = 31$ ($n = 8$ for KEB children, $n = 7$ for KEB adults, $n = 16$ for KM children). $*p < .01$, $**p < .05$.

Table 19

*The NSCM: Perceptual Confusions for the Six Consonants and the Three Vowels that Elicited Varying Responses in Korean-English Bilingual (KEB) children, KEB adults, and Korean Monolingual (KM) children*

| Response | Korean-English Bilingual Children[a] | Korean-English Bilingual Adults[b] | Korean Monolingual Children[c] | All Participants |
|---|---|---|---|---|
| | | Presented Stimulus / ʧ / | | |
| ʧ | 96 | 95 | 77 | 89 |
| ʤ | 4 | 3 | 6 | 4 |
| k | 1 | 1 | 2 | 2 |
| | | Presented Stimulus / **b** / | | |
| b | 85 | 89 | 69 | 81 |
| f | 2 | 3 | 14 | 6 |
| θ | 10 | 5 | 7 | 7 |
| | | Presented Stimulus / **f** / | | |
| f | 31 | 37 | 34 | 34 |
| b | 15 | 19 | 22 | 18 |
| θ | 47 | 33 | 18 | 33 |
| | | Presented Stimulus / θ / | | |
| θ | 75 | 61 | 33 | 56 |
| f | 12 | 19 | 24 | 18 |
| s | 3 | 8 | 13 | 8 |
| | | Presented Stimulus / **z** / | | |
| z | 89 | 70 | 43 | 67 |
| s | 7 | 10 | 18 | 12 |
| ʤ | 2 | 19 | 29 | 17 |

Table 19 (continued)

| Response | Korean-English Bilingual Children[a] | Korean-English Bilingual Adults[b] | Korean Monolingual Children[c] | All Participants |
|---|---|---|---|---|
| | | Presented Stimulus / ʤ / | | |
| ʤ | 88 | 79 | 55 | 74 |
| z | 2 | 13 | 25 | 13 |
| ʧ | 4 | 7 | 4 | 5 |
| | | Presented Stimulus / ɛ / | | |
| ɛ | 72 | 64 | 46 | 60 |
| ɪ | 6 | 4 | 10 | 7 |
| æ | 22 | 33 | 44 | 33 |
| | | Presented Stimulus / ɪ / | | |
| ɪ | 70 | 56 | 51 | 59 |
| ɛ | 26 | 39 | 32 | 32 |
| æ | 4 | 4 | 17 | 8 |
| | | Presented Stimulus / æ / | | |
| æ | 78 | 71 | 59 | 70 |
| ɪ | 1 | 0 | 7 | 3 |
| ɛ | 21 | 28 | 33 | 27 |

*Note.* The given value is the mean percentage of responses for the given target. KEB and KM children were 8 to 13 years old; and KEB adults were 21 to 33 years old. $N$ = 31; [a]n = 8; [b]n = 7; [c]n = 16.

Table 20

*The NSIT: Average Percent Correct for Familiar and Unfamiliar Consonants and Vowels by Korean-English Bilingual (KEB) Children, KEB Adults, and Korean Monolingual (KM) Children.*

| | Consonants | Vowels | Sound Type Combined |
|---|---|---|---|
| | | KEB Children[a] | |
| Familiar | 99 (3) | 79 (14) | 89 (6) |
| Unfamiliar | 91 (6) | 91 (14) | 91 (9) |
| Familiarity Combined | 94 (4) | 87 (13) | 91 (7) |
| | | KEB Adults[b] | |
| Familiar | 88 (7) | 91 (13) | 89 (9) |
| Unfamiliar | 86 (10) | 71 (8) | 78 (8) |
| Familiarity Combined | 87 (7) | 77 (8) | 82 (7) |
| | | KM Children[c] | |
| Familiar | 90 (6) | 92 (13) | 91 (7) |
| Unfamiliar | 77 (15) | 59 (8) | 68 (16) |
| Familiarity Combined | 83 (6) | 70 (14) | 79 (11) |

Table 20 (continued)

| | Consonants | Vowels | Sound Type Combined |
|---|---|---|---|
| | | All Participants | |
| Familiar | 92 (7) | 89 (14) | 90 (7) |
| Unfamiliar | 83 (14) | 73 (19) | 79 (16) |
| Familiarity Combined | 86 (9) | 76 (14) | 83 (11) |

*Note.* The given value is the mean percent correct with standard deviation in "( )." KEB and KM children were 8 to 13 years old; KEB adults were 21 to 33 years old. *N* = 35 (*n* = 9 for KEB children, *n* = 8 for KEB adults, *n* = 18 for KM children).

Table 21

*The NSIT: A Repeated Measures ANOVA with Sound Type and Familiarity as Within-group Factors in Korean-English Bilingual (KEB) Children, KEB Adults, and Korean Monolingual (KM) Children*

| | *df* | *F* | *p* |
|---|---|---|---|
| | | Between-group Factor | |
| Group (G) | 2 | 4.325* | .022 |
| | | Within-group Factors | |
| Sound Type (S) | 2 | 16.902** | .000 |
| S x G | 2 | 2.473 | .100 |
| Familiarity (F) | 2 | 23.590** | .007 |
| F x G | 2 | 10.818** | .000 |
| S x F | 2 | 1.395 | .246 |
| S x F x G | 2 | 21.228** | .000 |

*Note.* KEB and KM children were 8 to 13 years old; KEB adults were 21 to 33 years old.

$N = 35$ ($n = 9$ for KEB children, $n = 8$ for KEB adults, $n = 18$ for KM children).

*$p < .05$. **$p < .01$.

Table 22

*The NSIT: Paired t-test for Familiar and Unfamiliar Consonants and Vowels by Korean-English Bilingual (KEB) Children, KEB Adults, and Korean Monolingual (KM) Children.*

| Group | Production Performance | *t* | *p* |
|---|---|---|---|
| | | Consonants | |
| All participants | Familiar > Unfamiliar | $t_{34}$ = 4.081** | *p* = .000 |
| KEB children | Familiar > Unfamiliar | $t_{8}$ = 2.742* | *p* = .025 |
| KEB adults | Familiar ≈ Unfamiliar | $t_{7}$ = .337 | *p* = .746 |
| KM children | Familiar > Unfamiliar | $t_{17}$ = 3.511** | *p* = .003 |
| | | Vowels | |
| All participants | Familiar > Unfamiliar | $t_{34}$ = 4.040** | *p* = .000 |
| KEB children | Familiar < Unfamiliar | $t_{8}$ = - 3.337** | *p* = .010 |
| KEB adults | Familiar > Unfamiliar | $t_{7}$ = 5.128** | *p* = .001 |
| KM children | Familiar > Unfamiliar | $t_{17}$ = 6.806** | *p* = .000 |

*Note*. *N* = 34 (*n* = 8 for KEB children, *n* = 8 for KEB adults, *n* = 18 for KM children).

*$p$ < .05. **$p$ < .01.

Table 23

*The NSIT for Age: A Repeated Measures ANOVA with Sound Type and Familiarity as Within-group Factors by Korean-English Bilingual (KEB) Children and Adults*

| | *df* | *F* | *p* |
|---|---|---|---|
| | | Between-group Factor | |
| Group (G) | 1 | 8.957* | .009 |
| | | Within-group Factors | |
| Sound Type (S) | 1 | 11.635** | .004 |
| S x G | 1 | 2.631 | .126 |
| Familiarity (F) | 1 | 5.409* | .034 |
| F x G | 1 | 7.528* | .015 |
| S x F | 1 | .886 | .361 |
| S x F x G | 1 | 28.415** | .000 |

*Note.* KEB children were 8 to 13 years old; KEB adults were 21 to 33 years old. $N = 17$ ($n = 9$ for KEB children, $n = 8$ for KEB adults). *$p < .05$. **$p < .01$.

Table 24

*The NSIT: Average Percent Correct for Familiar and Unfamiliar Consonants and Vowels by English dominant (ED), Balanced (BAL), and Korean dominant (KD) Korean-English Bilingual (KEB) Children.*

| | Consonants | Vowels | Sound Type Combined |
|---|---|---|---|
| | | English dominant KEB Children[a] | |
| Familiar | 97 (5) | 78 (16) | 87 (6) |
| Unfamiliar | 90 (6) | 98 (1) | 94 (3) |
| Familiarity Combined | 93 (4) | 91 (5) | 92 (3) |
| | | Balanced KEB Children[b] | |
| Familiar | 97 (6) | 90 (4) | 93 (5) |
| Unfamiliar | 94 (4) | 96 (4) | 95 (3) |
| Familiarity Combined | 95 (2) | 94 (3) | 95 (2) |
| | | Korean dominant KEB Children[c] | |
| Familiar | 100 (0) | 73 (11) | 86 (6) |
| Unfamiliar | 93 (8) | 80 (20) | 87 (14) |
| Familiarity Combined | 96 (2) | 78 (8) | 87 (11) |

*Note.* The given value is the mean percent correct with standard deviation in "( )."  KEB children were 8 to 13 years old. *N* = 9; [a]n = 3; [b]n = 3; [c]n = 3.

Table 25

*The NSIT for Proficiency: A Repeated Measures ANOVA with Sound Type and Familiarity as Within-group Factors in English-dominant (ED), Balanced (BAL), and Korean-dominant (KD) Korean-English Bilingual Children*

| | *df* | *F* | *p* |
|---|---|---|---|
| | | Between-group Factor | |
| Group (G) | 1 | .747 | .513 |
| | | Within-group Factors | |
| Sound Type (S) | 1 | 16.425** | .008 |
| S x G | 1 | 4.579 | .062 |
| Familiarity (F) | 1 | .082 | .784 |
| F x G | 1 | .536 | .611 |
| S x F | 1 | 14.675** | .009 |
| S x F x G | 1 | .635 | .562 |

*Note.* KEB children were 8 to 13 years old. $N = 9$ ($n = 3$ for ED children, $n = 3$ for BAL children, $n = 3$ for KD children). *$p < .05$. **$p < .01$.

Table 26

*The NSCM Task and NSIT Correlation: Partial Pearson Product-Moment Correlation Coefficients between Perceptual Performance on the NSCM task and Production Performance on the NSIT by Korean-English Bilingual (KEB) children, KEB adults, and Korean Monolingual (KM) Children*

| | Consonants | Vowels | Sound Type Combined |
|---|---|---|---|
| | | Group Controlled | |
| Familiar | .03 (.899) | -.25 (.192) | -.25 (.200) |
| Unfamiliar | .32 (.087) | .25 (.200) | .36 (.079) |
| Familiarity Combined | .36 (.058) | .26 (.170) | .29 (.123) |
| | | Group Uncontrolled | |
| Familiar | .02 (.932) | - .33 (.072) | - .32 (.089) |
| Unfamiliar | .38* (.036) | - .30* (.013) | .40* (.027) |
| Familiarity Combined | .37* (.044) | .31 (.099) | .33 (.071) |

*Note.* Values of Pearson correlation coefficients were given with p-values in "( )." $N = 31$; ($n = 8$ for KEB children; $n = 7$ for KEB adults; $n = 16$ for KM children). Group controlled means participants were separated into the three named groups. Group uncontrolled means that participants were combined for the analysis. *$p < .05$. **$p < .01$.

Table 27

*16 English Nonwords as Stimuli at Each Length on the NRT*

| 1 syllable | 2 syllable | 3 syllable | 4 syllable |
|---|---|---|---|
| /nɑιb/ | /teιvɑk/ | /ʧin□ιtɑʊb/ | /veιtɑʧɑιd□ιp/ |
| /voʊp/ | /ʧoʊvæg/ | /nɑιʧoʊveιb/ | /dævoʊn□ιʧig/ |
| /tɑʊʤ/ | /væʧɑιp/ | /d□ιtaʊvæb/ | /nɑιʧ□ιtɑʊvub/ |
| /d□ιf/ | /n□ιtɑʊf/ | /teιv□ιʧɑιg/ | /tævɑʧιnɑιg/ |

*Note.* Adopted from “Nonword Repetition and Child Language Impairment” by Dollaghan and Campbell, 1998, *Journal of Speech, Language, and Hearing Research*, *41*, p. 1138.

Table 28

*The NRT: Average Percentage of Phonemes Correct (PPC) for each Nonword Length and for all 16 Nonwords on the NRT in Korean-English Bilingual (KEB) Children, KEB Adults, and Korean Monolingual (KM) Children.*

| | KEB Children[a] | KEB Adults[b] | KM Children[c] | All Participants |
|---|---|---|---|---|
| 1 Syllable | 83 (10) | 83 (5) | 70 (8) | 75 (9) |
| 2 Syllable | 88 (5) | 93 (5) | 78 (9) | 84 (10) |
| 3 Syllable | 91 (5) | 89 (3) | 80 (9) | 84 (9) |
| 4 Syllable | 81 (6) | 88 (8) | 67 (12) | 75 (14) |
| Total | 86 (4) | 89 (4) | 74 (7) | 80 (9) |

*Note.* The given value is the mean percent correct with standard deviation in "( )." KEB and KM children were 8 to 13 years old; KEB adults were 21 to 33 years old. *N* = 32; [a]n = 7; [b]n = 7; [c]n = 18.

Table 29

*The NRT: A Repeated Measures ANOVA with Nonword Length as a Within-group Factor in Korean-English Bilingual (KEB) Children, KEB Adults, and Korean Monolingual (KM) Children*

| | *df* | *F* | *p* |
|---|---|---|---|
| | | Between-group Factor | |
| Group (G) | 2 | 25.082** | .000 |
| | | Within-group Factors | |
| Length (L) | 4 | 10.482** | .000 |
| L x G | 4 | 1.617 | .127 |

*Note.* KEB and KM children were 8 to 13 years old; KEB adults were 21 to 33 years old. $N = 32$ ($n = 7$ for KEB children, $n = 7$ for KEB adults, $n = 18$ for KM children).

*$p < .05$. **$p < .01$.

Table 30

*The NSCM Task and NRT Correlation: Partial Pearson Product-Moment Correlation Coefficients between Perceptual Performance on the NSCM Task and Phonological Awareness, as Measured by NRT, in Korean-English Bilingual (KEB) children, KEB adults, and Korean Monolingual (KM) Children*

| | Consonants | Vowels | Sound Type Combined |
|---|---|---|---|
| | | Group Controlled | |
| Familiar | - .04 (.851) | - .15 (.456) | - .12 (.568) |
| Unfamiliar | .16 (.414) | .47* (.013) | .39* (.046) |
| Familiarity Combined | .09 (.646) | .29 (.142) | .15 (.460) |
| | | Group Uncontrolled | |
| Familiar | .28 (.149) | .22 (.254) | .26 (.176) |
| Unfamiliar | .57** (.002) | .54** (.003) | .63** (.000) |
| Familiarity Combined | .51** (.006) | .54** (.003) | .53** (.004) |

*Note.* Values of Pearson correlation coefficients were given with p-values in "( )." $N = 30$ ($n = 7$ for KEB children; $n = 7$ for KEB adults; $n = 16$ for KM children). Group controlled means participants were separated into the three named groups. Group uncontrolled means that participants were combined for the analysis. The NRT score entered into the correlations was always the total percent of phonemes correct (i.e., TOT-PPC). $*p < .05$. $**p < .01$.

Table 31

*The NSIT and NRT Correlation: Partial Pearson Product-Moment Correlation Coefficients between Production Performance on the NSIT and Phonological Awareness, as measured by NRT, in Korean-English Bilingual (KEB) children, KEB adults, and Korean Monolingual (KM) Children*

| | Consonants | Vowels | Sound Type Combined |
|---|---|---|---|
| | | Group Controlled | |
| Familiar | .11 (.568) | -.22 (.186) | -.18 (.337) |
| Unfamiliar | .24 (.204) | .17 (.370) | .21 (.263) |
| Familiarity Combined | .23 (.227) | .07 (.703) | .13 (.496) |
| | | Group Uncontrolled | |
| Familiar | .05 (.779) | -.33 (.073) | -.28 (.122) |
| Unfamiliar | .35 (.051) | .30 (.100) | .34 (.061) |
| Familiarity Combined | .30 (.108) | .18 (.324) | .23 (.206) |

*Note.* Values of Pearson correlation coefficients were given with p-values in "( )." $N = 30$; $n = 7$ for KEB children; $n = 7$ for KEB adults; $n = 16$ for KM children. Group controlled means participants were separated into the three named groups. Group uncontrolled means that participants were combined for the analysis. The NRT score entered into the correlations was always the total percent of phonemes correct (i.e., TOT-PPC). $^*p < .05$. $^{**}p < .01$.

Table 32

*Summary of findings from the Nonsense Syllable Confusion Matrix (NSCM) task and the Nonsense Syllable Imitation Task (NSIT)*

| | Sound Type by Familiarity | | Combined |
|---|---|---|---|
| | | the NSCM task | |
| All participants[a] | FC > UC | FV ≈ UV | F > U, C > V |
| KEB children[b] | FC > UC | FV ≈ UV | F > U, C > V |
| KEB adults[c] | FC > UC | FV ≈ UV | F > U, C > V |
| KM children[d] | FC > UC | FV ≈ UV | F > U, C > V |
| | | the NSIT | |
| All participants[e] | FC > UC | FV > UV | F > U, C > V |
| KEB children[f] | FC > UC | UV > FV | F ≈ U, C > V |
| KEB adults[g] | FC ≈ UC | FV > UV | F > U, C > V |
| KM children[h] | FC > UC | FV > UV | F > U, C > V |

*Note.* Korean-English bilingual (KEB) and Korean monolingual (KM) children were 8 to 13 years old; KEB adults were 21 to 33 years old. [a]$n$ = 31; [b]$n$ = 8; [c]$n$ = 7; [d]$n$ = 16; [e]$n$ = 35; [f]$n$ = 9; [g]$n$ = 8; [h]$n$ = 18. The symbol “>” represents that the left hand is significantly greater than the right hand; and the symbol “≈” represents that the difference between the two is not statistically significant.

# FIGURES

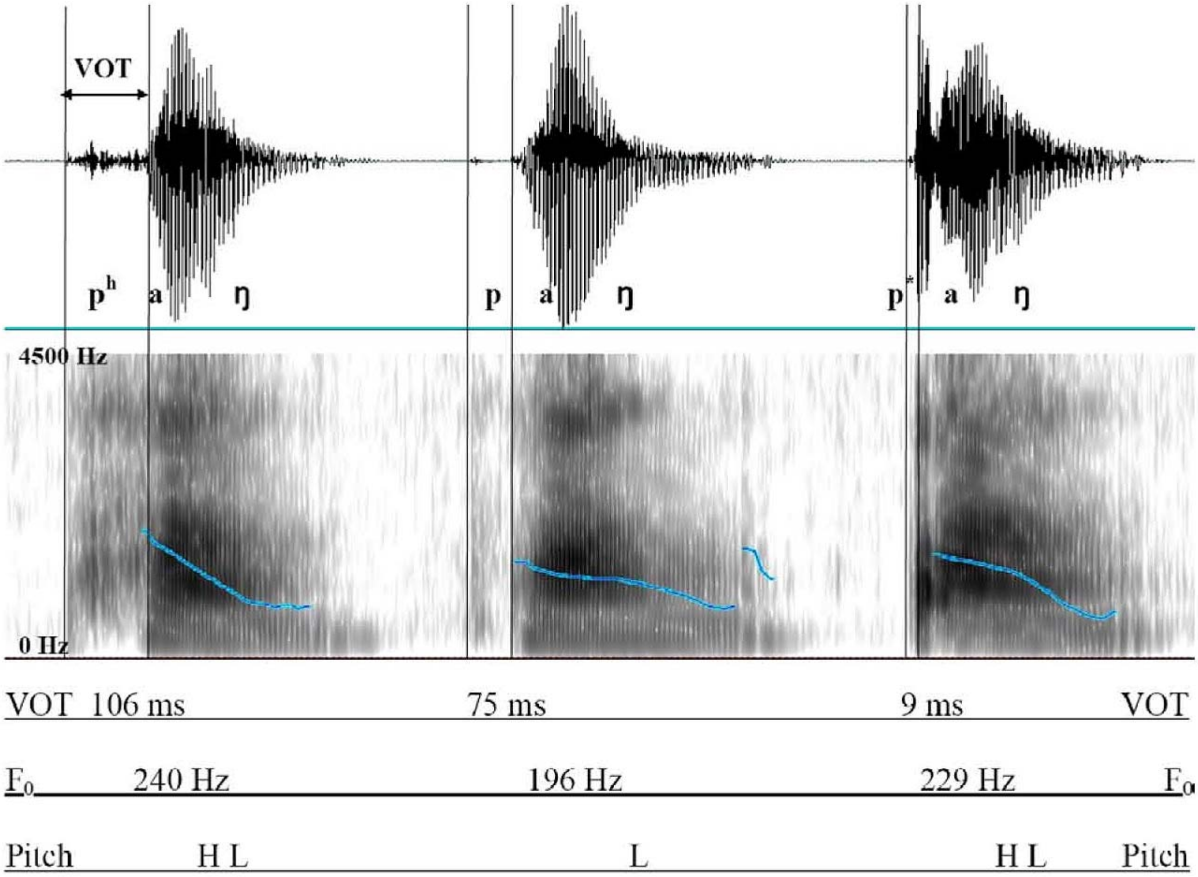


*Figure 1.* Production of three types of Korean voiceless bilabial stops by a female native speaker of Korean. The stops are /p$^h$, p, p$^*$/ in the words /p$^h$aŋ/, /paŋ/, and /p$^*$aŋ/. Voice onset time (VOT) is measured in ms; fundamental frequency ($F_0$) is in Hz; and pitch is described with H for a high tone and L for a low tone. Pitch contours are presented as a blue line on the spectrograms.

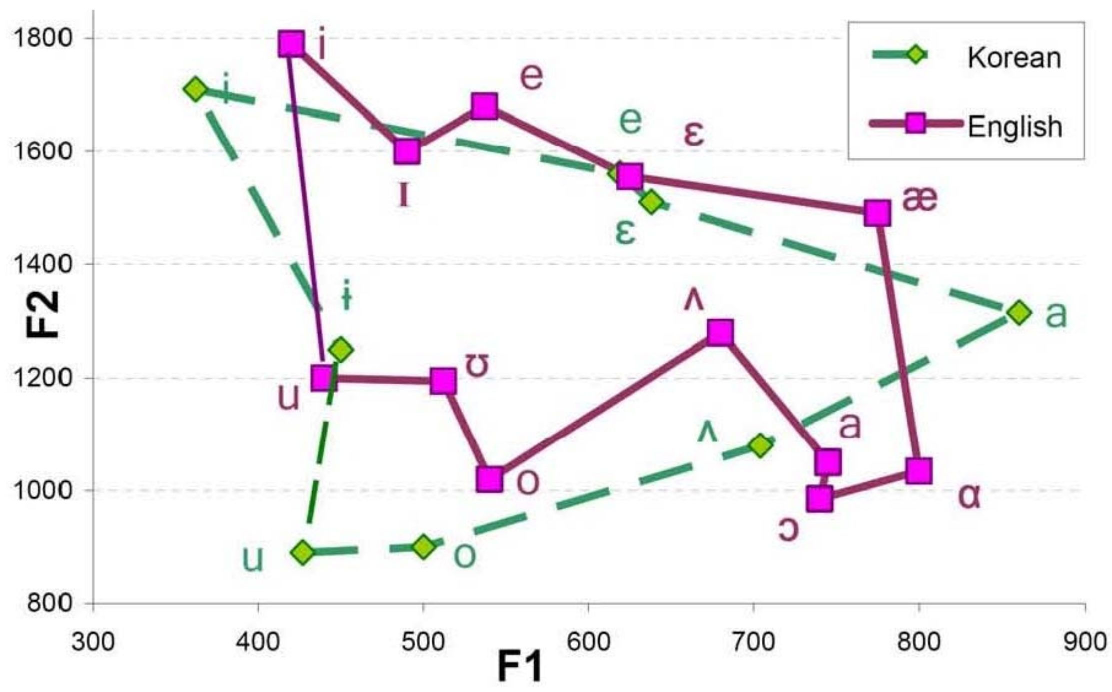


*Figure 2*. Vowel formants of American-English produced by American-English southern dialect speakers (solid line) and vowel formants of Korean by Korean Seoul-dialect speakers (dashed line).

---

Adapted from "A comparative study of American English and Korean vowels produced by male and female speakers" by Byunggon Yang, 1996, *Journal of Phonetics* 24, p. 258 with oral consent from the author.

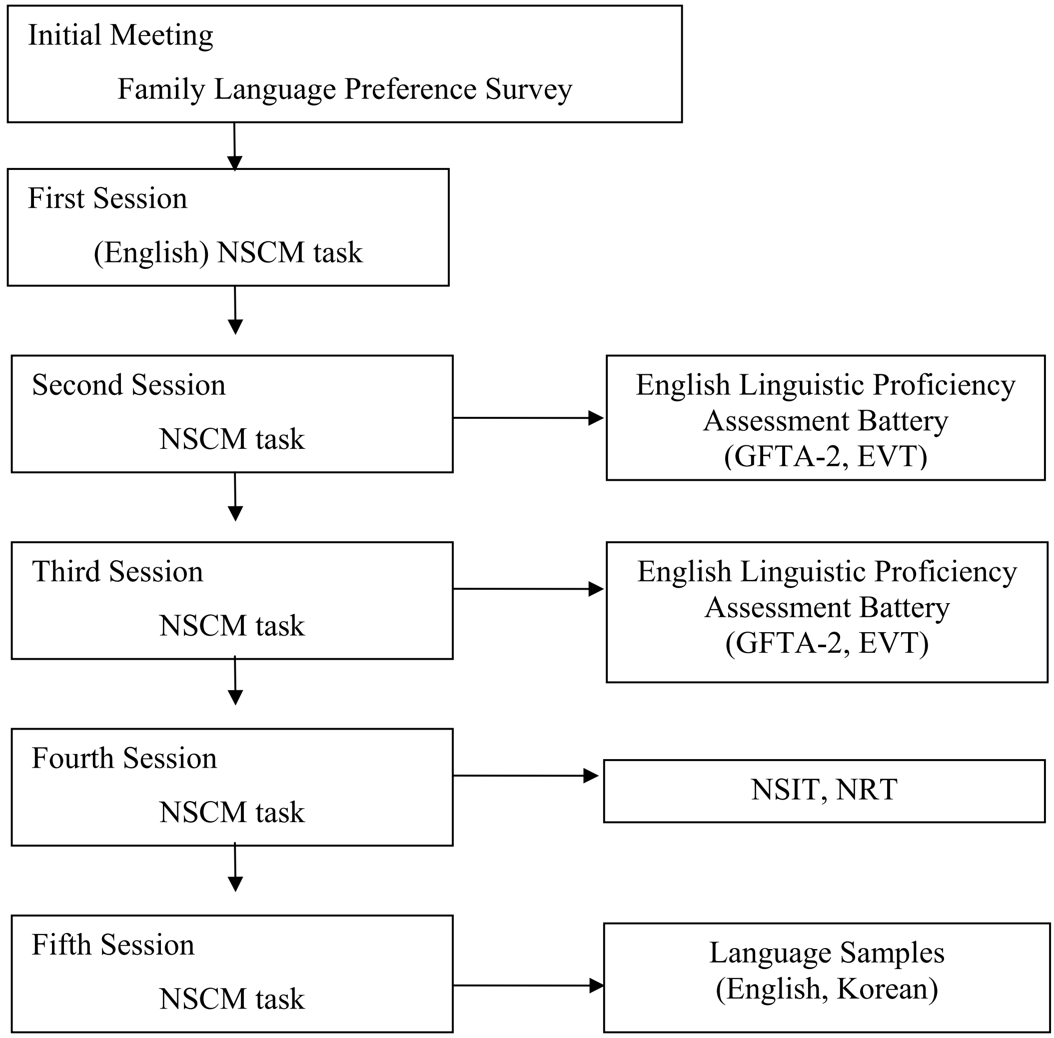


*Figure 3.* General procedure for completing a whole set of experimental sessions. The Nonsense Syllable Confusion Matrix (NSCM) task, the Nonsense Syllable Imitation Task (NSIT), the Nonword Repetition Task (NRT), the Goldman Fristoe Test of Articulation-2 (GFTA-2), and the Expressive Vocabulary Test (EVT).

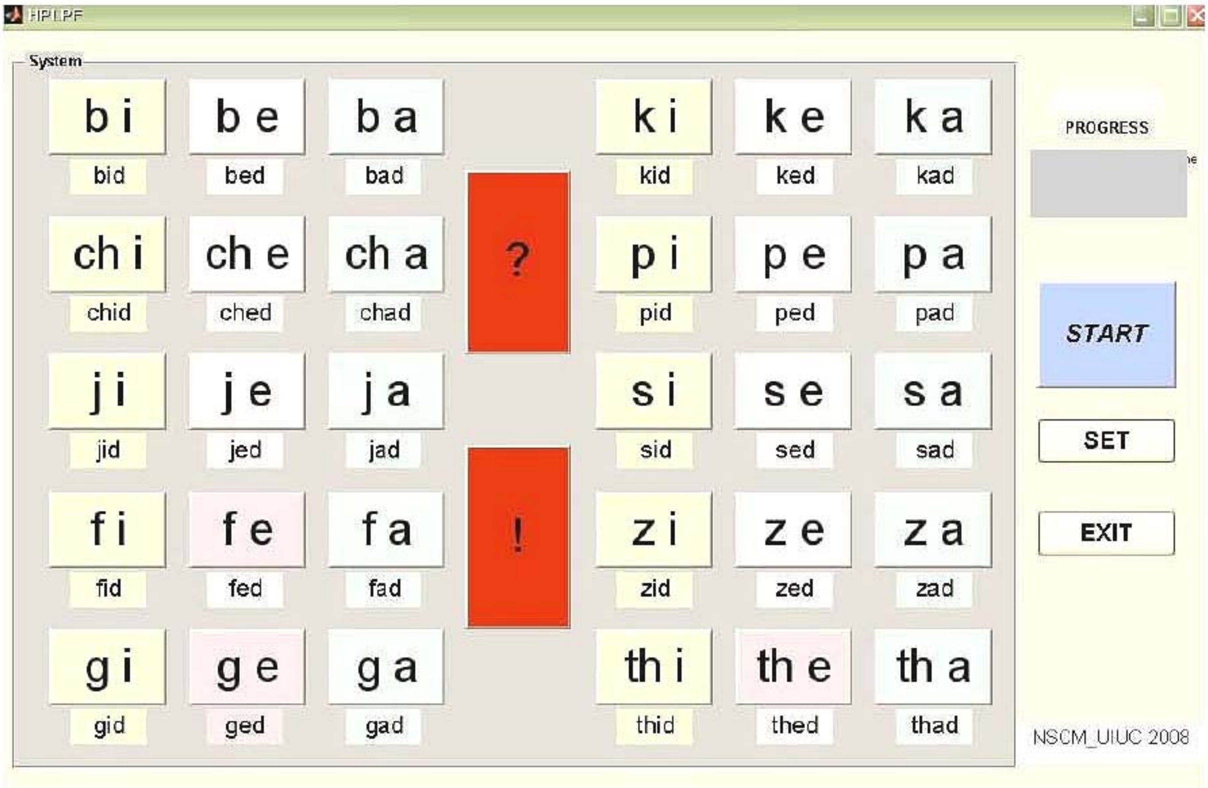


*Figure 4.* A display of the computer screen in the Nonsense Syllable Confusion Matrix (NSCM) task.

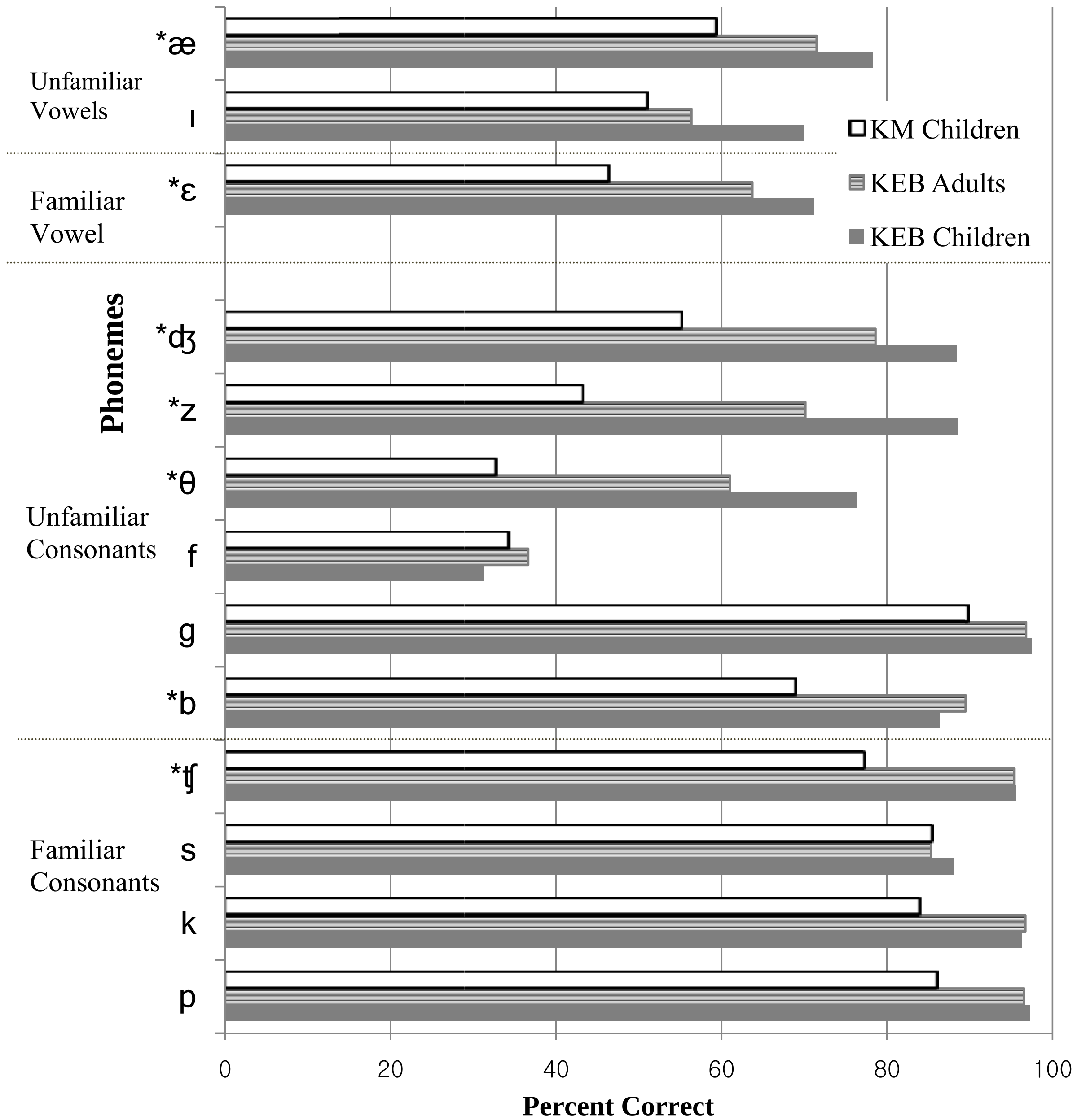


*Figure 5*. Average percent correct for 10 consonants and 3 vowels by Korean-English bilingual (KEB) children[a], KEB adults[b], and Korean monolingual (KM) children[c] on the NSCM Task. $N$ = 31; [a]n = 8; [b]n = 7; [c]n = 16. KEB and KM children were 8 to 13 years old; KEB adults were 21 to 33 years old. The phoneme with a symbol "*" to the left showed a significant group difference.

# REFERENCES


Abramson A.S. & Lisker, L. (1970). Discriminability along the voicing continuum: Cross-language tests. *Proceedings of the Sixth International Congress of Phonetic Sciences, Prague, 1967,* 569-573.

Abramson A.S. & Lisker, L. (1973). Voice-timing perception in Spanish word-initial stops. *Journal of Phonetics*, 1, 1-8.

Abu-Rabia, S. & Kehat, S. (2004). The critical period for second language pronunciation: Is there such a thing? Ten case studies of late starters who attained a native-like Hebrew accent. *Educational Psychology, 24*, 77-98.

Adler, S., & Birdsong, S. (1983). Reliability and validity of standardized testing tools used with poor children. *Topics in Language Disorders*, *3*, 76-88.

Ahn, S-C. (1998). *An Introduction to Korean Phonology*. Seoul: Hanshin

Abdelatty Ali, A.M., der Spiegel, J.V., & Mueller, P. (2001). Robust classification of stop consonants using auditory-based speech processing. *IEEE International Conference on Acoustics, Speech and Signal Processing*, *1*, 81-84.

Allen, Jont B. (2005). Consonant recognition and the Articulation index. *Journal of the Acoustical Society of America*, *117(4),* 2212-2223.

Allen, P., Wlghtman, F., Kiastler, D., & Dolan, T. (1989). Frequency resolution in children. *Journal of Speech and Hearing Research, 32*, 317-322.

Andrew, B.W. & Blumstein, S.E. (2009). Temporal integration in vowel perception. *Journal of Acoustical Society of America, 125(3)*, 1704-1711.

Aram, D., Morris, R., & Hall, N. (1993). Clinical and research congruence in identifying children with specific language impairment. *Journal of Speech and Hearing Research, 36*(3), 580-591.

Ardila, A., Rosselli, M., Ostrosky-Solis, F., Marcos, J., Granda, G., & Soto, M. (2000). Syntactic comprehension verbal memory and calculation abilities in Spanish-English bilinguals. *Applied Neuropsychology, 7,* 3-16.

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

Bedore, L. & Leonard, L. (1998). Specific language impairment and grammatical morphology: A discriminant function analysis. *Journal of Speech, Language, and Hearing Research, 41*, 1185-1192.

Benki, J. S. (2001). Place of articulation and first formant transition pattern both affect perception of voicing in English. *Journal of Phonetics, 29*, 1-22.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In C. Goodman & H. Nusbaum (Eds.), *The Development of Speech Perception*, 167-224. Cambridge: The MIT Press.

Best, C. T. (1995). A direct realist view of cross-language speech. In W. Strange (Ed.), *Speech Perception and Linguistic Experience*, 171-204. Baltimore: York.

Best, C. C. & McRoberts, G. W. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and Speech, 46*, 183–216.

Best, C. T., McRoberts, G., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America, 109(2)*, 775-794.

Best, C. T., McRoberts, G. W. & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology, 14(3)*, 345-360

Best, C. T. & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants, *Journal of phonetics, 20*, 305-330.

Bialystok E. & Miller, B. (1999). The problem of age in second-language acquisition: Influences from language, structure, and task. *Language and Cognition, 2(2)*, 127-145.

Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language, 68,* 706-755.

Bishop, D.V.M., North, T., & Donlan, C. (1996). Nonword repetition as a behavioral marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry, 37*, 391-403.

Bleile, K. (1995). *Manual of Articulation and Phonological Disorders: Infancy through Adulthood*. San Diego: Singular.

Blomert, L., Mitterer, H. & Paffen, C. (2004). In search of the auditory, phonetic, and/or phonological problems in dyslexia: context effects in speech perception. *Journal of Speech Language Hearing Research, 47(5)*, 1030-1047.

Boatman, D., Hall, C., Goldstein, M.H., Lesser, R., & Gordon, B. (1997). Neuroperceptual differences in consonants and vowel discrimination: as revealed by direct cortical electrical interference. *Cortex, 33*, 83-98.

Boersma, P. & Weenink, D. (2005). Praat: doing phonetics by computer (Version 4.3.14) [Computer program]. Retrieved May 26, 2005, from http://www.praat.org/

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

Boothroyd, A. (1984). Auditory perception of speech contrasts by subjects with sensorineural hearing loss. *Journal of Speech and Hearing Research, 28*, 185-196.

Borden, G. J., Harris, K. S., & Raphael, L. J. (1994). *Speech Science Primer: Physiology, Acoustics, and Perception of Speech*. Baltimore, MD: Lippincott Williams & Wilkins.

Brady S, Shankweiler D, & Mann, V. (1983). Speech perception and memory coding in relation to reading ability. *Journal of Experimental Child Psychology, 35*, 345–367.

Britton, B., Blumstein, S., Myers, E.B., & Grindrod, C. (2009). The role of spectral and durational properties on hemispheric asymmetries in vowel perception. *Neuropsychologia, 47(4)*, 1096-1106.

Browman, C. P. & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook, 3*, 219-252.

Browman, C. P. & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica, 49*, 155-180.

Brown, C. (2000). The interrelation between speech perception and phonological acquisition from infant to adult. In J. Archibald (Ed.), *Second Language Acquisition and Linguistic Theory*, 4-63. Wiley-Blackwell.

Burnham, D.K. (1986).Developmental loss of speech perception: Exposure to and experience with a first language. *Applied Linguistics, 7*, 201-240.

Campbell, T., Dollaghan, C., Needleman, H., & Janosky, J. (1997). Reducing bias in language assessment: Processing-dependent measures. *Journal of Speech, Language, and Hearing Research, 40,* 519–525.

Caramazza, A., Chialant, D., Capasso, R., & Miceli, G. (2000). Separable processing of consonants and vowels. *Macmillan Magazines Ltd*, 403, 428-430.

Cho, H.Y., Mo, Y., & Yoon, S-Y. (2004). An acoustic analysis of Korean speakers' production of English vowels. Unpublished manuscript, University of Illinois at Urbana-Champaign.

Cho, S-B. (1967). *A Phonological Study of Korea*. Uppsala: Almqvist and Wikells.

Cho, T. (1996). Vowel correlates to consonant phonation: An acoustic-perceptual study of Korean obstruents. Unpublished master's thesis, University of Texas at Arlington.

Cho, T., Jun, S-A., & Ladefoged, Peter. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics, 30,* 193-228.

Cho, T. & Keating, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics, 29*, 155-190.

Choi, Hyon-Pae. (1954). *Uli Malbon* (Korean Diction). Seoul: Jongumsa.

Choi, Hansook. (2002). Acoustic cues for the Korean stop contrast - Dialectal variation. *ZAS Papers in Linguistics, 28*, 1-12.

Christman, S.S. (1991). Uncovering phonological regularity in neologisms: contributions of sonority theory. *Clinical Linguistics and Phonetics, 6*, 219-247.

Cohen, L. B., Amsel, G., Redford, M. A., & Casasola, M. (1998). The development of infant causal perception. In A. Slater (Ed.), *Perceptual Development: Visual, Auditory, and Speech Perception in Infancy*, 167-209. East Sussex, UK: Psychology Press Ltd.

Cummins, J. (1992). Language proficiency, bilingualism, and academic achievement. In P. A. Richardmato & M. A. Snow (Eds.), *The Multicultural Classroom: Readings for Content-area Teachers*, 16-26. New York: Longman.

Cutler, A., D. Dahan, and W. van Donselaar. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40:141-201.

Dart, S. (1987). An aerodynamic study of Korean stop consonants: Measurements and Modeling. *Journal of the Acoustical Society of America, 81(1)*, 138-147.

Delattre, P. C., Cooper, F. S., & Liberman, A. M. (1955). Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America, 27*, 769-773.

DeThorne. L.S., Petrill, S.A., Schatschneider, C. & Cutting, L. (2010). Conversational language use as a predictor of early reading development: Language history as a moderating variable. *Journal of Speech, Language, and Hearing Research, 53*, 209-223.

Dodd, B., Holm, A., Hua, Z., & Crosbie, S. (2003). Phonological development: A normative study of British English-speaking Children. *Clinical Linguistics and Phonetics, 17(8)*, 617-643.

Dodd, B. & Lydia K.H. So. (1994). The phonological abilities of Cantonese-speaking children with hearing loss. *Journal of Speech and Hearing Research, 37*, 671-679.

Dollaghan, C. & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech Language Hearing Research, 41*, 1136-1146.

Dubno, J.R., Dirks, D.D., & Langhofer, L.R. (1992). Evaluation of hearing-impaired listeners using a nonsense-syllable test II, syllable recognition and consonant confusion patterns. *Journal of Speech and Hearing Research, 25*, 141-148.

Dunn, M., Flax, J., Sliwinski, M., & Arm, D. (1996). The use of spontaneous language measures as criteria in identifying children with specific language impairment: An attempt to reconcile clinical and research incongruence. *Journal of Speech and Hearing Research, 39*, 643-654.

Eimas, P.D. (1978). Developmental aspects of speech perception. In R. Held, H.W. Leibowitz, & H.L. Teuber (Eds.), *Handbook of Sensory Physiology, v8*. Berlin: Springer.

Eimas, P.D. & Miller, J.L. (1978). Effects of selective adaptation on the perception of speech and visual patterns: Evidence for feature detectors. In R.D. Walk and H.L. Pick Jr. (Eds.), *Perception and Experience*, 307-345. New York: Plenum.

Eimas, P.D. & Miller, J.L. (1980). Contextual effects in infant speech perception. *Science, 209*, 1140-1141.

Eimas, P.D., Siqueland, E.R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science, 171*, 303 – 306.

Elliott, L. L. (1979). Performance of children aged 9 to 17 on a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America, 66*, 651-653.

Elliott, L. L. (1986). Discrimination and response bias for CV syllables differing in voice onset time among children and adults. *Journal of the Acoustical Society of America, 80*, 1250-1255.

Elliott, L L, Busse, L., Partridge, R., Rupert, J., & DeGraaff, R. (1986). Adult and child discrimination of CV syllables differing in voice onset time. *Child Development, 57*, 628-635.

Elliott, L L, & Hammer, M.A. (1988). Longitudinal changes in auditory discrimination in normal children and children with language-learning problems. *Journal of Speech and Hearing Research, 53*, 467-474.

Elliott, L. L., Longinotti, C., Meyer, D., Raz, I., & Zucker, K. (1981). Developmental differences in identifying and discriminating CV syllables. *Journal of Acoustical Society of America, 70*, 669-677.

Epstein, S.D., Flynn, S., & Martohardjono, G. (1996). Second language acquisition: Theoretical and experimental issues in contemporary research. *Behavioral and Brain Sciences, 19*, 677-758.

Evans, J.L. & Miller, J. (1999). Language sample analysis in the 21st century. *Seminar in Speech and Language, 20(2)*, 101-115.

Fant, G. (1973). *Speech Sounds and Features*. Cambridge, MA: MIT Press.

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*, 47-65.

Flege, J. E. (1992). Speech learning in a second language. In C.A. Ferguson, L. Menn & C. Stoel-Gammon (Eds.), *Phonological Development: Models, Research, Implications*, 565-604. Baltimore: York.

Flege, J. E. (2002). Interactions between the native and second language phonetic systems. In Burmeister, P., Piske, T., & Rohde, A. (Eds.), *An Integrated View of Language Development: Papers in Honor of Henning Wode*. Wissenschaftlicher Verlag Trier.

Flege, J. E. (2003). Methods for assessing the perception of vowels in a second language. In E. Fava & A. Mioni (Eds.), *Issues in Clinical Linguistics,* 19-44. Padua, Italy:UniPress.

Flege, J. E., Bohn, O-S., & Jang, S. (1997). Effects of experience on nonnative speakers’ production and perception of English vowels. *Journal of Phonetics, 25*, 437-470.

Flege, J. E. & MacKay, Ian R. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition, 26*, 1-34.

Flege, J. E., Munro, M., & Fox, R. (1994). Auditory and categorical effects on cross-language vowel perception. *Journal of the Acoustical Society of America*, *95*, 3623-3641.

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

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

Flege, J. E., Takagi, N., & Mann, V. (1996). Lexical familiarity and English-language experience affect Japanese adults' perception of /r/ and /l/. *Journal of Acoustical Society of America*, *99* (2), 1161-1173.

Gallagher, A., Frith, U., & Snowling, M. J. (2000). Precursors of literacy delay among children at genetic risk of dyslexia. *Journal of Child Psychology and Psychiatry, 41*, 203–213.

Gavin, W.J. & Giles, L. (1996). Sample size effects on temporal reliability of language sample measures of preschool children. *Journal of Speech and Hearing Research, 39*, 1258-1262.

Gildersleeve-Neumann, C.E., Kester, E.S., Davis, B.L., & Pena, E.D. (2008). English speech sound development in preschool-aged children from bilingual English-Spanish environments. *Language, Speech, and Hearing Services in Schools, 39*, 314-328.

Goldstein, B.A. & Iglesias, Aquiles. (1996). Phonological patterns in normally developing Spanish-speaking 3- and 4-year olds of Puerto Rican descent. *Language, Speech, and Hearing Services in Schools, 27*, 82-90.

Goldstein, B.A. & Iglesias, Aquiles. (2001). The effect of dialect on phonological analysis: Evidence from Spanish-speaking children. *American Journal of Speech-Language Pathology, 10(4)*, 394.

Goldstein, B.A. & Kohnert, Kathryn (2005). Speech, language, and hearing in developing bilingual children: Current findings and future directions. *Language, Speech and Hearing Services in Schools 36*, 264-267.

Goldstein, B.A., Fabiano, L., & Iglesias, A. (2003). Phonological development in monolingual English, monolingual Spanish, and bilingual (Spanish–English) Latino children. 2003 May Poster session presented at the 4th annual International Symposium on Bilingualism, Tempe, AZ

Goldstein, B.A., Fabiano, L., & Washington, P.S. (2005). Phonological skills in predominantly English-speaking, predominantly Spanish-speaking, and Spanish-English bilingual children. *Language, Speech, and Hearing Services in Schools, 36*, 201-218.

Goldstein, B.A & Washington, P. (2001). An initial investigation of phonological patterns in 4-year-old typically developing Spanish–English bilingual children. *Language, Speech, and Hearing Services in Schools, 32*, 153-164.

Gottfried, T. L., Miller, J. L., & Payton, P.E. (1990). Effect of speaking rate on the perception of vowels. *Phonetica 47*, 155-172

Gottfried, T. L., & Strange, W. (1980). Identification of coarticulated vowels. *Journal of Acoustical Society of America, 68*, 1626-1635.

Gottlieb, G. (1981). Roles of early experience in species-specific perceptual development. In R.N. Aslin, J.R. Alberts, and M.R. Petersen (Eds.), *Development of Perception I.* New York: Academic Press.

Grieser, D. & Kuhl, P. K. (1989). Categorization of speech by infants: Support for speech-sound prototypes. *Developmental Psychology, 25*, 577-588.

Grosjean, Francois. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language, 36*, 3-15.

Haggard, M., Summerfield, O., & M. Roberts. (1981). Psychological and cultural determinants of phoneme boundaries: evidence from trading F0 cues in the voiced-voiceless distinction. *Journal of Phonetics, 9*, 49-62.

Hall, J. & Grose, J. (1991). Notched-noise measures of frequency selectivity in adults and children using fixed-masker level and fixed-signal level presentation. *Journal of Speech and Hearing Research, 34*, 651-660.

Halle, P.A. & Best, C.T. (2007). Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/ clusters. *Journal of Acoustical Society of America, 121(5)*, 2899-2914.

Halle, M. & Stevens, K. (1971). A note on laryngeal features. *Quarterly Progress Report of the Research Laboratory of Electronics 101*, 198-213.

Han, J-I. (1996a). The Phonetics and Phonology of ‘Tense’ and ‘Plain’ Consonants in Korean. Unpublished doctoral dissertation, Cornell University.

Han, J-I. (1996b). Perception of Korean tense and lax consonants: evidence for a geminate analysis of tense consonants. *Japanese/Korean Linguistics, 5*, 407-423.

Han, M.S. & Weitzman, R.S. (1970). Acoustic features of Korean /P, T, K/, /p, t, k/ and /ph, th, kh/. *Phonetica, 22,* 112-128.

Han, N. (1998). A comparative acoustic study of Korean by native Korean children and Korean-American children. Unpublished master's thesis, University of California at Los Angeles.

Hardcastle, W. J. (1973). Some observation on the tense-lax distribution in initial stops in Korean. *Journal of Phonetics, 1*, 263-272.

Harris, K. S. (1958). Cues for the discrimination of American English fricatives in spoken syllables. *Lang and Speech, 1*, 1-7.

Hillenbrand, J. M., Clark, M. J., & Houde, R. A. (2000). Some effects of duration on vowel recognition. *Journal of the Acoustical Society of America, 108*, 3013-3022.

Hillenbrand, J. M., Clark, M. J., and Nearey, T. M. (2001). Effects of consonant environment on vowel formant patterns. *Journal of the Acoustical Society of America, 109(2)*, 748-763.

Hillenbrand, J. M., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America, 97*, 3099-3111.

Hillenbrand, J. M., & Nearey, T. N. (1999). Identification of resynthesized /hVd/ syllables: Effects of formant contour. *Journal of the Acoustic Society of America, 105*, 3509-3523.

Hirose, Hajime, Lee, C-Y., & Ushijima, T. (1974). Laryngeal control in Korean stop production. *Journal of Phonetics, 2*, 145-152.

Holm, A. & Dodd, B. (1999). An intervention case study of a bilingual child with phonological disorder. *Child Language Teaching and Therapy, 15*, 139-158.

Holm, A., Dodd, B., Stow, C., & Pert, S. (1999). Identification and differential diagnosis of phonological disorder in bilingual children. *Language Testing, 16(3)*, 271-292.

Holt, Rachael Anneke Frush. (2003). Non-sensory factors in children's speech perception. Unpublished doctoral Dissertation, University of Minnesota.

de Houwer, A. (1995). Bilingual language acquisition. In P. Fletcher & B.MacWhinney (Eds.), *Handbook of Child Language*, 219-250. Oxford: Basil Blackwell Ltd.

Huang, R-J., Hopkins, J., & Nippold, M.A. (1997). Satisfaction with standardized language testing: A survey of speech-language pathologists. *Language, Speech, and Hearing Services in Schools, 28,* 12-29.

Hume, E., Johnson, K., Seo, M., & Tserdanelis, G. (1998). A cross-linguistic study of stop place perception. *Proceedings of the XIVth International Congress of Phonetic Sciences*. 2069-2072.

Hurford, J. (1991). The evolution of the critical period for language acquisition. *Cognition, 40(3)*, 159-201.

Iglesia, A., Miller, J.F., Nockerts, A. (2009). Using SALT Software to Assess the Language Production of Bilingual (Spanish-English) Children. [Computer Software and Manual]. University of Wisconsin-Madison & SALT Software, LLC. Retrieved from http://www.SALTSoftware.com.

Ingram, D. (1989). *Phonological Disability in Children: Studies in Disorders of Communication.* Second Edition. London: Whurr Publishers Limited.

Ingram, J. C. L., & Park, S. G. (1998). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics, 25(3)*, 343-370.

Iverson, P. & Kuhl, P. K. (1996). Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/. *Journal of the Acoustical Society of America, 99*, 1130-1140.

Jax, V. A. (1988). *Narrative construction by children learning English as a second language: A precursor to reading comprehension. Unpublished doctoral dissertation,.* University of California, Los Angeles.

Jiang, J., Chen, M., & Alwan, A. (2003). On the perception of voicing in syllable-initial plosives in noise, *Journal of Acoustical Society of America, 40*, 291-313.

Johnson, C.J., Phatak, S., Steele, S.D., Lobdell, B., Allen, J.B., Yoon, S-Y., & Frame, S. (2007). Reading disability profiles and consonant-vowel confusions in perception-production tasks. Presented in *2007 American Speech and Hearing Association Convention*.

Johnson, H.J. & Newport, E.L. (1987). Critical period effect in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60-99.

Jun, S-A. (1993). The phonetics and phonology of Korean prosody. Published doctoral dissertation, Ohio State University.

Jun, S-A. (1995) Asymmetrical prosodic effects on the laryngeal gesture in Korean. In B. Connell & A. Arvaniti, (Eds.), *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV*, 235-253. Cambridge, U.K.: Cambridge University Press.

Jun, S-A. (2000). K-ToBI (Korean ToBI) Labelling Conventions (version 3.1, in November 2000). Unpublished manuscript, University of California at Los Angeles: Department of Linguistics.

Jun, S-A., Beckman, M., & Lee, H. (1998). Fiberscopic evidence for the influence on vowel devoicing of the glottal configurations for Korean obstruents. *UCLA Working Papers in Phonetics, 96*, 43-68.

Junker, Heinrich F. J. (1953). Uber phoneme im Koreanischen. Wiss. Zeitschr. d. Humboldt-Universitat zu Berlin, *Gesellschafts-u. sprachwiss, Reihe* 3, 25-31.

Junker, Heinrich F. J. (1955). *Koreanische Studien*. Abhandlunge d. Deutschen Akademie d, Wissenschaften zu Berlin. Klasse f. Sprachen, Literatur u. Kunst. Jg. Nr. 5. Berlin.

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

Kagaya, Ryohei. (1974). A fiberscopic and acoustic study of the Korean stops, affricates and fricatives. *Journal of Phonetics, 2*, 161-180.

Kenting, P. A., Westbury, J. R., & Stevens, K. N. (1980). Mechanisms of stop-consonant release for different places of articulation. *Journal of the Acoustical Society of America*, *Suppl. 1, 67,* S93.

Keshavarz, M.H. & Ingram, D. (2002). The early phonological development of a Farsi-English bilingual child. *The International Journal of Bilingualism, 6*, 255-269.

Kewely-Port, D. (1983). Time-varying features as correlates of place of articulation in stop consonants. *Journal of Acoustical Society of America, 73*, 322-335.

Kewely-Port, D., Burkle, T.Z., & Lee, J.H. (2007). Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listeners. *Journal of Acoustical Society of America, 122(4)*, 2365-2375.

Kim, C-W. (1965). On the autonomy of the tensity feature in stop classification with special reference to Korean stops. In C-W. Kim (Ed.), *Sojourns in Language: Collected Papers (1988),* 12-33. Seoul, Korea: Tower Press.

Kim, C-W. (1970). A theory of aspiration. *Phonetica, 21*, 107-116.

Kim, H. (1999). The place of articulation of Korean affricates revisited. *Journal of East Asian Linguistics, 8(4)*, 313-347.

Kim, Midam. (2004). Correlation between VOT and F0 in the perception of Korean stops and affricates. *Interspeech, 2004*, 49-52.

Kim, M-R. (1994). Acoustic characteristics of Korean stops and perception of English stop consonants. Unpublished doctoral dissertation, University of Wisconsin, Madison.

Kim, M-R., Beddor, Patrice S., & Horroks, J. (2002). The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics, 30*, 77-100.

Kim, M-R., & Duanmu, San. (2004). “Tense” and “Lax” stops in Korean. *Journal of East Asian Linguistics, 13*, 59-104.

Kim-Renaud, Y-K. (Ed.) (1994). *Theoretical Issues in Korean Linguistics*. Center for the Study of Language and Inf.

Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology, 15*, 166-171.

Kimura, D. (1964). Left-right difference in the perception of melodies. *Quarterly Journal of Experimental Psychology*, *16*, 355-358.

Klatt, D. H., & Klatt, L. D. (1990). Analysis, synthesis and perception of voice quality variations among male and female talkers. *Journal of the Acoustical Society of America, 87*, 820-856.

Kohnert, K & Bates, E. (2002). Balancing bilinguals II: Lexical comprehension and cognitive processing in children learning Spanish-English. *Journal of Speech, Language, and Hearing Research*, *45*, 347-359.

Kolinsky, R., Lidji, P., Peretz, I., Besson, M., & Morais, J. (2009). Processing interactions between phonology and melody: Vowels sing but consonants speak. *Cognition, 112*, 1-20. (auditory examples of the stimuli can be heard on the reference website http://www.brams.umontreal.ca/plab/research/Stimuli/kolinsky_et_al/index.html).

Komaki, R., Akahane-Yamada, R., & Choi, Y. (1999). Effects of native language on the perception of American English /r/ and /l/: Cross-language comparison between Korean and Japanese. *Technical report of IEICE SP 99-45*, 39-46.

Krebs-Lazendic, L. & Best, C.T. (2007). Early and late bilinguals' vowel perception and production: English vowel contrasts that give Serbian-English bilinguals as a H(E)AD-ache. New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech, 282-292.

Kraus, N., Skoe, E., Parbery-Clark, A., & Ashley, R. (2009). Experience-induced malleability in neural encoding of pitch, timbre, and timing: Implications for language and music. *Annals of the New York Academy of Sciences, 1169 (1)*, 543-557.

Kuhl, P.K. (1992). Speech prototypes: Studies on the nature, function, ontogeny and phylogeny of the ‘‘centers’’ of speech categories. In Y. Tohkura, E. Vatikiotis-Bateson, & Y. Sagisaka (Eds.) *Speech Perception, Production and Linguistic Structure*, 239–264. Tokyo: Ohmsha.

Kuhl, P.K. (2000). A new view of language acquisition. *Proceedings of the National Academic of Sciences, 97,* 11850-11857.

Ladefoged, P. (2005). *A Course in Phonetics*. Fifth Edition. Orlando, Florida: Harcourt Brace Jo-vanovish.

Lado, Robert. (1957). *Linguistics across Cultures: Applied Linguistics Language Teachers.* Ann Arbor, MI: University of Michigan Press.

Leadholm, B. & Miller, J. F. (1992). *Language sample analysis: The Wisconsin guide.* Madison, WI: Department of Public Instruction.

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

Liberman, A.M., Delattre, P.C., Cooper, F.S., & Gerstman, L.J. (1954). The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psyhological Monogram, 68*, 1-13.

Liberman, A.M., Ingemann, F., Lisker, L., Delattre, P., & Cooper, F. (1959). Minimal rules for synthesizing speech. *Journal of the Acoustical Society of America, 31*, 1490-1499.

Lindblom, B.E.F. (1986). Phonetic universals in vowel systems. In Ohala, J.J. & Jaeger, J.J. (Eds.), *Experiemental Phonology*. Academic, Orlando.

Lindblom, B.E.F. (1996). Role of articulation in speech perception: Clues from production. *Journal of Acoustical Society of America, 99*, 1683-1692.

Lindblom, B.E.F. & Studdert-Kennedy, M. (1967). On the role of formant transitions in vowel recognition. *Journal of the Acoustical Society of America, 42*, 830-843.

Lindsay, G., Pather, S. & Strand, S. (2006). *Special Educational Needs and Ethnicity: Issues of over- and under-represenation* (Research Report RR757). University of Warwick, department for education and skills.

Lisker, L. & Abramson, A.S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*, 384-422.

Lively, S.E. (1993). An examination of the perceptual magnet effect. *Journal of the Acoustical Society of America, 93*, 2423.

Loban, W. (1976). *Language development : kindergarten through grade twelve*. Urbana, IL: National Council of Teachers of English.

Lotto, A.J., Kluender, K.R., & Holt, L.L. (1998). Depolarizing the perceptual magnet effect. *Journal of the Acoustical Society of America, 103*, 3648-3655.

Lynch, M.P., Eilers, R.E., Oller, D.K., & Urbano, R.C. (1990). Innateness, experience, and music perception. *Psychological Science, 1(4)*, 272-276.

Manis, F.R., McBride-Chang, C., Seidenberg, M.S., Keating, P., Doi, L.M., Munson, B., & Petersen, A. (1997). Are speech perception deficits associated with developmental dyslexia? *Journal of Experimental Child Psychology, 66*, 211–235.

Mayberry, R.I. & Eichen, E.B. (1991). The long-lasting advantage of learning sign language in childhood: Another look at the critical period for language acquisition. *Journal of Memory and Language, 30*, 486-512.

Mayo, Catherine & Turk, Alice. (2001). Is the development of cue weighting strategies in children's speech perception context-dependent? Unpublished manuscript, University of Edinburgh.

McGregor, K., Williams, D., Hearst, S., & Johnson, A. (1997). The use of distinctive analysis in distinguishing difference from disorder: A tutorial. *American Journal of Speech-Language Pathology, 6*, 45-56.

McLaughlin, B. (1978). *Second-language Acquisition in Childhood*. Mahwah, NJ: Ealbaum.

Mennen, I., Stansfield, J., & Johnston, S. (2005). Speech and language therapy services for bilingual children in England and Scotland: A tale of three cities. In J. Cohen, K.T. McAlister, K. Rolstad, & J. MacSwan (Eds.), *Proceedings of the 4th International Symposium on Bilingualism*,1578-1596. Somerville, MA: Cascadilla Press.

Miller, J. & Chapman, R. (1991). SALT: A computer program for the Systematic Analysis of Language Transcripts. University of Wisconsin: Madison, WI.

Miller, G. A. & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America, 34(11)*, 1689-1697.

Minagawa-Kawai, Y., Mori, K., Naoi, N., & Kojima, S. (2006). Neural attunement processes in infants during the acquisition of a language-specific phonemic contrast. *Journal of Neuroscience, 27(2),* 315-321.

Mody, M., Studdert-Kennedy, M., & Brady, S. (1997). Speech perception deficits in poor readers: Auditory processing or phonological coding? *Journal of Experimental Child Psychology, 64*, 199-231.

Morrongiello, B.A., Kulig, W.K., & Clifton, R.K. (1984). Developmental changes in auditory temporal perception. *Child Development, 55*, 461-471.

Morrongiello, B. A., Robson, R.C., Best, C.T., & Clifton, R.K. (1984). Trading relations in the perception of speech by five-year-old children. *Journal of Experimental Child Psychology, 37*, 231-250.

Munson, B., Edwards, J., & Beckman, M. (2005). Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. *Journal of Speech Language and Hearing Research, 48*, 61-78.

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences*, *104(40)*, 15894-15898.

Nearey, T.M. & Assmann, P. (1986). Modeling the role of inherent spectral change in vowel identification. *Journal of the Acoustic Society of America, 80*, 1297-1308.

Neuman , A.C. & Hochberg, I. (1983). Children's perception of speech in reverberation. *Journal of the Acoustical Society of America*, *73*, 2145-2149.

Nish, K., Lewis, D. E., Hoover, B. M., Choi, S, & Stelmachowicz, P. G. (2010). Children’s recognition of American English consonants in noise. *Journal of Acoustical Society of America*, *127*, 3177-3188.

Nittrouer, S. (1992). Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. *Journal of Phonetics, 20*, 1-32.

Nittrouer, S. (1996). Discriminability and perceptual weighting of some acoustic cues to speech perception by 3-year-olds. *Journal of Speech and Hearing Research, 39*, 278-297.

Nittrouer, S. & Crowther, Court S. (1998). Examining the role of auditory sensitivity in the developmental weighting shift. *Journal of Speech and Hearing Research, 41*, 809-818.

Nittrouer, S. & Miller, M.E. (1997). Predicting developmental shifts in perceptual weighting schemes. *Journal of the Acoustical Society of America, 101(4)*, 2253-2266.

Nittrouer, S. & Studdert-Kennedy, M. (1987). The role of coarticulatory effects in the perception of fricatives by children and adults. *Journal of Speech and Hearing Research, 30*, 319-329.

Ohde, R. N. (1984). Fundamental frequency as an acoustic correlate of stop consonant voicing. *Journal of the Acoustic Society of America*, *75(1)*, 224-230.

Oller, D., Eilers, R.E., Rubano, R., & Cobo-Lewis, A.B. (1997). Development of precursors to speech in infants ex[posed to two languages. *Journal of Speech and Hearing Research, 28*, 47-62.

Oller, D. Kimbrough & Ramsdell, Heather L. (2006). A weighted reliability measure for phonetic transcription. *Journal of Speech and Hearing Research, 49(6)*, 1391-1411.

Onishi, K.H., Chambers, K.E., & Fisher, C. (2001). Learning phonotactic constraints from brief auditory experience. *Cognition, 83(1)*, 13-23.

Owren, M.J. & Cardillo, G.C. (2005). The relative roles of vowels and consonants in discriminating talker identity versus word meaning. *Journal of Acoustical Society of America, 119(3)*, 1727-1739.

Owens, R.E. Jr. (2005). *Language Development: An Introduction*. MA: Allyn and Bacon.

Paradis, J. (2001). Do bilingual children have separate phonological systems? *International Journal of Bilingualism, 5(1)*, 19-38.

Patkowski, M.S. (1990). Age and accent in a second language: a reply to James Emil Flege. *Applied Linguistics, 11*, 73-89.

Paul, R. (2007). *Language disorders from infancy through adolescence: Assessment and intervention*. St. Louis, MO: Mosby.

Paulson, D.M. (1991). *Phonological systems of Spanish-speaking Texas preschoolers*. Thesis, Texas Christian University.

Peña, E., Bedore, L.M., & Rappazzo, D. (2003). Comparison of Spanish, English, and bilingual children's performance across semantic tasks. *Language, Speech and Hearing Services in Schools, 34*, 5-16.

Phatak, S.A. & Allen, J.B. (2007). Consonant and vowel confusions in speech-weighted noise. *Journal of the Acoustical Society of America, 121(4)*, 2312 – 2326.

Ploog, D. (1984). Comment on J. Leiber's paper. In R Harre & V. Reynolds (Eds.), *The Meaning of Primate Signals*, 88. Cambridge: Cambridge Univ. Press.

Poole, I (1934). Genetic development of articulation of consonants sounds in speech. *Elementary English Review, 11*, 159–161.

Poole, E. (1934). The Genetic Development of the Articulation of Consonant Sounds, Ph.D. Dissertation, University of Michigan.

Price, P. (1981). *A cross-linguistic study of flaps in Japanese and in American English*. Ph.D. Dissertation, University of Pennsylvania.

Romani, C., Granna, A., & Semenz, C. (1996). More errors on vowels than consonants: an unusual case of conduction aphasia. *Brain Language, 55*, 144-146.

Roseberry-McKibbin, C., Brice, A., & O'Hanlon, L. (2005). Serving English language learners in public school settings: A national survey. *Language, Speech, and Hearing Services in Schools, 36*, 48-61.

Roux, Justus. (2007). Unresolved issues in the representation and phonetic description of click articulation in Xhosa and Zulu. *Language Matters, 38*, 8-25.

Scarcella, R. & Chin, K. (1993). Literacy practices in two Korean-American communities. *Center for Research on Education, Diversity, and Excellence Research Reports, 8*.

Schnitzer, M.L. & Krasinski, E. (1994). The development of segmental phonological in a bilingual child. *Journal of Child Language, 21*, 585-622.

Schnitzer, M.L. & Krasinski, E. (1996). The development of segmental phonological production in bilingual child: A contrasting second case. *Journal of Child Language*, *23*, 547-571.

Schulte-Körne, Demiel, Bartling, & Remschmidt. (1999). The role of phonological awareness, speech perception, and auditory temporal processing for dyslexia. *European Child and Adolescent Psychiatry, 8(3)*, 28-34.

Seymour, H. (1992). The invisible children: A reply to Lahey's perspective. *Journal of Speech and Hearing Research, 35*, 638-641.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, *10(3),* 209-231.

Selinker, L. & Lakshamanan, U. (1992). Language transfer and fossilization: the multiple effects principle. In Gass, S.M. & Selinker, L. (Eds.), *Language transfer in language learning*, 197-216. Amsterdam: John Benjamins.

Shriberg, L.D., Austin, D., Lewis, B.A., McSweeny, J.L., & Wilson, D.L. (1997). The percentage of consonants correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, and Hearing Research, 40*, 708-722.

Shriberg L., & Kwiatkowski, J. (1982). Phonological disorders III: A procedure for assessing severity of involvement. *Journal of Speech and Hearing Disorders, 47*, 256-270.

Silverman, K.E.A. (1987). The structure and processing of fundamental frequency contours. Unpublished doctoral dissertation, University of Cambridge.

Silverman, D. & Jun, J. (1994). Aerodynamic evidence for articulatory overlap in Korean. *Phonetica, 51*, 210-220.

Smit, A., Hand, L., Freilinger, J., Bernthal, J., & Bird, A. (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders 55*, 779–798.

Snow, C. (1983). Literacy and language: Relationships during the preschool years. *Harvard Educational Review, 53*, 165-189.

Spellacy, F. & Blumstein, S. (1970). The influence of language set on ear preference in phoneme recognition. *Cortex*, *6(4)*, 430-439.

Starkweather, C.W. (1987). *Fluency and Stuttering*. Englewood Cliffs, NJ: Prentice Hall.

Steriade, D. (1992). Complex Onsets as Single Segments: The Mazateco Pattern. In J. Cole &. C. Kisseberth, (Eds.), *Perspectives in Phonology*. Stanford: CSLI.

Stockman, Ida J. (1996). The promises and pitfalls of language sample analysis as an assessment tool for linguistic minority children. *Language, Speech and Hearing Services in Schools 27*, 355-366.

Streeter, L. A. & Landauer, T. K. (1976). Effects of learning English as a second language on the acquisition of a new phoneme contrast. *Journal of the Acoustical Society of America, 59*, 448-451.

Stevens, K. N. (1995). Applying phonetic knowledge to lexical access. Proceedings of 4th European Conference on Speech Communication and Technology, 1, Madrid, 3-11.

Stevens, K. N. (1996). Critique: Articulatory-acoustic relations and their role in speech perception. *Journal of the Acoustical Society of America, 99(3)*, 1693-1694.

Stevens, K. N. (1998). *Acoustic Phonetics*. MIT Press, Cambridge.

Stevens, K. N. (2000). Diverse acoustic cues at consonantal landmarks. *Phonetica, 57*, 139-151.

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

Stevens, K. N., & Keyser, J.K. (1989). Primary features and their enhancement in consonants. *Language, 65*, 81-106.

Stoel-Gammon, C. (1985). Phonetic inventories, 15-24 months: A longitudinal study. *Journal of Speech and Hearing Research, 28*, 505-512.

Strange, W., Edman, T.R., & Jenkins, J.J. (1979). Acoustic and phonological factors in vowel identification. *Journal of Experimental Psychology: Human Perception and Performance, 5(4)*, 643-656.

Strange, W., Verbrugge, R.R., Shankweiler, D.P., & Edman, T.R. (1976). Consonant environment specifies vowel identity. *Journal of the Acoustic Society of America, 60*, 213–224.

Sussman, J.E. (1991). Stimulus ratio effects on speech discrimination by children and adults. *Journal of Speech and Hearing Research, 34*, 671-678.

Sussman, J.E. (1993a). Auditory processing in children’s speech perception: Results of selective adaptation and discrimination tasks. *Journal of Speech and Hearing Research, 36*, 380-395.

Sussman, J.E. (1993b). Perception of formant transition cues to place of articulation in children with language impairment. *Journal of Speech and Hearing Research, 36*, 1286-1299.

Sussman, J.E. & Carney, A.E. (1989). Effects of transition length on the perception of stop consonants by children and adults. *Journal of Speech and Hearing Research, 32*, 151-160.

Sussman, J.E. & Lauckner-Morano, V. (1995). Further tests of the 'perceptual magnet effect' in the perception of [i]: Identification and change/no-change discrimination. *Journal of the Acoustic Society of America, 97*, 539-552.

Takagi, N. (1993). *Perception of American English /r/ and /l/ by adult Japanese learners of English: A unified view*. Ph.D. Dissertation, University of California, Irvine.

Tallala, P. & Gaab, N. (2006). Dynamic auditory processing, musical experience and language development. *Trends in Neurosciences, 29(7)*, 382-390.

Taylor, O.L. & Payne, K.T. (1983). Culturally valid testing: A proactive approach. *Topics in Language Disorders, 3*, 1-7.

Templin, M. (1957). Certain language skills in children: Their development and interrelationships. *Institute of Child Welfare Monographs, 26*. Minneapolis, MN: University of Minnesota Press.

Trehub, S.E., Schneider, B.A., Morronglello, B.A., & Thorpe, L.A. (1988). Auditory sensitivity in school-age children. *Journal of Experimental Child Psychology, 46*, 273-285.

Tsukada, K., Birdsong, D., Bialystock, 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*, 263-290.

Tzivinikou, Sotiria. (2004). Developmental speech problems and bilingualism: The difficulties of identification. *International Education Journal, 5(4)*, 466-474.

de Thorne, L.S., Johnson, C.J., Walder, L., & Mahurin-Smith, J. (2009). When "Simon Says" doesn't work: Alternatives to imitation for facilitating early speech development. *American Journal of Speech-Language Pathology*, *18*, 133-145.

van Kleeck, A. (1994). Metalinguistic development. In G. Wallach & J. Butler (Eds.), *Language Learning Disabilities in School-age Children and Adolescents*, 53-101. New York: Macmillan.

Vera, C.S., Messbauer, de Jong, & Peter F. (2003). Word, nonword, and visual paired associate learning in Dutch dyslexic children. *Journal of Experimental Child Psychology,* 84(2), 77-96.

Vihman, M.M. (1996). *Phonological Development: The Origins of Language in the Child*. Oxford: Basil Blackwell.

Vihman, M.M. (2002). Getting started without a system: From phonetics to phonology in bilingual development. *International Journal of Bilingualism, 6(3)*, 239-254.

Walley, A. & Carrell, T. (1983). Onset spectra and formant transitions in the adult's and child's perception of place of articulation in stop consonants. *Journal of the Acoustical Society of America, 73*, 1011-1022.

Wang M.D. & Bigler, R. C. (1973). Consonant confusions in noise: A study of perceptual features. *Journal of the Acoustical Society of America, 54*, 1248- 1266.

Weismer, G., Laures, J.S., Jeng, N.Y., Kent, R.D., & Kent, J.F. (2000). The effect of speaking rate manipulations on acoustic and perceptual aspects of the dysarthria in amyotrophic lateral sclerosis. *Folia Phoniatrica et Logopaedica 52*, 201-219.

Wellman, B., Case, I., Mengert, I., & Bradbury, D. (1931). Speech sounds of young children. University of Iowa study, *Child Welfare, 5*, 1-82.

Werker, J.F. (1995). Exploring developmental changes in crosslanguage speech perception. In D. Osherson, L. Gleitman, & M. Liberman, (Eds.), *An Invitation to Cognitive Science, Part I: Language,* 87–106. MIT Press.

Werker, J.F. & Baldwin (1991). Speech perception and lexical comprehension. Poster presented at the Society for Research in Child Development, Seattle, Washington., April 1991.

Werker, J.F., Gilbert, J.H.V., Humphrey, K. & Tees, R.C. (1981). Developmental aspects of cross-language speech perception. *Child Development, 52*, 349-355.

Werker, J.F. & Lalonde, Chris E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology, 24(5)*, 672-683.

Werker, J.F. & Tees, R.C. (2002). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development, 25*, 121-133.

Weismer, G., Laures, Jacqueline S., Jeng, J-Y., Kent. Ray D., & Kent, Jane F. (2000). Effect of speaking rate manipulations on acoustic and perceptual aspects of the dysarthria in amyotrophic lateral sclerosis. *International Journal of Phoniatrics, Speech Therapy and Communication Pathology, 52(5)*, 201-219.

Whalen, D.H., Abramson, A.S., Lisker, L., & Moody, M. (1990). Gradient effects of fundamental frequency on stop consonant voicing judgments. *Phonetica, 47*, 36-49.

Whalen, D.H., Best, Catherine T., & Irwin, Julia R. (1997). Lexical effects in the perception production of American English /p/ allophones. *Journal of Phonetics, 25*, 501-528.

Williams, G. L. (1974). Speech perception and production as a function of exposure to a second language. Unpublished doctoral dissertation. Harvard University.

Wyatt, T. (1995). Language development in African-American English child speech. *Linguistics and Education, 7(1)*, 7-22.

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

Yansen, E. & Shulman, E. (1996). Language assessment: multicultural considerations. In L. Suzuki, P. Meller, & J. Ponterotto, (Eds.), *The Handbook of Multicultural Assessment*, 353–393. San Francisco: Jossey-Bass.

Yeni-Komshian, Grace., Flege, J., & Liu, Serena. (2000). Pronunciation proficiency in the first and second languages of Korean-English bilinguals. *Bilingualism: Language and Cognition, 3(2)*, 131-149.

Yoon, S-Y. (2009). Nonsense syllable confusion matrices and nonsense syllable imitation tasks in Korean-English bilingual children and adults. Unpublished thesis, University of Illinois at Urbana-Champaign.

# APPENDIX A

## ABBREVIATIONS AND DISTINCTIVE FEATURES

Table A.1

*Abbreviations for Terms in the Present Study*

| Abbreviations | Full names |
|---|---|
| ANOVA | Analysis of Variance |
| AOA | Age of Arrival |
| ASL | American Sign Language |
| BAL | balanced KEB |
| C&I | Complete & Intelligible |
| CV | Consonant-Vowel (structure) |
| DL | Difference Limens |
| E | English-speaking monolingual |
| ES | balanced English-Spanish bilingual |
| ED | English-dominant KEB |
| EM | English-monolingual |
| EVT | Expressive Vocabulary Test |
| F | Familiarity effect |
| FC | Familiar Consonants |
| FV | Familiar Vowels |
| $F_0$ | Fundamental frequency |
| FL | Foreign Language |
| G | Group effect |
| GFTA-2 | the Goldman-Fristoe Test of Articulation-Second Edition |
| IPA | International Phonetic Alphabet |
| JND | Just Noticeable Difference |
| KD | Korean-dominant KEB |
| KEB | Korean-English Bilingual |
| KM | Korean monolingual |
| L | Length effect |

(Table A.1 continued)

| Abbreviations | Full names |
|---|---|
| L1 | First Language |
| L2 | Second Language |
| LOR | Length of Residence |
| LSA | Language Sample Analysis |
| M | Mean |
| MANOVA | Multivariate Analysis of Variance |
| MLU | Mean Length of Utterance |
| NDW | Number of Different Words |
| the NRT | the Nonword Repetition Task |
| the NSCM task | the Nonsense Syllable Matrix Task |
| the NSIT | the Nonsense Syllable Imitation Task |
| NTW | Number of Total Words |
| PAM | Perceptual Assimilation Model |
| PE | Predominantly English-speaking bilingual |
| PPC | Percentage of Phonemes Correct |
| S | Sound type effect |
| SALT | Systematic Analysis of Language Transcripts |
| SD | Standard Deviation |
| SLM | Speech Learning Model |
| SLP | Speech Language Pathology |
| TOT-PPC | Total Percentage of Phonemes Correct |
| UC | Unfamiliar Consonant |
| UV | Unfamiliar Vowel |
| VOT | Voice Onset Time |
| WPM | Words per Minute |

Table A.2

*Terms of Distinctive Features*

| Terms | Definitions |
|---|---|
| Obstruents vs. Sonorants | Obstruents are speech sounds that are made either a total or partial closure of the vocal tract, causing air frictions. Typically stops, fricatives, and affricates are categorized as obstruents.<br>Sonorants are speech sounds that are produced without turbulent airflow in the vocal tract, such as vowels, nasals, glides, and liquids. |
| Anterior vs. Posterior Consonants | Anterior sounds are produced with a constriction at or in front of the alveolar ridge, while posterior sounds with a constriction behind the alveolar ridge. Labials, labio-dentals, and alveolar belong to anterior sounds; however, palatal, velar, and glottal belong to posterior sounds. |
| Apical, Coronal, Dorsal Consonants | Apical consonants are made by using the tip of the tongue (inter-dentals); coronal consonants are made by using the front part of the tongue (alveolars); and dorsal consonants are made by using the mid-body part of the tongue (palatals and velars). |
| Laryngeal Feature (Aspirated and Glottalized) | Laryngeal feature describes the glottal states of sounds. Sounds with vibration of the vocal folds are called voiced, otherwise voiceless. Sounds with the spread of the vocal folds are called aspirated, which are often represented with the symbol ‘$^{h}$.’ Sounds with the closure of the glottis so that air cannot pass through momentarily are called constricted or glottalized. |
| Monophthongs vs. Diphthongs | Monophthongs refer to pure vowels such as /i/ or /e/, whereas diphthongs are vowels that change the quality within the same syllable such as /eɪ/ in ‘*pay*.’ |

# APPENDIX B

# ENGLISH CONSONANT INVENTORY

Table B.1

*English Consonant Inventory*

| | | Place | | | | | |
|---|---|---|---|---|---|---|---|
| Manner | Bilabial | Labio-dental | Inter-dental | Alveolar | Palatal | Velar | Glottal |
| Stop | p, b | | | t, d | | k, g | |
| Fricative | | f, v | θ, ð | s, z | ʃ, ʒ | | h |
| Affricates | | | | | ʧ, ʤ | | |
| Nasal | m | | | n | | ŋ | |
| Glides | w | | | | j | | |
| Liquids | | | | l, r | | | |

*Note*. Voiced consonants are on the right.

# APPENDIX C

## PREDICTIONS OF THE SPEECH LEARNING MODEL AND PERCEPTUAL ASSIMILATION MODEL

Table C.1

*Predictions of SLM and PAM for Perceptual and Production Performance of Experienced and Inexperienced Korean Speakers of L2 English in the Present Study*

| | SLM | | PAM | |
|---|---|---|---|---|
| | easier | harder | easier | harder |
| **Experienced** | | | | |
| Perception | Unfamiliar Sounds<br>b, g, f, θ, z, ʤ, ɪ, æ | Familiar Sounds<br>p, k, s, ʧ, ε | Familiar Sounds<br>p, k, s, ʧ, ε | Unfamiliar Sounds<br>b, g, f, θ, z, ʤ, ɪ, æ |
| Production | Unfamiliar Sounds<br>b, g, f, θ, z, ʤ, ɪ, æ | Familiar Sounds<br>p, k, s, ʧ, ε | Familiar Sounds<br>p, k, s, ʧ, ε | Unfamiliar Sounds<br>b, g, f, θ, z, ʤ, ɪ, æ |
| **Inexperienced** | | | | |
| Perception | Familiar Sounds<br>p, k, s, ʧ, ε | Unfamiliar Sounds<br>b, g, f, θ, z, ʤ, ɪ, æ | Familiar Sounds<br>p, k, s, ʧ, ε | Unfamiliar Sounds<br>b, g, f, θ, z, ʤ, ɪ, æ |
| Production | Familiar Sounds<br>p, k, s, ʧ, ε | Unfamiliar Sounds<br>b, g, f, θ, z, ʤ, ɪ, æ | Familiar Sounds<br>p, k, s, ʧ, ε | Unfamiliar Sounds<br>b, g, f, θ, z, ʤ, ɪ, æ |

*Note*. For *inexperienced* Korean speakers of L2 English, Speech Learning Model (SLM) makes the same predictions as Perceptual Assimilation Model (PAM), i.e., familiar sounds are easier and unfamiliar sounds are harder. Predictions of PAM presented in this table are extended by the author, where PAM focuses on *discrimination* of a pair of two nonnative phones rather than on *identification* of nonnative phones. Therefore, familiar sounds with more of familiar contrasts or features may be easier to identify and less confusing than unfamiliar sounds with more of unfamiliar contrasts.

# APPENDIX D

# FAMILY LANGUAGE PREFERENCE SURVEY

D.1. For a Child Participant

## I. GENERAL INFORMATION

Child

NAME AGE BIRTH DATE BIRTH PLACE

If your child was born outside of the U. S.,

Age of arrival ______ Years in the U.S. ______

Child's Age of arrival Years in the U.S.

Father

NAME AGE BIRTH DATE BIRTH PLACE

HIGHEST DEGREE EARNED OR HIGHEST LEVEL OF EDUCATION REACHED OCCUPATION WORKPLACE (if applicable)

Mother

NAME AGE BIRTH DATE BIRTH PLACE

HIGHEST DEGREE EARNED OR HIGHEST LEVEL OF EDUCATION REACHED OCCUPATION WORKPLACE (if applicable)

Person filling out the questionnaire

PRINT NAME & RELATIONSHIP TO THE CHILD SIGNATURE DATE

*All information remains confidential; it is shared only with our researchers.*

## II. LANGUAGE PREFERENCES

**Farther**

KOREAN

Language Proficiency (please check the box which applies)

| **Speaking** | **Listening** | **Writing** | **Reading** |
|---|---|---|---|
| ☐ Beginner | ☐ Beginner | ☐ Beginner | ☐ Beginner |
| ☐ Intermediate | ☐ Intermediate | ☐ Intermediate | ☐ Intermediate |
| ☐ Advanced | ☐ Advanced | ☐ Advanced | ☐ Advanced |
| ☐ Native | ☐ Native | ☐ Native | ☐ Native |

ENGLISH

Language Proficiency (please check the box which applies)

| **Speaking** | **Listening** | **Writing** | **Reading** |
|---|---|---|---|
| ☐ Beginner | ☐ Beginner | ☐ Beginner | ☐ Beginner |
| ☐ Intermediate | ☐ Intermediate | ☐ Intermediate | ☐ Intermediate |
| ☐ Advanced | ☐ Advanced | ☐ Advanced | ☐ Advanced |
| ☐ Native | ☐ Native | ☐ Native | ☐ Native |

| | | |
|---|---|---|
| Which language do you use more frequently at home? | ☐ Korean | ☐ English |
| Which language do you use more frequently at work? | ☐ Korean | ☐ English |
| Which language do you speak other than the two languages? | | |

**Mother**

| | Language Proficiency (please check the box which applies) | | | |
|---|---|---|---|---|
| KOREAN | **Speaking** | **Listening** | **Writing** | **Reading** |
| | ☐ Beginner<br>☐ Intermediate<br>☐ Advanced<br>☐ Native | ☐ Beginner<br>☐ Intermediate<br>☐ Advanced<br>☐ Native | ☐ Beginner<br>☐ Intermediate<br>☐ Advanced<br>☐ Native | ☐ Beginner<br>☐ Intermediate<br>☐ Advanced<br>☐ Native |
| ENGLISH | Language Proficiency (please check the box which applies) | | | |
| | **Speaking** | **Listening** | **Writing** | **Reading** |
| | ☐ Beginner<br>☐ Intermediate<br>☐ Advanced<br>☐ Native | ☐ Beginner<br>☐ Intermediate<br>☐ Advanced<br>☐ Native | ☐ Beginner<br>☐ Intermediate<br>☐ Advanced<br>☐ Native | ☐ Beginner<br>☐ Intermediate<br>☐ Advanced<br>☐ Native |

| | | |
|---|---|---|
| Which language do you use more frequently at home? | ☐ Korean | ☐ English |
| Which language do you use more frequently at work? | ☐ Korean | ☐ English |
| Which language do you speak other than the two languages? | | |

**Child**

| | Language Proficiency (please check the box which applies) | | | |
|---|---|---|---|---|
| **KOREAN** | **Speaking** | **Listening** | **Writing** | **Reading** |
| | ☐ Beginner<br>☐ Intermediate<br>☐ Advanced<br>☐ Native | ☐ Beginner<br>☐ Intermediate<br>☐ Advanced<br>☐ Native | ☐ Beginner<br>☐ Intermediate<br>☐ Advanced<br>☐ Native | ☐ Beginner<br>☐ Intermediate<br>☐ Advanced<br>☐ Native |

Where does your child use Korean?

(Please check all that apply)

☐ Home ☐ School ☐ Day Care

☐ Play Place ☐ Church/Chapel ☐ Club

☐ After-school Program (please specify what kind of program it is)

: -----------------------------------------------------

☐ Other (please list, if there are other important places where your child speaks English)

: -----------------------------------------------------

| **ENGLISH** | Language Proficiency (please check the box which applies) | | | |
|---|---|---|---|---|
| | **Speaking** | **Listening** | **Writing** | **Reading** |
| | ☐ Beginner<br>☐ Intermediate<br>☐ Advanced<br>☐ Native | ☐ Beginner<br>☐ Intermediate<br>☐ Advanced<br>☐ Native | ☐ Beginner<br>☐ Intermediate<br>☐ Advanced<br>☐ Native | ☐ Beginner<br>☐ Intermediate<br>☐ Advanced<br>☐ Native |
| Where does your child use English?<br><br>(Please check all that apply) | | ☐ Home<br>☐ Play Place<br>☐ After-school Program (please specify what kind of program it is) :<br>---------------------------------------------------<br><br>☐ Other (please list, if there are other important places where your child speaks English)<br>:<br>----------------------------------------------------- | ☐ School<br>☐ Church/Chapel | ☐ Day Care<br>☐ Club |

## III. OTHER INFORMATION

1. Does your child have problems with hearing? ☐ Yes ☐ No

2. Does your child have problems with sleeping or eating? ☐ Yes ☐ No

3. Does your child have any problems with:

| | | |
|---|---|---|
| Mental impairment | ☐ Yes | ☐ No |
| Social-emotional disorder | ☐ Yes | ☐ No |
| Neurological disorder | ☐ Yes | ☐ No |
| Sensory impairment | ☐ Yes | ☐ No |

4. Does your child have allergies? Are there any foods he or she cannot eat or that you would not like them to have during our sessions? If so, please list them below (e.g., peanuts):

________________________________________________________

________________________________________________________

Thank you so much!
We appreciate for you patience and cooperation.

D.2. For an Adult Participant

## I. GENERAL INFORMATION

| NAME | | BIRTH DATE | BIRTH PLACE |
|---|---|---|---|
| | | | |
| AGE | Age or years of arrival in the U.S. | Years in the U.S. | |
| | | | |
| HIGHEST DEGREE EARNED OR HIGHEST LEVEL OF EDUCATION REACHED | | OCCUPATION | WORKPLACE (if applicable) |
| | | | |
| CURRENT ADDRESS | | TELEPHONE | EMAIL ADDRESS |

*All information remains confidential; it is shared only with our researchers.*

-----------------------------------------------------------------------------------------------

## II. LANGUAGE PREFERENCES

**KOREAN**

Language Proficiency (please check the box which applies)

| **Speaking** | **Listening** | **Writing** | **Reading** |
|---|---|---|---|
| ☐ Beginner | ☐ Beginner | ☐ Beginner | ☐ Beginner |
| ☐ Intermediate | ☐ Intermediate | ☐ Intermediate | ☐ Intermediate |
| ☐ Advanced | ☐ Advanced | ☐ Advanced | ☐ Advanced |
| ☐ Native | ☐ Native | ☐ Native | ☐ Native |

Where do you use Korean?

(Please check all that apply)

☐ Home ☐ School ☐ Work
☐ Social clubs ☐ Church/Chapel/Temple
☐ Language Institution
☐ Other

-------------------------------------------------------

**English**

Language Proficiency (please check the box which applies)

| **Speaking** | **Listening** | **Writing** | **Reading** |
|---|---|---|---|
| ☐ Beginner | ☐ Beginner | ☐ Beginner | ☐ Beginner |
| ☐ Intermediate | ☐ Intermediate | ☐ Intermediate | ☐ Intermediate |
| ☐ Advanced | ☐ Advanced | ☐ Advanced | ☐ Advanced |
| ☐ Native | ☐ Native | ☐ Native | ☐ Native |

Where do you use English?

(Please check all that apply)

☐ Home ☐ School ☐ Work
☐ Social clubs ☐ Church/Chapel/Temple
☐ Language Institution
☐ Other

-------------------------------------------------------

**III. EXPERIENCE IN ENGLISH**

1. At what age did you start studying English? ________________________________

2. How long have you studied English including formal instructions at school?
(e.g., including 6 years in middle & high schools, 2 yrs and 4 months in language institute)

________________________________________________________________________

________________________________________________________________________

3. How long have you lived in any country where English is its official language?
(e.g., 2 yrs and 5 mo. in the U.S.)

________________________________________________________________________

4. What was the purpose of visiting a country where English is spoken?
(e.g., studying English; or acquiring degrees; or business/working, etc)

________________________________________________________________________

5. In your country, have you ever been taught by native speakers of English? If so, how long, how often, and where? (e.g., twice a week for 2 yrs in language program)

________________________________________________________________________

**IV. OTHER INFORMATION**

1. Does you have problems with hearing? ☐ Yes ☐ No

2. Does you have problems with sleeping or eating? ☐ Yes ☐ No

3. Does you have any problems with:

| | | |
|---|---|---|
| Mental impairment | ☐ Yes | ☐ No |
| Social-emotional disorder | ☐ Yes | ☐ No |
| Neurological disorder | ☐ Yes | ☐ No |
| Sensory impairment | ☐ Yes | ☐ No |

4. Does you have allergies that we need to know to improve experimental environment? If so, please list them below (e.g., peanuts; dusts):

________________________________________________________________________

________________________________________________________________________

# APPENDIX E

# CONFUSION MATRICES FOR THE 10 CONSONANTS AND THE 3 VOWELS IN THE NONSENSE SYLLABLE CONFUSION MATRIX TASK IN KOREAN-ENGLISH BILINGUAL AND KOREAN MONOLINGUAL CHILDREN AND ADULTS

Table E.1

*Confusion Matrix for the 10 Consonants in All Participants*

| Stimuli ↑ | Responses → | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | p | b | k | g | ʧ | ʤ | s | z | f | θ |
| p | **93** | | | | | | | | 4 | 2 |
| b | 4 | **81** | | | | | | | 6 | 7 |
| k | 1 | | **92** | 2 | 3 | | | | | |
| g | | | | **95** | | | | | | 2 |
| ʧ | | | 2 | | **89** | 4 | 2 | 2 | | 2 |
| ʤ | | | | 6 | 5 | **74** | 1 | 13 | | |
| s | | | | | | | **86** | 2 | 2 | 9 |
| z | | | | 2 | | 17 | 12 | **67** | | 2 |
| f | 8 | 18 | | 1 | 1 | | 5 | | **34** | 33 |
| θ | 5 | 11 | | | | 1 | 8 | 2 | 18 | **56** |

Note. Tables 1 through 13: Scores that are percent of responses less than 1 is not presented; percent correct greater than 1 is rounded. $n = 31$.

Table E.2

*Confusion Matrix for the 10 Consonants in English-dominant Korean-English Bilingual Children*

| Stimuli ↑ | Responses ⟶ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | p | b | k | g | ʧ | ʤ | s | z | f | θ |
| p | **99** | | | | | | | | | |
| b | | **97** | | | | | | | 1 | 2 |
| k | | | **97** | 1 | | | | | | |
| g | | 1 | | **99** | | | | | | |
| ʧ | | | | | **94** | 5 | | | | |
| ʤ | | | | 6 | 1 | **93** | | | | |
| s | | | | | | | **92** | 1 | | 6 |
| z | | | | | | | 6 | **89** | | 4 |
| f | 3 | 21 | | | | | | | **37** | 37 |
| θ | | 11 | | | | | 3 | | 16 | **69** |

*Note*. *n* = 3.

Table E.3

*Confusion Matrix for the 10 Consonants in Balanced Korean-English Bilingual Children*

| Stimuli ↑ | Responses ⟶ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | p | b | k | g | ʧ | ʤ | s | z | f | θ |
| p | **94** | | | | | | | | | 4 |
| b | 5 | **70** | | | | | | | 1 | 21 |
| k | | | **92** | 3 | | | | | | 2 |
| g | | 1 | | **95** | | | | | | |
| ʧ | | | 2 | | **96** | 1 | | | | |
| ʤ | | | | 8 | 6 | **84** | | 2 | | |
| s | | | | | | | **90** | 2 | | 8 |
| z | | | | | | 1 | 6 | **92** | | 1 |
| f | 7 | 12 | 1 | | | | 1 | | **15** | 63 |
| θ | 3 | 5 | | | | | 2 | | 4 | **84** |

*Note*. *n* = 2.

Table E.4

*Confusion Matrix for the 10 Consonants in Korean-dominant Korean-English Bilingual Children*

| Stimuli ↓ / Responses → | p | b | k | g | ʧ | ʤ | s | z | f | θ |
|---|---|---|---|---|---|---|---|---|---|---|
| p | **98** | | | | | | | | | 2 |
| b | 2 | **92** | | | | | | | 3 | 4 |
| k | | | **99** | | | | | | | |
| g | | | | **98** | | | | | | 1 |
| ʧ | | | 1 | | **97** | 2 | | | | |
| ʤ | | | | 2 | 3 | **91** | | 4 | | |
| s | | | | | | | **91** | 2 | | 6 |
| z | | | | | | 6 | 6 | **85** | | 2 |
| f | 3 | 14 | | | | | | | **50** | 31 |
| θ | 1 | 4 | | | | | 2 | | 21 | **71** |

*Note*. *n* = 3.

Table E.5

*Confusion Matrix for the 10 Consonants in Korean-English Bilingual Children*

| Stimuli ↓ / Responses → | p | b | k | g | ʧ | ʤ | s | z | f | θ |
|---|---|---|---|---|---|---|---|---|---|---|
| p | **97** | | | | | | | | | 2 |
| b | 2 | **85** | | | | | | | 2 | 9 |
| k | | | **96** | 1 | 1 | | | | | 1 |
| g | | 1 | | **97** | | | | | | 1 |
| ʧ | | | 1 | | **96** | 4 | | | | |
| ʤ | | | | 6 | 4 | **88** | | 2 | | |
| s | | | | | | | **88** | 2 | | 10 |
| z | | | | | | 2 | 7 | **89** | | 2 |
| f | 4 | 14 | | | | | 1 | | **31** | 47 |
| θ | 1 | 6 | | | | | 3 | 1 | 12 | **75** |

*Note*. *n* = 8.

Table E.6

*Confusion Matrix for the 10 Consonants in Korean-English Bilingual Adults*

| Stimuli ↓ / Responses → | p | b | k | g | ʧ | ʤ | s | z | f | θ |
|---|---|---|---|---|---|---|---|---|---|---|
| p | **97** | | | | | | | | 2 | 2 |
| b | 3 | **90** | | | | | | | 3 | 5 |
| k | 1 | | **97** | 1 | 1 | | | | | 1 |
| g | | | 2 | **97** | | | | | | 1 |
| ʧ | | | 1 | | **95** | 3 | | | | |
| ʤ | | | | 1 | 7 | **79** | 1 | 13 | | |
| s | | | | | | | **85** | 4 | 1 | 9 |
| z | | | | | | 19 | 10 | **70** | | 1 |
| f | 6 | 19 | | | 1 | | 4 | | **37** | 33 |
| θ | 3 | 8 | | | | 1 | 8 | 1 | 19 | **61** |

Note. *n* = 7.

Table E.7

*Confusion Matrix for the 10 Consonants in Korean Monolingual Children*

| Stimuli ↓ / Responses → | p | b | k | g | ʧ | ʤ | s | z | f | θ |
|---|---|---|---|---|---|---|---|---|---|---|
| p | **86** | 1 | 1 | | | 1 | 1 | 1 | 6 | 2 |
| b | 5 | **69** | | 1 | | 1 | 1 | 1 | 14 | 7 |
| k | 1 | | **84** | 3 | 6 | 1 | 1 | 1 | 1 | 1 |
| g | 2 | | 1 | **90** | | 1 | 1 | 1 | 1 | 3 |
| ʧ | 1 | | 2 | 1 | **77** | 6 | 4 | 3 | 2 | 3 |
| ʤ | 1 | | | 9 | 4 | **55** | 2 | 25 | 1 | 2 |
| s | 1 | | | | 1 | 1 | **86** | 1 | 3 | 7 |
| z | 1 | | | 6 | | 29 | 18 | **42** | 1 | 1 |
| f | 11 | 22 | 1 | 2 | 2 | 1 | 8 | 1 | **34** | 18 |
| θ | 7 | 17 | 1 | 1 | 1 | 2 | 13 | 2 | 24 | **33** |

*Note*. *n* = 16.

Table E.8

*Confusion Matrix for the 3 Vowels in All Participants*

| Stimuli ↓ / Response → | æ | ɛ | ɪ |
|---|---|---|---|
| æ | **70** | 27 | 3 |
| ɛ | 33 | **60** | 7 |
| ɪ | 8 | 32 | **59** |

*Note*. *n* = 31.

Table E.9

*Confusion Matrix for the 3 Vowels in English-dominant Korean-English Bilingual Children*

| Stimuli ↓ / Response → | æ | ɛ | ɪ |
|---|---|---|---|
| æ | **74** | 26 | |
| ɛ | 14 | **76** | 10 |
| ɪ | | 12 | **87** |

*Note*. *n* = 3.

Table E.10

*Confusion Matrix for the 3 Vowels in Balanced Korean-English Bilingual Children*

| Stimuli ↓ / Response → | æ | ɛ | ɪ |
|---|---|---|---|
| æ | **74** | 26 | |
| ɛ | 14 | **77** | 9 |
| ɪ | | 12 | **87** |

*Note*. *n* = 2.

Table E.11

*Confusion Matrix for the 3 Vowels in Korean-dominant Korean-English Bilingual Children*

| Stimuli ↓ / Response → | æ | ɛ | ɪ |
|---|---|---|---|
| æ | **78** | 22 | |
| ɛ | 32 | **65** | 3 |
| ɪ | 8 | 47 | **45** |

*Note*. *n* = 3.

Table E.12

*Confusion Matrix for the 3 Vowels in Korean-English Bilingual Children*

| Stimuli ↓ / Response → | æ | ɛ | ɪ |
|---|---|---|---|
| æ | **78** | 21 | 1 |
| ɛ | 22 | **72** | 6 |
| ɪ | 4 | 26 | **70** |

*Note*. *n* = 8.

Table E.13

*Confusion Matrix for the 3 Vowels in Korean-English Bilingual Adults*

| Stimuli ↓ / Response → | æ | ɛ | ɪ |
|---|---|---|---|
| æ | **72** | 28 | |
| ɛ | 33 | **64** | 4 |
| ɪ | 4 | 39 | **56** |

*Note*. *n* = 7.

Table E.14

*Confusion Matrix for the 3 Vowels in Korean Monolingual Children*

| | | Response → | | |
|---|---|---|---|---|
| | | æ | ɛ | ɪ |
| Stimuli → | æ | **59** | 33 | 7 |
| | ɛ | 44 | **46** | 10 |
| | ɪ | 17 | 32 | **51** |

*Note*. $n = 18$.