

Copyright

by

Jin Huh

2007

The Dissertation Committee for Jin Huh certifies that this is the approved  
version of the following dissertation:

**A Comparison Study on the Effects of Two Explicit Pronunciation Syllabi  
on Korean Adult EFL Learners' Learning of English Sounds**

**Committee:**

---

Zena T. Moore, Supervisor

---

Elaine K. Horwitz

---

Diane L. Schallert

---

Daniel H. Robinson

---

Orlando R. Kelm

**A Comparison Study on the Effects of Two Explicit Pronunciation Syllabi  
on Korean Adult EFL Learners' Learning of English Sounds**

**by**

**Jin Huh, B.S., B.A., M.A.**

**Dissertation**

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**Doctor of Philosophy**

**The University of Texas at Austin**

**May, 2007**

## DEDICATION

To my wife Young-Min, who has always been there for me and who  
gives her endless support and encouragement to me.

## ACKNOWLEDGEMENTS

I am deeply grateful to my wife Young-Min Kim, who has always been there for me. And, I'd like to express my thankfulness to my daughters Goom-Young, Jee-Young, and Che-Young for their warm-hearted support and relentless laughter that have been energizing me throughout my life with them. I thank my parents for giving birth to me.

I can never thank my supervisor Dr. Zena Moore too much for her ever-present support for me ever since the moment she offered her insightful inspiration for language education to me in my first course here at the University of Texas at Austin. I would love to extend my deep appreciation to my committee Dr. Elaine Horwitz, Diane Schallert, Dan Robinson, and Orlando Kelm for their enthusiastic care and insightful guidance. I am also greatly indebted to Dr. Scott Myers who instilled in me zeal for teaching pronunciation.

I thank the senior professors in the Department of English at the Korea Military Academy for having provided me with the prestigious opportunity to study here at the University of Texas at Austin and to teach my wonderful cadets at my alma mater.

I regret that I did not have an opportunity to fully express my heartfelt gratitude to Shin-Myung Lee, Jen Gammon, and Joel Hart for their help with the experiment. Lastly, I'd like to give my appreciation to Denise Warburton for aiding me with her exceptional insights while I was finishing up my dissertation.

**A Comparison Study on the Effects of Two Explicit Pronunciation Syllabi  
on Korean Adult EFL Learners' Learning of English Sounds**

Publication No. \_\_\_\_\_

Jin Huh, Ph.D.

The University of Texas at Austin, 2007

Supervisor: Zena T. Moore

This study compared two explicit approaches to teaching English pronunciation to Korean adult EFL learners. One approach involved teaching agendas based on the analysis of pronunciation errors that were considered to be typically observed in the target students. The other approach involved rather general teaching agendas that considered no specific L1 background. The experimental group, the error-analysis-based group, was taught by two Korean instructors of English, while the control group was taught by two native English-speaking instructors. The experimental group teachers used an in-house workbook that consisted of segmental sound items from the clinical data obtained from the error analysis. These Korean instructors of English presented short authentic video clips that the researcher edited from feature movies, situational comedies, news casts, etc. to the learners as model L2

utterances. The native English-speaking teachers used a pre-existing pronunciation workbook published by two native English-speaking language practitioners and presented relevant parts of the supplementary tutorial video to their learners. These four 75-minute pronunciation lectures were conducted as a part of an English intensive course at a tertiary institute in Korea. The two groups did not significantly differ in terms of L2 identification improvements. Based on the real-word and fake-word reading-aloud tests, the experimental group indicated a significantly higher production accuracy rate than the accuracy mean score observed in the control group. Mean accuracy scores of the identification and production tests of each individual sound were compared using Post Hoc ANCOVA techniques for any significant mean differences. Qualitative data from relevant surveys were also included.

## TABLE OF CONTENTS

<b>List of Tables</b> .....	xiii
<b>List of Figures</b> .....	xvi
<b>Chapter 1: Introduction</b> .....	1
<b>Situation</b> .....	1
L2 Pronunciation Difficulty Attributable to L1 Interference .....	3
Difficulty Attributable to Imperfect Interlanguage Decoding Rules .....	6
Difficulty Attributable to Socio-Linguistic Factors .....	9
<b>Possible Solutions Hinted at by Research</b> .....	12
<b>What to Investigate</b> .....	14
<b>Expected Benefits Offered by the Study</b> .....	17
<b>Chapter 2: Review of Literature</b> .....	19
<b>Difficulties That Foreign Accents May Pose</b> .....	19
Intelligibility Problems Caused by Foreign Accents .....	19
Possible Fossilization in EFL Learning Environment .....	22
<b>Considerations of FL Sound System Learning</b> .....	25
Explicit Learning vs. Implicit Acquisition of an FL Sound System .....	25
Segmental vs. Supra-Segmental Aspects of an FL Sound System .....	27
Word- vs. Discourse- Level Aspects of an FL Sound System .....	32
<b>Relevant Speech Perception Issues</b> .....	36
Speech Perception Studies: Differentiation and Identification .....	36
Adult's Abilities to Perceive Sounds .....	38
Perceptual Assimilation Model .....	40



Other Relevant Issues of L2 Speech as Perceived by L2 Learners .....	41
L1 Sound System as the Basis for L2-Sound Perception .....	41
Likelihood of Effective Speech Perception Through Training .....	42
Foreign-Accented Speech as Perceived by Native Speakers .....	43
Effects of Rate of Speech on L2 Utterance Perception .....	46
<b>Revision of Contrastive Analysis Hypothesis .....</b>	<b>50</b>
Overview of the Three Versions of CAH .....	50
Implications of the Moderate CAH .....	53
<b>Age Issues Concerning Second Language Learning .....</b>	<b>59</b>
Adults' Disadvantages Suggested by the Critical Period Hypothesis .....	59
Adults' Disadvantages Caused by Affective Filter .....	65
Adults' Cognitive Superiority .....	66
Adults' High Motivation .....	68
<b>Roles of Non-Native English-Speaking Teachers in the Classroom ....</b>	<b>69</b>
Negative Self-Images of NNESTs .....	69
NESTs' Roles That Tend to Be Unavailable to NNESTs .....	70
Advantages of NNESTs .....	71
<b>Applicability of Socio-Constructivistic Views to Pronunciation Learning .....</b>	<b>73</b>
Orthography as a Social Artifact .....	74
Socio-Constructivistic Review of EFL Learning Environments .....	76
EFL Pronunciation Learning Goals Reconsidered .....	79
Roles of NNESTs as Facilitators of Socio-Constructivistic FL Learning ...	80
Class Grouping and Learning Content Reconsidered Based on SC Suggestions ....	81
<b>Framework for Adult FL Phonology Learning: Noticing Hypothesis ...</b>	<b>82</b>
Emergence of Noticing Hypothesis .....	82
Relevant Claims of Noticing Hypothesis .....	83
Roles of Processing Instruction and Output .....	85
Syllabus Based on Processing Instruction and Output .....	86
<b>Use Of Authentic Video Clips for Teaching Pronunciation .....</b>	<b>89</b>
<b>Use of HTML for In-Class Multimedia Presentations .....</b>	<b>90</b>

<b>Chapter 3: Methodology</b> .....	92
<b>Overview of the Experimental Design</b> .....	92
<b>Participants</b> .....	94
<b>Pronunciation Teaching Approaches</b> .....	96
Textbook for the Control Group .....	96
Textbook for the Experimental Group .....	97
Agendas Based on L1-Interference-Induced Errors .....	99
Agendas Based on Interlanguage-Induced Errors .....	100
Agendas Based on Socio-Culturally-Induced Errors .....	102
Pronunciation Learning Situation of the Current Study .....	103
<b>Data Collection</b> .....	109
Quantitative Data to Address Research Question 1 .....	109
Quantitative Data to Address Research Question 2 .....	111
Quantitative Data to Address Research Question 3 .....	112
Quantitative Data to Address Research Question 4 .....	113
Qualitative Data .....	114
<b>Data Analysis</b> .....	116
Procedure to Analyze Experimental Test Results .....	116
Validity Issues .....	118
Measures to Address Validity Issues .....	119
<b>Chapter 4: Results and Discussions</b> .....	123
<b>EFL Learning Background Survey Results</b> .....	123
<b>L2-Sound Identification Test Results</b> .....	129
Total Score of the Identification Tests .....	130
Improved Identification Accuracy by the Control Group Syllabus .....	131
Improved Identification Accuracy by the Experimental Group Syllabus ...	136
<b>L2-Word Reading-Aloud Test Results</b> .....	140
Validity and Reliability Verified .....	140
Overview of the Reading-Aloud Test Results .....	144

Total Score of the Reading-Aloud Tests.....	152
Achievement Patterns in the Individual Sounds .....	157
Classification of Observed Patterns of Individual Sounds .....	158
Sounds With Significantly Higher Means and Possible Knowledge Transfer .....	159
Sounds With Significantly Higher Means but No Knowledge Transfer .....	162
Sounds With Questionable Gain Patterns .....	163
Sounds With No Significant Mean Differences .....	166
<b>Results of the Language Tests Based on Difficulty Levels .....</b>	<b>168</b>
Relationship Between Identification Results and Difficulty Levels .....	168
Relationship Between Reading-Aloud Results and Difficulty Levels .....	172
<b>Chapter 5: Findings and Pedagogical Implications .....</b>	<b>179</b>
<b>Research Questions Answered .....</b>	<b>179</b>
Improvement in the L2-Sound Identification Ability .....	179
Improvement in the L2-Sound Production Ability .....	181
Role of the Artificial-Word Test in L2-Production Testing .....	182
L2-Sound Producing Skills Transferable to Novel Words .....	184
CAH's Predictions Compared to the Observed L2-Sound Difficulties .....	186
<b>Modifications Considered for CAH-Based Explanations .....</b>	<b>190</b>
Orthography as a Factor of L2-Sound Learning Difficulty .....	190
Phonological Context as a Factor of L2-Sound Learning Difficulty .....	193
<b>Findings Concerning Socio-Constructivist Aspects .....</b>	<b>194</b>
Noticing Approaches Reflected on the Survey Results .....	194
Importance of Instructor's Actual Classroom Conduct .....	197
<b>Findings Concerning Educational Implications .....</b>	<b>198</b>
Utility of EFL Learners' L1 and Interlanguage .....	198
Need for Careful Evaluation of Observed Imperfect Interlanguage .....	199
Possibility of NNESTs' Hypercorrection .....	200
Teacher Education for NNESTs .....	202

Need for Timely Update of Teaching Agendas .....	203
Utility of Explicit Pronunciation Syllabi .....	204
<b>Chapter 6: Conclusions .....</b>	<b>205</b>
<b>Summary of Findings .....</b>	<b>205</b>
<b>Limitations of the Study .....</b>	<b>207</b>
Limitations of the Measurement Instruments .....	207
Limitations in Terms of External Validity .....	209
<b>Suggestions for Future Research .....</b>	<b>210</b>
<b>Appendix .....</b>	<b>212</b>
<b>Bibliography .....</b>	<b>236</b>
<b>Vita .....</b>	<b>251</b>

## LIST OF TABLES

Table 1: Experimental Design .....	94
Table 2: Features of Experimental Syllabi .....	99
Table 3: Reading-Aloud Items on Pre- and Post- Tests .....	114
Table 4: Coding System for Pronunciation Assessment .....	122
Table 5: No. of Semesters of English at Elementary School .....	124
Table 6: No. of Semesters of English Taught by an NEST .....	126
Table 7: Priority of English Skills to Improve .....	127
Table 8: Amount of Time per Day Spent in Using English Skills Before EIC ...	128
Table 9: Amount of Time per Day Spent in Using English Skills During EIC ...	129
Table 10: ANCOVA Source Table for Identification Total Score .....	131
Table 11: Identification ANCOVA Results Showing Higher Improvements in the Control Group .....	134
Table 12: Mean Changes on 'bid' in Each Instructor's Cell .....	135
Table 13: Identification ANCOVA Results Showing Higher Improvements in the Experimental Group .....	137
Table 14: Mean Changes on 'breathe' in Each Instructor's Cell .....	138
Table 15: Mean Changes on 'jealous' in Each Instructor's Cell .....	139
Table 16: Mean Changes on 'zealous' in Each Instructor's Cell .....	140
Table 17: Source Table of Intraclass Correlation Coefficient for Pilot Post-Test Items .....	142
Table 18: Source Table of Intraclass Correlation Coefficient for Pilot Pre-Test Items .....	143

Table 19: Source Table of Test-Retest Reliability (ICC) for Artificial-Word Grading .....	143
Table 20: Source Table of Test-Retest Reliability (ICC) for Real-Word Grading .....	144
Table 21: Analyzed Sounds and Words in the Word Reading-Aloud Tests .....	152
Table 22: Source Table for the Real-Word Reading-Aloud Total Score .....	153
Table 23: Source Table for the Artificial-Word Reading-Aloud Total Score .....	154
Table 24: Descriptive Statistics of the Reading-Aloud Total Score .....	155
Table 25: Estimated Marginal Means of the Reading-Aloud Total Score .....	156
Table 26: Interpretation of Reading-Aloud Test Results .....	160
Table 27: Consistent ANCOVA Results of the Real-Word and Artificial-Word Reading-Aloud Tests .....	161
Table 28: Inconsistent ANCOVA Results of the Real-Word and Artificial-Word Reading-Aloud Tests .....	163
Table 29: ANCOVA Results With Questionable Gain Patterns .....	166
Table 30: ANCOVA Results With No Significant Mean Differences .....	167
Table 31: Descriptive Statistics of the Sounds Showing Significant Group X Pre-Test Interaction .....	168
Table 32: Identification Items Categorized Based on Prator's Levels of Difficulty .....	169
Table 33: Identification Pre-Test Means and Levels of Difficulty .....	170
Table 34: Identification Post-Test Means and Levels of Difficulty .....	171
Table 35: Identification Gain-Score Means and Levels of Difficulty .....	172
Table 36: Reading-Aloud Items Categorized Based on Prator's Level of Difficulty .....	173

Table 37: Mean Scores of the Reading-Aloud Pre-Test Categorized Into Difficulty Levels .....	174
Table 38: Mean Scores of the Reading-Aloud Post-Test Categorized Into Difficulty Levels .....	175
Table 39: Mean Scores of the Reading-Aloud Gain Categorized Into Difficulty Levels .....	177
Table 40: Case Summaries of the Words That Showed Orthography-Sensitive Performance .....	192
Table 41: Case Summaries of the Words That Showed Context-Sensitive Performance .....	194

## LIST OF FIGURES

Figure 1: Estimated Marginal Mean Comparisons of Individual Sounds .....	158
Figure 2: Identification Pre-Score Comparison Based on the Proposed Levels of Hierarchy .....	187
Figure 3: Identification Gain-Score Comparison Based on the Proposed Levels of Hierarchy .....	188



## Chapter 1

### INTRODUCTION

#### SITUATION

In the era of current second language (L2) learning environment where learners are viewed as individuals who need teachers' care-taking in order to achieve optimal learning outcomes, some L2 pronunciation teaching agendas seem to have been organized without considering the typical origins of difficulty that L2 learners face. The consequences this situation may bring forth can be quite serious, especially to EFL (English as a foreign language) learners who have relatively few opportunities to expose themselves to model English utterances.

One type of such phonology acquisition difficulty involves problems caused by differences between the phonological distribution patterns of an L2 and L1 (first language). For the past 50 years, researchers have attempted to shed light on the L2 learners' difficulties caused by dissimilarities between an L1 and an L2 (Lado, 1957; Stockwell & Bowen, 1965; Oller & Ziahosseiny, 1970; Wardhaugh, 1970; Goto, 1971; Acton, 1984; Takagi, 2002). However, it does not appear very hard to find pronunciation teaching agendas that still reveal the lack of understanding of such interlingual differences.

Another major source of difficulty in overcoming foreign accents is the set of false interlanguage decoding rules, probably based on hasty generalizations, established in the mind of EFL learners. One may think decoding skills exclusively concern the area of teaching reading rather than of teaching pronunciation. However, at least for EFL learners, whose first encounter with English was with a textbook in a classroom, rather than with English utterances in a real life outside the classroom, orthography may be one of the most influential factors that affect their L2 pronunciation.

Thirdly, there may be at least one more source of L2 pronunciation difficulty that is the socio-linguistic aspect of specific L2 pronunciation patterns. Speaking a language involves taking on an identity. Korea is known to have been influenced by Confucianism for hundreds of years. In a collective society like the Korean society, Korean adult learners' value system concerning English pronunciation may affect their actual conduct when they speak English. The author of this dissertation found several instances of such a filtering value system, especially in relation to the various realization patterns of the phoneme /t/ as in the word "water" or "Manhattan." In what follows, specific examples of the errors that can be caused by the above-mentioned origins and related pronunciation teaching issues will be briefly talked about.

## L2 PRONUNCIATION DIFFICULTY ATTRIBUTABLE TO L1 INTERFERENCE

It is common to observe EFL learners replace an L2 sound, which their L1 sound system does not have, with an L1 sound that they believe to be identical or closest to the L2 sound. The author of this dissertation has an episode to relate that can clearly show how seriously this undesirable tendency may affect English pronunciation learning. The author of this dissertation started to receive formal English education at school when he was 13 years old. Having spent 13 years in a classroom-based EFL learning environment in Korea, he became an English instructor at a college-level institute. After about three years of teaching English, he went to the U.S.A. for his graduate studies. It is surprising to note that it took as long as 16 years for his habit of replacing /z/ with /dʒ/ in words such as “organization” to be detected by a native speaker of English. Due to this undesirable tendency, which might have been caused by the absence of /z/ in Korean, he had been unable to effectively teach the difference between /z/ and /dʒ/ to his EFL students. His students would have difficulty differentiating the following two sentences thinking that the underlined words are supposed to be pronounced identically as [ˈdʒɛtəs].

I don't know why he is so jealous about it.

I don't know why he is so zealous about it.

Unfortunately, the importance of including such phonology items mentioned above in pronunciation teaching agendas is sometimes not perceived by language

practitioners. One of the English conversation textbooks that were used at the Korean college where the researcher taught provides an example of this unawareness. This textbook was written in English and published in 2006 by one of the most renowned private English institutes in Korea. This book seemed to be intended for both native English-speaking and non-native English instructors. An actual teaching item that focuses on distinguishing a pair of L2 sounds that most target L2 learners generally find easy to differentiate between can be taken as an example of an ineffective pronunciation agenda. In a subsection entitled “Pronunciation” of a unit in the textbook, the following sentences were presented to exemplify the minimal pair of /s/ and /z/.

Six **razors** for \$1. It’s a good **price**.

Six **racers** for \$1,000. It’s a good **prize**.

This pronunciation agenda seemed to be intended for activities that are expected to highlight the presence and absence of voicing in the two alveolar stop sounds [s] and [z], the latter of which is typically problematic to Korean learners. However, even Korean adult learners who are not able to sense the voicing are likely to aurally differentiate between the minimal pairs in the above sentences with ease. The real difficulty arises from their tendency to recognize or pronounce the word “razor” as [rɛɪdʒə] with the Korean palatal affricate [dʒ] in it, instead of [rɛɪzɹ] with the alveolar fricative [z], as explained above. This confusion in aural and oral representations is probably due to the lack of fricative sounds other than [s], [ʃ], and

[h] in the Korean language. In fact, it may be more important to note that Korean adults can experience more difficulties in understanding the differences among the voiced sounds /ʒ/, /z/, and /dʒ/, while they have far fewer problems with their voiceless counterparts [s], [ʃ], and [tʃ]. Therefore, minimal pairs such as “pleasure” vs. “pledger” and “Zale’s” vs. “jails” may constitute a pronunciation agenda that is more helpful for the target learners than the aforementioned one containing the minimal pair of “razor” vs. “racer.”

Some English instructors may not be able to realize the limitations of the above mentioned agenda when they use it to teach adult EFL students. This is because some expertise both in Korean and English phonology is needed in order to understand such interlingual phonology issues. Even native English-speaking teachers (NESTs), who are capable of providing accurate model utterances on their own, are not likely to sense the limitations of aforementioned agenda. Adult students are often prone to fossilized patterns of inaccurate English pronunciation (Acton, 1984), for example, a habit of replacing /z/ with /dʒ/. Korean adult EFL students of this sort would probably remain deaf to the articulation pattern of the fricative /z/ if their teacher explains only about the voicedness of it using the example sentences that contain the minimal pair of “razor” and “racer,” without drawing their attention to the continuous air stream of the /z/.

## DIFFICULTY ATTRIBUTABLE TO IMPERFECT INTERLANGUAGE DECODING RULES

Another origin of L2 pronunciation problems involves undesirable interlanguage rules formed by EFL learners in relation to L2-letter decoding<sup>1</sup>. As widely noted, it is very challenging to generalize English spelling rules, partly because of the orthographic conventions that seem arbitrary as seen in the case of spelling patterns such as “hot” decoded as [hot]<sup>2</sup>, rather than as [hot<sup>2</sup>]. In addition, as can be seen in loan words such as “fiancé” ([fianseɪ]<sup>3</sup>), the various origins of vocabulary can make the learners commit even more errors in pronunciation.

The following episode shows the negative interference of inaccurate decoding rules frequently observed in Korean learners of English. A group of words that contained same vowel [ou], as in “know,” “no,” “sew,” “goat,” “aloe,” “code,” and “cocoa,” were written on the blackboard in an EFL classroom. There were fifteen college freshmen in the class. They said that they were familiar with the seven words due to the formal English classes they had had in Korea for at least six years. Then, they were requested to quietly read these seven different words to themselves in order to count the number of different vowel sounds that the

---

<sup>1</sup> Decoding processes are rarely investigated in the field of phonology because linguists in this field are generally not interested in orthographical representation of sounds. However, the relationships between certain English spelling patterns and their corresponding sounds need to be considered as meaningful to EFL learners who tend to base their pronunciation on the generalizations they make about English sound-spelling relationships.

<sup>2</sup> This pattern is confusing to non-native speakers of English in that the letter ‘o’ is pronounced as [ɑ] in English while, in other languages such as Spanish or German, ‘o’ has a different phonological distribution. For example, the same letter ‘o’ is pronounced as [o] as in the French word “rose.”

<sup>3</sup> The two most typical Korean-accented patterns of the word “fiancé” are expected to be [p<sup>h</sup>ianseɪ] and [p<sup>h</sup>ians].

underlined parts contained. When they were asked about the number of vowels in the group of words, their answers varied from two to five, with most students answering three or four. This decoding confusion is expected to cause intelligibility problems in both oral and aural exchanges. For example, a student who expects [seu] from the word “sew” is highly prone to failure in identifying the word when they *hear* the desirable pronunciation pattern, [sou]. Similar undesirable situations are expected when they *say* the word with such a Korean accent to a native speaker of English.

In the classroom, remedial instructions based on the analyses of student’s errors, not to mention the insights into their imperfect interlanguage decoding rules, are hardly found. In the course of ten years of EFL classes taught by Korean instructors of English, the author of this dissertation received no such explicit tutorials. Also none of the native English-speaking instructors who taught six semesters of his ESL or EFL classes employed this kind of remedial approach. If the English programs provided to him can be considered to be typical, the lack of instructions based on interlanguage decoding rules is likely to be the reason that the EFL students failed to learn the plausible pronunciation patterns of the seven words mentioned above. In fact, Native English-speaking instructors who are not familiar with the characteristics of the native language of the learners (Canagarajah, 1999), such as Korean orthography as compared to English orthography are unlikely to consider orthography or interlanguage decoding rules in order to help their students. This tendency may pose difficulties that could lead to EFL learners’ low

degree of intake of model pronunciation patterns or to the development of fossilized pronunciation errors.

The even more fundamental question, as to whether or not these clinical data concerning the mispronunciation of the seven words are related to different orthographic systems of the L1 and L2, should be first addressed. Fortunately, there are a few studies (Suarez & Meara, 1989; Share, 2004) that deal with the effects of the differences between the two different degrees of regularity that the L1 and L2 orthographic systems in question have. It has been suggested that learners with an L1 that has a relatively regular orthography tend to be less sensitive to exceptions or specifics that are unique to certain L2 words (Share, 2004).

The aforementioned error examples, which involved the words “know,” “no,” “sew,” etc., can be accounted for based on Share’s observation as follows: First it should be noted that the Korean language is believed to have relatively easy-to-generalize relationships between spelling and phonetic rendition. On the other hand, English orthography is considered to be replete with irregularities and exceptions (Kreidler, 1971; Suarez & Meara, 1989; Share, 2004). Therefore, Korean learners may establish hypotheses on the relationships between certain English spelling patterns and their corresponding sounds based on the analogy between their L1 orthography and the L2 orthography. If they mistakenly rely on these hypotheses even when they decode L2 words that involve arbitrary spelling patterns, they are likely to commit pronunciation errors similar to the ones previously illustrated by the seven words.



What aggravates this situation is the EFL environment where there are only few opportunities to encounter model utterances that may help them realize that their L1-influenced interlanguage decoding rules may not systematically work for the L2 system. If this is the case, it may be worthwhile to consider interlanguage decoding patterns of Korean learners of English for a better explanation of their pronunciation learning difficulties.

#### **DIFFICULTY ATTRIBUTABLE TO SOCIO-LINGUISTIC FACTORS**

Some Korean EFL learners tend to intentionally choose a less nativelike pattern of L2 pronunciation based on their value systems. A few instances of socio-culturally affected mispronunciation were found during a pilot study that was conducted prior to the current study. One of the graduate ESL participants in the pilot test seemed unable to pronounce the English flap [ɾ] in the word “water,” although he appeared to have a decent control over segment-level English pronunciation. Even though his pattern of pronouncing this word as [wʌtʰɪ]<sup>4</sup> may be completely intelligible, American English speakers may rarely say the word ‘water’ this way unless they try to emphasize or carefully enunciate the word.

What was intriguing was that the ESL student instantly switched to a more nativelike pronunciation [wʌtɾ] when the researcher asked him to imitate an American English speaker. He revealed that he, at first, thought that he should not say the word in a highly nativelike fashion, even if he was able to do so, because his

---

<sup>4</sup> Some Koreans often replace the English phoneme /ɔ/ with /ʌ/, the latter of which sounds close to /ɿ/ in Korean.

self-perceived English fluency was not very high. One of the other interviewees of the pilot study provided a similar suggestion that Korean learners might hesitate to speak English in a nativelike manner simply because they do not like to appear to be flaunting their fluency.

This kind of trait may bring forth situations where Korean people explicitly recognize highly nativelike patterns of English pronunciation as remarkable if not prestigious when they hear them. On a Korean T.V. show, a group of singers were talking to the host about their career as a musician. The singers were Korean Americans from Los Angeles. Even though their songs were written in Korean, their Korean contained traces of foreign accent because they had only spoken Korean as their heritage language. When they relied on code mixing<sup>5</sup> to say a sentence in Korean, the crowd expressed a sense of awe in a loud voice. The host made a comment, “Your English pronunciation is so authentic. This makes me think you’re really from the States,” even though they were not speaking English. The word that impressed them was “style” ([staɪl]), which was inserted in place of the Korean word “스타일” ([st<sup>h</sup>ail]) because they did not know the Koreanized pronunciation of this loan word.

Sometimes, highly natural American way of pronunciation is perceived as symbolic of identity. A famous Korean comedian was interviewed on a Korean television show right after his visit to the U.S.A. He was one of the greatest

---

<sup>5</sup> Code-mixing means mixing L1 and L2 words in the same sentence.

entertainers in Korea, who had been very poor and had had little education before he distinguished himself on T.V. The hostess asked him where in the States he went. He answered, “맨-햇-은” ([mɛnhaʔŋ]) in Korean, which sounded very close to what would be expected from an American speaker. The hostess echoed what he said, “Oh, 맨하탄 ([mɛnhat<sup>h</sup>an]),” using a Koreanized fashion of pronunciation. Then he intentionally told her in a more exaggerated way, “It is 맨-햇-은” ([mɛnhaʔŋ], not 맨하탄 ([mɛnhat<sup>h</sup>an]). His pronunciation made the hostess and audience laugh aloud maybe because they thought it to be ironic that such a poorly educated person was trying to imitate an American with no shyness. In other words, such a highly authentic pronunciation pattern is regarded as so prestigious that even well-educated English learners would rather not use this pattern when speaking to Koreans for fear of appearing immodest.

These kinds of value-laden reactions may sometimes influence Korean instructors of English in an undesirable way. Concerning the highly nativelike pronunciation of the phoneme /t/ that was discussed earlier, there was a Korean teacher who believed that teachers need to encourage their Korean students to opt to say the fully aspirated allophone of [t<sup>h</sup>] instead of the authentic allophone of flap ([ɾ]) when they say the word “letter.” One of the reasons for his belief was that the English flap was so complex to say that learners with low fluency would rather not learn how to say it and need to choose the aspirated one for clarity. This prejudice against authentic pronunciation patterns seems pointless when we consider the fact

that the Korean language has an allophone of /ɾ/ as in “라면” which is virtually identical to the English flap.

A similar phenomenon was reported by Munro and Derwing (1999). Their study suggests that nativelike L2 pronunciation seems to be unduly underrated, especially in some EFL or ESL classes taught by non-native English-speaking instructors. For some non-native English-speaking teachers, pursuit of a nativelike accent may simply be avoided because they believe clear pronunciation would be enough to suffice. Being unable to sense and produce this natural and realistic pattern, however, may cause communication breakdowns when EFL learners communicate with English native speakers in authentic situations.

#### **POSSIBLE SOLUTIONS HINTED AT BY RESEARCH**

The sources of these difficulties and their examples are not likely to be effectively overcome even with extensive ESL immersion. The pilot study that was previously talked about involved four adult participants, three of which had studied for their graduate degrees in the U.S.A for more than seven months. The fact that even the participant, who had had about two and a half years of immersion experience in the target culture, demonstrated the Korean accents mentioned earlier has led the researcher to wonder what prevented these ESL learners from improving their accents. These participants showed quite a few pronunciation errors that could be regarded as fatal in terms of communication even after they had received 75% of the semester’s English pronunciation lab classes, taught by an English

native speaking instructor. According to the American lab instructor, these Korean adult participants' inability to sense and pronounce some English sounds was repeatedly observed in the three semesters that she had taught the same course.

This chronic insensibility may be accounted for by the following points of view. 1) It may be possible that adult language learners' high affective filters (Krashen, 1981) hinder these post-puberty learners' full immersion in the L2 environment. That is to say, the high language anxiety or low self-esteem in terms of L2 fluency that they probably had might have prevented effective language processing even through extended contact with L2 speakers. Or, 2) they might simply be too old to catch and digest the language input they face everyday (Johnson & Newport, 1989). A third possibility is that their Korean accents were too stabilized to be improved by the model language input they were exposed to while staying in the U.S.A. Foreign-accented pronunciation patterns are likely to be fossilized (Flege, 1981) when the speakers happen to succeed in getting their messages across even with their imperfect pronunciation habits. The participants might have been convinced that their Korean accents in English were workable through repeated instances of successful language tasks that they had conducted in the U.S.A. This explanation is probable because, without request, most native speakers would not explicitly point out or correct a foreigner's imperfect accent once they understand what is being told to them by the foreigners.

Fortunately, there is a hypothesis that might help overcome the modification-proof disposition of English learners that has just been mentioned.

The Noticing Hypothesis, which suggests the importance of conscious noticing for language intake to occur, proposed by Schmidt (1990) appears to be promising. This is even more promising if what the above-observed Korean adult learners lacked was effective intake when they were interacting with English native speakers. Teaching approaches suggested by the Noticing Hypothesis may enhance learners' intake of L2 pronunciation.

In addition to this promising hint of how to teach, there is another hypothesis that could potentially help language practitioners decide on what to teach in order to relieve foreign accents. The contrastive analysis hypothesis (CAH) from the 1970s (Prator, 1971; Lado, 1957; Wardhaugh, 1970 and so on) might work positively for preparing a remedial syllabus to improve EFL learners' foreign accents. This old attempt to shed light on the difficulty of L2 learning is expected to play an important role in the framework for analyzing observed FL pronunciation errors that are related to the L1, when prudently applied. Moreover, the moderate version of the CAH even takes interlanguage rules into consideration. The reflections on these idiosyncratic, but many times systematic, learner-language items are likely to offer pedagogically valuable insights into some of the problems that are beyond the accountability of the strong version of CAH.

#### **WHAT TO INVESTIGATE**

In order to investigate the effects of a pronunciation syllabus based on the analysis of errors as hinted at by relevant research discussed so far, the effects of

another pronunciation approach, which was modeled after the pronunciation lab classes taught by an American instructor, the teacher informant in the researcher's pilot study, was used as the comparison group's syllabus. The reason for choosing this lab course for the comparison with the newly proposed pronunciation approach is as follows:

First, the pronunciation agendas were not based on the typical errors influenced by the learners' L1, Korean, nor considered as having reflected the learners' interlanguage decoding rules. This judgment was based on the review of the textbook for this lab classes and the verbal descriptions of the course provided by the teacher and the participant students in the pilot study. Second, this lab course was the only one that offered explicit pronunciation lecture on a regular basis among the English classes that this dissertation's author had experienced, whether in Korea or in the U.S.A. Lastly, this lab course was taught by a native English-speaking teacher (NEST). That is to say, the two salient aspects of the newly-proposed approach, which are mainly the L1 and interlanguage, are regarded to be tough for NESTs to be familiar with, at least thus far. This is because their L1 is not Korean and most of them are expected to have subconsciously mastered English, which means they rely on their intuition, rather than interlanguage decoding rules. Moreover, it may be NESTs, not non-native English-speaking teachers (NNEST), that many directors of English programs in Korea would hire in order to improve their students' pronunciation, unlike other skills such as reading or grammar. In fact, it has become easier to recruit NESTs now in Korea than it used to be in the

past. In this sense, it can be said that this lab course is *typical* compared to the one based on the analysis of errors.

In most cases, the learner-language-based curriculum proposed in this study is believed to be suitable to be taught by a Korean native-speaking instructor of English, given they have adequate control over L2 pronunciation. This is because they have the same L1 backgrounds as their learners' and many of them share a previous EFL learning experience similar to that of their learners, including the experience with improper interlanguage rules and the socio-linguistic values concerning L2 pronunciation.

To effectively compare these two approaches, the researcher has narrowed the current study's main focus down to segment-level pronunciation items. Patterns of word stress and rhythm in the English language are often regarded as supra-segmental aspects. Nevertheless, to explain allophonic variations including the schwa /ə/, the allophone of unstressed low vowels in English, the concept of word stress and rhythm will be presented. As outlined at the beginning of this chapter, there seem to be quite a few adult EFL learners who have these segment-level problems for various reasons. In order to tackle these problems, this dissertation will identify some patterns in these undesirable tendencies, then explore possible common causes of them, and finally seek a working approach to grappling with them. Then, discussions of the investigated effects that the approach has on the pronunciation patterns of Korean adult learners of English will be presented.



## EXPECTED BENEFITS OFFERED BY THE STUDY

Careful interpretation of the intended comparison results may offer the following implications. First, more customized language placement can be possible if this study can yield some insightful findings from the comparison among different listening comprehension levels. It is not uncommon to find EFL students who are stuck at a plateau in terms of oral and aural proficiency. This undesirable stagnation may result from the discrepancy between their language learning potential and the challenges imposed by the often-overwhelming communicative language activities they have to face. Students who may suffer from this misplacement could benefit more from conversation courses after they fine-tune their fundamental, but crucial, pronunciation skills through the scaffolding remedial curriculum.

Second, analyses of the participants' improvement concerning the highly nativelike patterns of pronunciation are expected to help make decisions as to whether these patterns need to be included in the FL phonology agenda.

Then, closely related to the inclusion of authentic pronunciation, a set of priorities may be determined for certain specific pronunciation items based on the changes in the learners' pronunciation performance in this study. For example, the determination that the minimal pair of "bead" and "bid" should be given higher priority than the priority of the minimal pair of "bait" and "bet" for Korean learners could be made if the latter turns out to be less problematic for the Korean adult

learners of English. If information on the relative degree of such difficulty is gleaned, ESL accent modification programs could be tailored to suit specific L1 backgrounds.

## Chapter 2

### REVIEW OF LITERATURE

#### DIFFICULTIES THAT FOREIGN ACCENTS MAY POSE

##### INTELLIGIBILITY PROBLEMS CAUSED BY FOREIGN ACCENTS

The first, and most fundamental, issue to discuss is the possible negative effects of foreign accents on communication. This issue leads to the question, “Do foreign accents actually hinder communication?” as suggested by the clinical data on Korean accents mentioned in the previous chapter. According to Dieterich et al. (1979), native speakers appear to have tolerance toward the foreign accents in the speech of learners who learn the native speakers’ language as a foreign language (FL). This observation points to the possibility that foreign accents do not actually hinder communication. In addition, it has been found that even perceived foreign accents do not undermine comprehensibility when native speakers of the target language listen to foreigners’ recorded utterances (Munro & Derwing, 1999). Native speakers are generally better at guessing the meaning even when they can only recognize some content words from utterances with a foreign accent (VanDijk & Kintsch, 1983), than are learners of the same language as their L2. For example, let us imagine that a native speaker has just heard an English sentence

from a foreigner. Due to his foreign accent, the native English speaker was able to recognize, say, only 70% of the words in the sentence. As the native speaker is an expert in terms of L1 processing including listening comprehension, he may succeed in deciphering the foreigner's intention based on this limited amount of information, relying on his L1 intuition and various linguistic or meta-linguistic strategies.

Nevertheless, when it comes to the comprehensibility perceived by FL learners when listening to native speakers of their foreign language, the gap between these FL learners and native language (NL) speakers in terms of accents may matter. We can take an imaginary situation similar to the previous one as an example. This time, imagine that a native speaker has just said a sentence (which is perfectly intelligible to another native speaker) to a learner of the native speaker's L1. Even though this sentence is totally acceptable as an authentic utterance, the FL learner might have recognized about 70% of the words in the sentence, possibly because he failed to identify the rest of the words due to his foreign accented pronunciation habits. The 30% information loss may result in failure to guess the meaning intended by the native speaker because the foreigner has no nativelike intuition and is not as good at linguistically or meta-linguistically processing partial information as is the native speaker.

It was suggested by Tatsuki (1999) that the degree to which foreign-accented affect intelligibility when foreigners listen to native speakers' authentic utterances is greater, than the degree to which intelligibility is affected when native speakers listen to foreigners' foreign-accented utterances. In her study, a group of EFL

students were observed while they were watching a feature movie in English. Among several causes of speech misperception reported by Tatsuki, there were “misperceptions based on phonological dialect or *foreign accent differences*” (p. 7). The students commonly experienced listening comprehension breakdowns when they watched a French-accented character talking. The replacement of [ð] with [z] in the words “that” and “there” may have caused the students to miss these words. It should be noted that the French-accented dialog in this film is supposed to be understood generally by adult native speakers of English although they would recognize the character as not depicted as a native speaker of English. On the contrary, it can be argued that the EFL students, say, might not have been fluent or skillful enough to effectively handle the same foreign-accent differences. This phenomenon points to the possibility that the same foreign-accent differences can have a relatively large effect on a FL learner’s listening comprehension but only a hindering effect, on native speakers, that is so small that they may easily guess the meaning.

This potential communication breakdown is also hinted at in Ingram and Park’s (1998) experiment. They found out that identifying certain FL sounds, which is a typical comprehension task that FL listeners need to perform, was the task that was most affected by phonological signal processing and prior phonology learning. In general, EFL learners tend to have a phonological signal handling ability that is less efficient than the one achieved by native speakers. We can infer, therefore, that EFL learners may fail to identify certain FL sounds if they are not

familiar with the nativelike pronunciation patterns of the sounds.

When we talk about oral communication in an L2 between native speakers of the L2 and the L2 learners, we should equally consider the flow of information from the native speakers to the non-native speakers, and the flow of information from the L2 learners to the native speakers. Munro and Derwing (1999) seem to have considered only the latter direction of aural comprehension when they argued that native speakers would have ignorable intelligibility problems when they listen to foreign-accented utterances. Based on the rarely explored aspect of a foreign accent's negative effect on a FL learner's listening comprehension alluded by Tatsuki, it seems hasty to assert that foreign accents would not considerably affect intelligibility.

#### **POSSIBLE FOSSILIZATION IN AN EFL LEARNING ENVIRONMENT**

The next issue to discuss is the difference in the likelihood of fossilization between FL phonology learning situations and SL phonology acquisition situations. According to Flege (1981), even child FL speakers, not to mention adult learners, may end up being satisfied with their less than desirable pronunciation if they succeed in getting their messages across through their phonological translation strategy. For example, the loan word “로펌” ([ʰrɔpʰəm]) from the word “law firm” ([lɔʰ ʰfɪm]) typically illustrates this phonological translation. Out of the five sounds in this combined word, four have been phonologically translated and widely accepted by most Korean speakers who know this word.

Some of these translation patterns may lead to phonological fossilization in Korea. The author of this dissertation had instructed three to four times in one semester that the word “law” should be pronounced as [lɔ] rather than [ro] to a group of Korean college students before he started this study. Many of them remained unable to say the word correctly and failed to identify the same word when they heard it in the middle of a sentence, despite these repeated tutorials. Based on the description of the term fossilization (Vigil & Oller, 1976 in Brown, 1994a), this example can be regarded as an instance of fossilization. The possible reasons for such fossilization are as follows: First, as can be seen in the example of the loan word “로펌” (“law firm”), Korean people do not feel the need to reconsider their Korean accent pattern because “로펌” and its oral representation [ʰroʰpʰəm] have been widely used instead of other possible transcription candidates such as “러펌” ([ʰrʌʰpʰəm]) or “라펌.” ([ʰraʰpʰəm]). Second, interestingly enough, many Korean learners may more effectively identify the word “law” when they hear the Korean-accented pronunciation [ro] or [lɔ] than when they hear the correct one [lɔ].

One may say that there is no effective way to transcribe such words without the help of this phonological translation strategy. Others can say that those loan words should not be considered as English words, but rather as Korean words. These observations are right. However, discussions about such troublesome loanword transcription are not intended here for a transcription reform movement. The purpose of such discussions is to draw language practitioners’ attention to the

possible causes of fossilization so that they can precisely understand the sources of pronunciation problems that their EFL learners tend to have. It is important to note that, although the above-mentioned loan words are not English words, they can be problematic with respect to FL pronunciation because Korean EFL learners tend to rely on the loan word pronunciation patterns when they say the original English words, such as “law” or “law firm” even in English sentences.

Moreover, the fossilized *ESL* pronunciation patterns that Acton (1984) claimed to be modifiable are different in the degree of seriousness from the instances of fossilization typically observed in *EFL* settings. The target participants of Acton’s treatment were foreign professionals that worked in an ESL immersion environment, who appeared to be at least functional, and who wanted to polish their English. On the contrary, most EFL learners are not likely to possess as functional an EFL proficiency as the proficiency that Acton’s ESL professionals possessed, due to definitely insufficient verbal/non-verbal EFL input. That is, EFL learners may suffer more seriously from pronunciation errors than do foreign ESL speakers, who are normally equipped with relatively handy communication strategies and cultural understanding.

In order to help the target Korean EFL learners realize their often fossilized L2 pronunciation patterns, explicit instructions based on concrete examples that are likely to show how such fossilization is caused, were used in the current study. These examples included Korean words loaned from English, foreign movie titles transcribed into Korean, screen captures of web pages that contained problematic



transcription, etc.

## **CONSIDERATIONS OF FL SOUND SYSTEM LEARNING**

### **EXPLICIT LEARNING VS. IMPLICIT ACQUISITION OF AN FL SOUND SYSTEM**

The most fundamental issue of FL sound learning is the need for explicit foreign language (FL) phonology learning as compared to the need for implicit acquisition. One of the most influential proponents of the latter, subconscious attainment, is Krashen. Krashen and Pon (1975) argue that grammatical items that are explicitly learned might only be exploited through monitoring, which requires processing time. This assertion, in turn, means that consciously learned rules may not be retrieved in real-time FL communication situations, as these types of conditions are normally not expected to afford adult FL learners sufficient processing time. Based on Krashen's argument, it seems ironic that several studies attempted to measure the language monitor's efficacy (Ponterotto, 1990; Maun, 1986; Adamson, 1984; Ciske, 1984), hoping that the monitor is a facilitating device.

On the other hand, Maken and Ferguson (1987) argue that the procedure to internalize a phonological distribution may involve "active" processes, rather than "automatic" ones (p. 17). This claim points to the positive role of explicit FL phonology learning. Moreover, it has been proposed that even linguistic items organized as explicit knowledge may be passed on to an implicit linguistic repertoire (Crawford, 1987). This can be possible, according to Crawford, through automatization processes that are implemented by repeated use of the monitor.

These two positions suggest that carefully employing the monitor as the second best option might at least lead to the achievement of automatic, implicit linguistic collection, if the use of the monitor does not instantly benefit authentic communication directly. Acton (1984) argues that “*post hoc* monitoring” (p. 76) provides opportunities for SL learners to recollect their own utterances after they produce them in order to locate errors. Another advantage of this after-action monitoring is that it does not involve negative affective consequences on the spot, for example, communication breakdowns caused by too much self-consciousness.

Jones et al. (1994) point out an additional positive effect of conscious learning strategies. They observed that language learners who had been exposed to linguistic rules tended to have a higher awareness of their own speech and its weaknesses. This observation appears to shed light on helping adult FL learners who often tend to remain unable to realize their problems in terms of their utterances’ intelligibility, despite their extended FL schooling. Therefore, it may be worthwhile to pay attention to Jones’ (1997) suggestion that if linguistic information can help self-improvement, there is no justification for denying such information.

It appears reasonable to be cautious in discarding conscious FL phonology learning for the following reason. As evident in many EFL learning situations such as Korea, EFL settings rarely provide out-of-class FL contact. This unavailability poses difficulty in exposing FL learners to enough comprehensible FL input for unconscious, implicit acquisition of FL phonology. As a result, EFL

learners are inclined to consider the same vowel sound that is spelled differently in a number of words to have more than one sound, as exemplified using the words “know,” “no,” “cocoa,” “aloe,” etc. in the previous chapter. If sufficient outside-the-class language input had been available, they might have known how to properly pronounce the words. As a second resort, explicit instruction, “The letters in the sequences ‘ow,’ ‘oa,’ and ‘oe’ are not individually pronounced as [ou], [oa], and [oe],” is expected to work positively in removing their misconceptions in decoding such letters.

#### SEGMENTAL VS. SUPRA-SEGMENTAL ASPECTS OF AN FL SOUND SYSTEM

Choosing which elements of the foreign language to focus on is another salient issue. One of the hottest debates in teaching FL phonology has been about whether to focus on segmental sounds or on supra-segmental features of the target language. In general, more research recently highlights the importance of prosodic aspects of FL phonology. For example, McNerney and Mendelsohn (1987) and Gilbert (1993) suggest that the FL utterances’ comprehensibility or the clarity of individual sounds relies more on stress, rhythm, and intonation than on segmental accuracy.

However, a survey study among college level ESL<sup>6</sup> students (Derwing & Rossiter, 2002) yielded an observation that is somewhat contrary to the position that

---

<sup>6</sup> Since these participants were attending *college* ESL classes, unlike more genuine ESL students at *elementary* or *middle* schools, they may reasonably be regarded as having acquired a considerable portion of their FL phonology probably in their EFL countries. This is because students who had attended US pre-college schools, probably in immersion settings, are hardly expected to take ESL classes at college.

emphasizes these supra-segmental aspects. According to Derwing and Rossiter's study, only 10% of the participants perceived prosodic deficiency as the cause of their speaking problems, while 84% found difficulty in segment level pronunciation to be their problem. Determination that this survey's results actually reflected what was going on in terms of FL phonology dynamics, of course, needs to be carefully postponed, as this survey is only based on the self-assessment of language learners' oral performance.

Fortunately, Goh (2000) presents a somewhat more convincing argument that favors the importance of segment-level accuracy. Based on her analysis of the transcriptions and listening logs of 40 international students from China in an ESL setting, Goh names, as the second most frequently observed problem, the inability to identify individual words they *hear* as the words they already know *visually*. She points out that incorrect or idiosyncratic phoneme and orthographic representations may be caused by undesirable learning habits such as memorizing the spelling of words without paying attention to their pronunciation.

Especially in cases where English is the target language, the discrepancy between sound and orthography can yield even more complicated difficulties in terms of parsing individual sounds, compared to other FL sound systems. For example, the case of the phoneme /a/ and its "irrational" orthographic representation by the letter 'o' in "hot," not 'a' in "hat," due to the Great Vowel Shift demonstrates

one confusing instance of this sort. Moreover, examples, such as /f/, /ø/<sup>7</sup>, and /g/ as in “tough,” “through,” and “ghost,” in which more than one sound is related to one identical orthographic sequence puzzle English learners.

In the same sense, the schwa /ə/, one of the most frequent sounds in English, may perplex many FL learners who expect one-to-one FL spelling and sound relationships. The schwa represents the most variations in English spelling.<sup>8</sup> The fact that this highly confusing sound was the most improved sound through explicit treatment for EFL learners suggests the importance of discrete learning of segment-level sounds such as activities to differentiate or pronounce them (Kendrick, 1997). In an attempt to help resolve confusion due to such non-one-to-one relationships between spelling and sound, International Phonetic Alphabet (IPA) symbols were created more than one hundred years ago. Nonetheless, EFL learners are still subject to difficulty posed by this discrepancy because most FL dictionaries contain no allophonic transcription, such as the ones concerning the complex one-to-many relationship between the letter ‘t’ and the group of sounds such as the aspirated, unaspirated, and glottalized /t/ ([t<sup>h</sup>] in ‘take’, [t] in ‘steak’, and [t̚] in ‘worked’ respectively) and even the glottal stop [ʔ] in ‘button’ and the flap [ɾ] in ‘entitled.’

Celce-Murcia et al. (1996) offer a recommendation that both segmental and prosodic components be included in a FL agenda. Reminded by this recommendation, it must be noted that respecting prosodic aspects does not

---

<sup>7</sup> /ø/ means silence in this study.

<sup>8</sup> The schwa is the allophone of various unstressed non-low vowels in the North American variety of English.

necessarily mean that segmental elements should be disregarded. What Gilbert (1993) warned against seems to be language students' prevailing preference for *extended drills* of individual sounds and postponement of their attention to supra-segmental features until they *master* individual FL sounds.

It is, however, difficult to imagine that this argument is even intended to exclude *short-term remedial* instruction to help enhance learners' self-awareness of chronically problematic pronunciation habits. Some of the observed problems, such as the ones observed in the pilot study, need only *basal* attention, as opposed to the aforementioned prolonged efforts to *master* FL sounds, because the domain of the problem seems to be as low as the phonics level. These kinds of fundamental clarifications can be metaphorically compared to the "zeroing in" process of shooting a rifle that should have been secured at an earlier stage of the adult learners' FL learning. Successfully produced and understood utterances can be understood to be the equivalent of hitting the bull's eye.

The metaphor of "zeroing in" makes more sense when we consider the rather simple distinctions between clear and unclear or intelligible or unintelligible patterns of pronouncing individual sounds. This is because we normally do not expect a shooter to master "zeroing in," in percentages. We would not, for example, assess "zeroing-in as 75% mastered," but simply check to see if the rifle is zeroed in, with no degree of achievement attached to it. Similarly, determining whether a segment FL sound pronounced by an FL learner is intelligible or not is a relatively simple task compared to assessing the degrees of other kinds of fluency in

higher domains.

The actual reason that a language learner may insist on mastering an individual sound and on postponing working on supra-segmental aspects can perhaps be because the learner has been meddling with irrelevant adjustment points, not because the learner has been gradually polishing his pronunciation of the individual sound. Especially for overly self-conscious learners or adults with a low tolerance for ambiguity, expediting the “zeroing in” process for the segment-level sound may be more effective in steering their attention toward proper supra-segmental features. This expedition might be possible by offering explicit error-based instruction and helping to raise their awareness of delicate segment-level L2 sounds. Without this type of help, the learner may be endlessly waiting for the advent of holistic intelligibility through accumulated unconscious contacts with a myriad of supra-segmental instances. This is because these types of learners can be too obsessed with their oral or aural unintelligibility to reach the equilibrium needed to manage such overflowing input.

It needs to be highlighted that the “all or nothing” approach, whether it is for segmental sounds or prosody, may pose problems, especially when it is extensively employed across a number of FL levels. Discussions of segment or supra-segment level phonology so far appear to point to the need for temporary scaffolding attention to the segmental sound system, prior to extensive language using situations such as conversation classes, which require mastery of language skills involved in higher domains than the segment level.

Among the various areas of segment-level phonology learning, a relatively rare discussion of phonemic or allophonic acquisition appears to be important especially for FL learners, if not SL learners. According to Prator (1971), allophonic aspects, for example English syllabic consonants, may not be ignored because incapability of pronouncing these consonants leads to adding an excrescent vowel, which may cause an unnatural rhythmic variation that affects intelligibility. Because allophonic variations are generally unavailable in FL dictionaries, many EFL learners experience communication breakdowns when they converse informally with native speakers of English who use variations unfamiliar to the EFL learners.

Based on the discussions of the need for segment-level pronunciation learning so far, teaching agendas that focused on individual English sounds and some allophonic distributions of English sounds that were likely to be confusing for Korean EFL learners have been organized. These agendas included, for example, English phonemes that did not exist in Korean such as /v/, /f/, and /ɪ/, various allophonic realizations of English /t/, changes in actual vowel qualities depending on word stress patterns, and so on.

#### **WORD- VS. DISCOURSE- LEVEL ASPECTS OF AN FL SOUND SYSTEM**

The argument on whether to emphasize word-level pronunciation or sentence/discourse level pronunciation has also drawn language practitioners' attention. Kendrick (1997) has found that word-level accuracy often does not



guarantee the same accuracy for a longer or a more contextualized discourse such as free conversations or story telling. This lack of transferability of achieved accuracy can be substantially troublesome because most language learners aim at being able to communicate by uttering whole sentences, not by speaking only a single word per turn.

However, it needs to be noted that the lack of transferability does not necessarily mean that the word-level competence achieved through practice suddenly disappears in the sentence/discourse pronunciation domain. This higher domain might simply need more linguistic dexterity than what is required for word-level accuracy. That is, longer and more contextualized discourses additionally require simultaneous processing of multiple tasks other than enunciating FL words. Discarding approaches to enhance word-level accuracy may be as faulty, for example, as avoiding the first step of taking medication, worrying that this first type of medication may not at once completely cure a disease that needs a *series* of medical treatments.

Furthermore, there seems to be one more rationale for not dispensing with word-level practice for accuracy. Researchers like Field (2004) pointed to the need for including word-level phonology, arguing that difficulty in word recognition causes low level L2 students to have an undesirable tendency to rely heavily on top-down listening strategies. On the contrary, there is also a notion that lower-level listeners tend to have problems in affording sufficient spare cognitive or temporal resources to handle higher-level units of meaning because they become too busy

concentrating on processing individual sounds and words (Shohamy & Inbar, 1988). As he also stated in the same paper, Field's argument may seem contradictory to this more widely acknowledged notion suggested by Shohamy & Inbar (1988).

A close look into Tsui and Fullilove's study (1998) on reading and listening comprehension, nevertheless, suggests that Field's argument may also be valid. They hypothesized that learners with relatively low skills may devote most of their cognitive resources to top-down schemata processing rather than to bottom-up linguistic processing, *unlike* the opposing view of Shohamy and Inbar. Tsui and Fullilove's hypothesis appears to be plausible because the less skilled learners obtained significantly lower scores on items that involved new information incompatible with existing schemata. These types of questions required them to pay attention to and interpret new linguistic input that was contradictory to their previously activated schemata formed through top-down processing, for example, decoding contextual cues. Their failure may have been caused by relying too much on top-down processing that consequently led them to incorrect guessing with little attention paid to the new *contradictory* linguistic input.

The fact that the same group of low skilled learners could achieve higher scores when they were faced with corresponding-schema items suggests that even low-level learners opt to rely on top-down processing, probably in order to compensate for their weak bottom-up processing abilities.

The three points of view concerning word-level bottom-up processing

discussed so far can be understood as follows: The group of low-level students in Shohamy and Inbar's study and Field's study might have had dissimilar dispositions in terms of linguistic processing. It is possible that these two groups of students were only different in that one group adopted the strategy of "guessing" made possible by top-down processing, probably, because they had realized that relying merely on bottom-up processing was not very effective. On the other hand, Shohamy and Inbar's students may well have been still struggling with information provided by individual words hoping that they would become efficient enough someday at decoding them without employing any top-down approaches. In this sense, Tsui and Fullilove's study (1998) can be said to have depicted the low-level students' problems using top-down strategy, like the ones in Field's study, still had to face.

In short, approaches that purport to improve word-level pronunciation accuracy are likely to be beneficial for low-level FL listeners. Naturally, enhanced individual-word recognition is expected to help FL learners who mainly resort to individual word comprehension. In addition, improved word-level pronunciation accuracy can also help top-down strategy students properly identify a new word even if its information seems incompatible with an existing schema. Furthermore, this may lead to balanced reliance on bottom-up aural processing and top-down schema-based processing. This desirable balance, in turn, could yield abilities to more accurately and expansively understand utterances exchanged in real-time aural/oral communication. Nevertheless, it appears to be hard to expect word-level

intelligibility to be automatically transferred to a higher domain such as the domain of sentences, maybe because of the cognitive overloads involved with the increased needs for simultaneous cue processing.

Therefore, the scope of pronunciation teaching agendas for the current study excluded issues such as sentential stress or intonation patterns. On the other hand, some issues that involve considerations beyond the word-level such as linking patterns between word boundaries (“first stop”) and ellipsis phenomena (“hand towel”) were incorporated as the same phenomena were considered as occurring within a word’s boundaries.

#### **RELEVANT SPEECH PERCEPTION ISSUES**

Even though speech perception may not seem directly related to FL pronunciation, speech perception issues are worth discussing because pronunciation generally affects the listener’s perception of the utterances.

#### **SPEECH PERCEPTION STUDIES: DIFFERENTIATION AND IDENTIFICATION**

Laboratory-based attempts to make FL learners improve L2 sound perception have a long history (Yeni-Komshian, 1968; Garnes, 1977; Sharf et. al., 1988; Best, McRoberts, & Goodell, 2001; Takagi, 2002; Levey, 2004). Most of these studies involve behavioral approaches used to *train* the subjects rather than *teach* them. Therefore, training sessions of this sort mostly consisted of linguistic stimuli and feedback based on the trainees’ responses to the stimuli. It is this

feedback that is believed to change the trainees' perceptual behavior, specifically differentiation and/or identification of certain unfamiliar phonological contrasts. Some of the studies only investigate the trainees' perceptual changes in their responses to the given stimuli over the course of the training sessions. Other studies also measure the degree of transfer in terms of the training effects on the trainees' changes in their differentiation or identification abilities by testing the same trainees with another set of FL words that had not been used for the stimuli.

A typical way of testing FL learners' discrimination abilities is to have them listen to stimuli that contain three tokens of the intended FL sound contrasts. In order to control for the possibility that testees can determine the stimuli simply by paying attention to acoustic identity, the order of stimuli is often manipulated. For example, a series of tokens "AAB," "ABA," "BBA," or "BAA" are presented and the testee are requested to choose the token, the first or the last, that is perceived to be the same as the middle token (Best, McRoberts, & Goodell, 2001).

One of the strengths of these laboratory-based perception studies is that, as factors other than a few interested causes were strictly excluded, they offered insights into the causal effects that change the perceptual behavior of the trainees. For example, Levey's (2004) study suggests that English vowels that the Spanish language does not have, cause the most difficulty in English/Spanish bilinguals' discrimination of the vowels.

Identification studies employ testing methods that typically involve one

stimulus at a time. Testees are supposed to push one of the two buttons before them after listening to the stimulus. Identification tasks are different from discrimination tasks in that the former need FL learners to have internalized certain constructs of the target sounds in their minds in order to compare the construct with the stimulus and make identification judgments. In discrimination tests, testees can rely on their short-term memory to compare the given stimuli to each other because they are given in a sequence.

#### **ADULT'S ABILITIES TO PERCEIVE SOUNDS**

Compared to infants' universal perception capabilities, adults are known to possess abilities to perceptually discriminate only phonetic distinctions that are reflected on their native language phonemic systems (Werker & Lalonde, 1988, Werker & Tees, 1984a; Werker & Tees, 1984b). Based on these studies, by the end of the first year of their lives, infants seem to finish restructuring their perception dispositions in such a way that it resembles the typical limited dispositions of adults. Thus, adults often tend to categorize FL sound input based on their restructured sound mapping system.

The adult tendency to process and identify FL sounds using L1 phonemic classification is likely to cause failure to sense the phonological contrasts between certain L2 phonemes if these phonemes do not exist as phonemes in the L1. Accordingly, this insensitivity frequently prevents adult FL learners from realizing the need to polish their L2 pronunciation, as shown in the clinical data concerning

the errors due to phonological translation of [z], [dʒ], and [ʒ] into [dʒ] in Chapter 1.

If this is the case across all the stages of L2 learning, an adult's FL pronunciation is doomed to failure especially when the L1 and L2 have little in common in terms of phonemic distribution. This is because, no matter how much aural input they are exposed to, they tend to filter the input through their L1 phonemic awareness and end up processing limited phonological information.

However, there is some research that claims that even adults can make the use of such universal discriminating sensitivity they used to possess when they were infants. According to the results of Morosan & Jamieson's (1989) laboratory discrimination training, adult monolingual speakers were able to differentiate certain L2 contrasts in the stimuli. They even succeeded in transferring their differentiation ability to tell the difference between unfamiliar L2 words. Some other studies (Pisoni et al., 1982; Werker & Tees, 1984b; MacKain, Best, & Strange, 1981) also claim that adults can recover this universal ability through training or extensive L2 exposure. Even though these studies suggest optimistic learning outcomes in only one of the two oral communication aspects, perception and production, such enhanced sensitivity is expected to benefit adult L2 learners by helping them extract self-teaching agendas from FL input. In what follows, one important notion in the area of speech perception, such as identification or discrimination, will be discussed even though laboratory-based behavioral training was not incorporated in the teaching agendas of the current study.

## **PERCEPTUAL ASSIMILATION MODEL**

An important framework for the understanding of L2 speech perception has been provided by one laboratory-based perception study. Based on Best, McRoberts, & Goodell (2001), the Perceptual Assimilation Model (PAM), which had been previously proposed by Best (1995), turns out to have provided predictions that explain the English speakers' discriminating patterns of Zulu and Tigrinya contrasts. According to PAM, there are three different degrees of difficulty involved in discriminating FL contrasts. First, certain FL contrasts are best discriminated if these contrasts are perceived as identical to a set of contrasts in the learners' native language. The second easiest level of discrimination is expected where one of the FL contrasts can be identified as a native language sound while the other FL contrast can be considered a bad example of the same native sound. Thirdly, FL sound discrimination involves most difficulty if certain FL contrasts are perceived by FL learners as equivalents to a single native segment. A framework for classifying FL sound difficulty that is similar to these predictions can be found in Prator's Hierarchy of Difficulty discussed later in this chapter.

This PAM framework played a role in the basis used to sort out the Korean-accented English example words when the author of this dissertation organized the pronunciation teaching agendas for this experiment. In fact, facilitating accurate discrimination abilities was not the ultimate goal of the pronunciation teaching syllabus for this study. However, teaching items for FL sound discrimination were thought to be a good starting point and were expected to enhance the target students'



sensitivity to FL contrasts that would otherwise remain confusing to them. The items classified using the PAM predictions were actually incorporated into a section called “Eye Openers” in the student workbook for the experimental group (see Chapter 3 for details).

#### **OTHER RELEVANT ISSUES OF L2 SPEECH AS PERCEIVED BY L2 LEARNERS**

Among the many perception studies that investigated patterns of FL sound identification, the contrast in English phonemes /ɪ/ and /I/ perceived by Japanese English learners has been one of the most frequently investigated FL contrasts (Yamada, 1993; Takagi & Mann, 1995; Takagi, 2002; Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997).

**L1 SOUND SYSTEM AS THE BASIS FOR L2-SOUND PERCEPTION:** It is possible that FL learners base the judgment they use to perceive certain L2 contrasts on the similarity that these contrasts bear to the closest phoneme of their native languages. Upon the request to describe the difference between the /ɪ/ and /I/ contrasts (Takagi, 2002), the Japanese subjects explained it in relation to the relative proximity that the two FL contrasts were considered to take to Japanese /ɾ/. Specifically, four out of the five subjects believed that Japanese /ɾ/ was closer to English /I/<sup>9</sup> than it was to /ɪ/. It can be said that Japanese learners of English use a native sound as a reference point when they identify words that contain certain

---

<sup>9</sup> In Takagi’s (2002) study, English [ʃ], one of the allophones of English /I/, was excluded.

confusing English contrasts. Interestingly enough, there is a study that suggests even child learners rely on their L1 sound system to acquire L2 phonology (Anderson, 2004). If this is the case, the L1 sound system should not be considered as undermining the effectiveness of FL sound perception.

#### LIKELIHOOD OF EFFECTIVE SPEECH PERCEPTION THROUGH TRAINING:

Laboratory-based identification training has been reported to be ineffective (Takagi, 1995; Takagi, 2002; Sharf, 1998). According to Takagi, making Japanese learners of English able to identify English /ɹ/ and /l/ in a truly nativelike fashion seems virtually impossible. For Japanese learners of English, in particular, even more than 12 years of extensive everyday exposure to the target language turned out to be fruitless (Takagi & Mann, 1995).

This “dead-end” conclusion is not very surprising when we consider the aforementioned Perceptual Assimilation Model. Despite the fact that the majority of the Japanese subjects in Takagi (2002) perceived English /l/ as closer to Japanese /r/, it would be too hasty to conclude that they are considered identical to each other in the subjects’ mind. Then, the next possibility is that Japanese speakers believe that both phonemes are identical, or at least quasi-identical, to /r/, probably classifying both contrasts into the same category of the only similar native sound, /r/. If this is the case, perceiving these English contrasts would be one of the hardest tasks to Japanese native speakers based on PAM predictions.

Another explanation is possible when taken in relation to the training

approach used in the identification training studies. In the same study, where Takagi doubted the possibility of the attainment of nativelike /ɹ/ and /l/ identification, an implicit training approach based on a simple stimulus-response scheme was used. That is, when the subjects pushed the wrong button after hearing a stimulus, they were provided with the right answer followed by the same stimulus in repetition. The lack of explicit instruction might have been the cause of such ineffective training outcomes.

Fortunately, according to PAM predictions, Koreans' identification of the same English phonemes would not be as difficult as it is for Japanese native speakers. Unlike the Japanese language, Korean has [l], one of the allophones of English /l/. Therefore, one of the two target contrasts can be easily identified as the Korean [l] by most Korean learners of English. This leads to the expectation that the Korean case pertains to the second toughest perception task based on PAM's classification. However, when it comes to identifying English [ɹ] and [ɻ], the latter of which is another allophone of English /l/, the case may be different. Because Korean has neither of the two contrasts, Korean learners of English would experience a level of difficulty similar to that of Japanese learners of English.

#### **FOREIGN-ACCENTED SPEECH AS PERCEIVED BY NATIVE SPEAKERS**

It is worthwhile to discuss findings on the perception issues of foreign-accented speech as this aspect of FL learners' utterances is more likely to be affected by their pronunciation accuracy. First of all, it has been generally noted that

foreign accents cause negative influence on language processing when native speakers of the target language listen to FL speakers (Johnson, 1997; Clarke, 2000; Clarke & Garrett, 2004; Floccia et al., 2006). Based on Clarke & Garrett (2004), foreign-accented speech turned out to cause delays of 100 to 150ms, which were three to five times longer than delays caused by regional accents. It has also been agreed upon that native speakers adapt themselves to such foreign accents through repeated exposure to them.

Another important issue concerning the processing of foreign accents has been centered on whether adaptation to foreign accents takes place across different speakers with the same foreign accent. Research presents different suggestions on this matter. Clarke & Garrett (2004) presents findings that point to the possibility that native speakers become better foreign-accent processor after exposure to utterances from different speakers who have the same foreign accent. On the contrary, Floccia et al. (2006) and Clarke (2000) are dubious about the existence of the cross-speaker transfer of adaptation. In this sense, the thin argument for native speakers' adaptation to foreign accents would hardly be an excuse for assigning a low priority to accent reduction agendas.

Moreover, foreign-accented speech may be more difficult for native speakers to understand when it is perceived in noisy environment. Munro (1998) reported that cafeteria noise caused a greater decrease in the intelligibility of Mandarin-accented English utterances than in the decrease of native English utterances in the same environment. This observation is worth considering as a noise-proof

environment is unavailable in most daily-life oral communication situations.

There is one more possible disadvantage of speaking with foreign accent. Based on Horani's (1995) report, foreign accents may undermine the speakers' professional competence as perceived of by native speakers. A Japanese physician was rated by native English-speaking nurses as possessing the lowest level of medical competence as compared to the competence of American and Persian physicians. It is interesting to note that this perceived deduction in the physician's efficacy occurred when the nurses heard his speech in a formal context rather than in an informal context. If this finding can be generalized to other professions and speakers of English with different L1 backgrounds, accent-modification programs should be included in FL agendas.

One of the studies on perceived foreign accents has hinted at the feasibility of organizing accent-modification agendas based on a specific native language. According to Major (2002), native Spanish speakers' lecturing in English yielded the highest comprehension scores when other native Spanish speakers listened to it, while the lowest score was obtained in the cases of Chinese-accented lecturing as perceived by native Chinese speakers. It should be noted that the comprehension scores obtained from the case of Spanish-accented lecture perceived by native Speakers of Spanish turned out to be higher than the case of Spanish-accented lecture perceived by native English speakers. This result suggests that FL learners with the same native language share similar foreign accent patterns. If PAM predictions are applied to this observation, this case pertains to the easiest

perception situation. That is to say, the lecturer and the listeners may have shared the same Spanish-influenced English contrasts and the listeners thus might have had little trouble in matching each of the FL contrasts spoken by the lecturer with a common FL phoneme available in the listeners' minds. This means these FL learners' interlanguage, which may otherwise be considered highly arbitrary, can be generalized, to a degree, for other FL learners who have the same L1 background.

Based on the discussions so far, it can be said that language practitioners would better not hope for native speakers to be able to effectively understand or adjust to foreign accents. In other words, if there is a gap between native speakers and FL learners in terms of accentedness, it should be FL learners that need to seek ways to close the gap. One of the gap-filling approaches is to have FL learners become familiar with nativelike patterns of pronunciation rather than merely clear pronunciation patterns that some language teachers consider as an imperfect, but safe oral gadget.

#### **EFFECTS OF RATE OF SPEECH ON L2 UTTERANCE PERCEPTION**

In addition to segmental errors (Munro & Derwing, 1999), L2 learners' rate of speech has been known to influence the overall foreign-accentedness and comprehensibility perceived of by native speakers. A number of studies have investigated the effects of different rates of speech on perception (Munro & Derwing, 2001; Anderson-Hsieh & Dauer, 1997; Anderson-Hsieh & Koehler, 1988; Zhao, 1997).

The first question that should be asked is whether foreign language learners actually tend to speak their L2 slowly. Guion et al. (2000) investigated the relationship between age of initial L2 learning and rate of L2 speech. In particular, for the Korean learners in the study, it turned out that the factor of age at which they arrived in America was the most salient factor influencing the duration of their English utterances. Therefore, it may be said that Korean EFL learners tend to speak English more slowly than Korean ESL learners, who are younger and have been exposed to ESL utterances earlier in their lives.

As a result, studies that purported to explore the effects of slow L2 utterances on how native speakers react to them are worth mentioning. Based on Anderson-Hsieh & Koehler's (1988) findings, native speakers tend to rate FL learners as having stranger foreign accents and being harder to comprehend when the native speakers hear the FL learners' L2 utterances at a high rate of speed. This observation appears to be the case if we consider the aforementioned finding (Clarke & Garrett, 2004) that foreign-accented speech can cause native speakers to need longer processing times. Furthermore, this predicted disadvantage of high-rate L2 utterances conforms to Zhao's (1997) finding that slowed-down model L2 utterances are likely to be considered as easier for *L2 learners* to comprehend than L2 sentences uttered by *native speakers* at a normal speed.

However, some other studies point to more a complex relationship between the rate of L2 speech and native speakers' perception. Munro & Derwing (1998) revealed that the native English-speaking raters felt comfortable when they heard

Mandarin-accented English utterances that were delivered at a slightly faster speed than the original rate of speech chosen by the Mandarin learners of English. Munro & Derwing's (2001) later investigation also confirms this intriguing observation. They carefully hypothesize that the optimal rate of speech is reached when FL learners speak the L2 at a speed that is slightly higher than the speed at which they normally do and slightly lower than the native speakers' normal speech. That is to say, either very fast or very slow speech can be problematic in terms of perceived foreign-accentedness and comprehensibility. The basis for this counter-intuitive prediction is the possibility that native speakers need to store previously heard information in temporary memory for a relatively longer time and wait for the next information to be uttered as they listen to unduly slow speech (Munro & Derwing, 2001). The problem occurs when the amount of time during which the information should be kept is considerably greater for prolonged FL speech than the length of time needed for processing normal speed speech. If this happens, it may require additional cognitive capacity and cause difficulty in comprehension.

It should be noted that, in the same study, Munro and Derwing warn against the decisions to advise FL learners that they should increase or decrease the rate of their FL speech. This warning is based on the possibility that other factors, such as prosody, segment-level accuracy, or discourse-wise performance, may have interaction with the rate of speech. In fact, many FL learners are unable to talk considerably faster without inflating the chance of slips or segment errors more than they ordinarily would.



Nevertheless, it appears to be beneficial for comprehensibility to help FL learners avoid undesirably slowing down their speech as long as prosody or segment pronunciation performance is maintained. This possibility has been hinted at by the aforementioned fact that native speakers favored slightly higher-rated speech samples that were electronically compressed (Munro & Derwing, 1998). More specifically, the digitally accelerated speech is expected to have maintained the same segment-level and prosody accuracy as the accuracy of the sample originally involved.

This speculation led the author of the current study to the inclusion of authentic speech models into the pronunciation teaching agendas. An important characteristic of authentic video clips, such as parts of movies, T.V. commercials, or situation comedies, is the relatively high rate of the recorded speech, compared to typical pronunciation samples, for example, from electronic dictionaries or careful native English-speaking teachers. By exposing the intended EFL learners to such high-speed utterances and explicitly instructing them how to aurally/orally familiarize themselves with the authentic models, the current study's pronunciation syllabus is expected to aid them in enhancing segment-level intelligibility. As the learners will be trained with high-rate speech models, this enhanced intelligibility, once secured, is more likely to make them perform stably at a high speech rate.

## REVISION OF CONTRASTIVE ANALYSIS HYPOTHESIS

The contrastive analysis hypothesis (CAH), which dates back over 40 years, deserves some attention. This overly ambitious linguistic attempt to theorize the difficulties of L2 learning drew a number of language practitioners' attention back in the 1960s. Language teachers may criticize that CAH is spotted with positivist perspectives based on mechanistic and behavioral analyses. Nevertheless, it is worthwhile to reconsider CAH for this study because this hypothesis can provide one of the frameworks in which to analyze observed FL errors. This working framework is expected to allow for inter-lingual insights into typical EFL pronunciation problems. In what follows, an overview of the CAH and its implications for the current study will be presented.

### OVERVIEW OF THE THREE VERSIONS OF CAH

This hypothesis can be classified into three versions: the strong, weak, and moderate versions. Lado (1957) claims that the ease or difficulty of learning a foreign language depends on how similar or dissimilar the L2 and L1 are to each other. That is, only dissimilar features were regarded as causing L2 learning difficulties. This tendency seems to be one of the causes that made language practitioners disapprove of the CAH approaches because it focused only on the interference of the L1 with the L2 system. Moreover, the strong version dictates that *every difficulty* of this sort can be predicted *in an a priori fashion* through a comparison and contrast between the two languages. This, in turn, falsely led to

the belief that making a complete list of differences, across various domains, between two languages was only a matter of a linguist's time and effort.

Following this attempt, the ideas of negative transfer, positive transfer, and zero transfer of L1 linguistic items to the learner's L2 system have been suggested by Stockwell and Bowen (1965). Although this three-fold categorization appears to have been applied by a number of language professionals, a more elaborate model articulated by Prator (1967) is likely to offer deeper insights into inter-lingual foreign-language-learning difficulties. This model contains 6 levels from Level 0, "Transfer," to Level 5, "Split," which are going to be illustrated in more detail later in this section.

The weak version of CAH is different from the former version in that the weak version observes that most of FL learning difficulties arise from the L2 items that L2 learners do not already know (Oller & Ziahosseiny, 1970), while the strong version posits that only dissimilarities between the L2 and the L1 cause FL errors. However, these two versions have a common limitation in that they only focus on inter-lingual interference, although the weak version has been regarded as having a certain usefulness unlike the stronger version, which has been criticized as unrealistic and impracticable (Wardhaugh, 1970).

These two versions of CAH were almost jeopardized by Oller and Ziahosseiny's study on spelling errors (1970). They pointed out the fallacy that the strong version of the CAH would predict more spelling errors if the native language

of the learners featured a non-Roman script system. Based on the results from the same study, they also identified the fallacy that the weak version of the CAH would anticipate *equal* or greater difficulty in spelling if the native language of the learners possessed a non-Roman script system. What their experiment revealed was that the prior knowledge of a Roman letter system caused additional difficulty in the learning of another Roman alphabet system. The results of this study imply that both the strong and weak versions may have been based on hasty generalizations, which were hardly grounded in empirical data.

Specifically, the two early CAHs may have been unsuccessful in reflecting the difficulties caused by subtleties between two language systems. These strong and weaker versions of CAH also seem to have neglected the fact that spelling errors are not solely caused by individual sound-letter relationships. In addition, the two former versions seem to have failed to consider the complexity of the interlanguage spelling rules, which second language learners occasionally revise in their minds. The weaknesses outlined so far may be the major reasons for rejecting the CAH after the Chomskian view of language competence emerged.

The moderate version of CAH seems to have been established as these two former versions turned out to be unreliable. The most salient aspects of this new version are two fold. First, it recognizes the seriousness of FL learning difficulties caused by *subtle* distinctions between the target language and the native language, while according to the two previous versions, the greater the difference the two items have, the more serious the difficulty would be. Second, similar to

interlanguage theorists' observations, even the intra-lingual sources of difficulty, subtleties between two items in the *same* L2, has been acknowledged by this moderate version. This acknowledgement seems to be the first perspective in comprehending the intricacies of subtle distinctions. This moderate position appears to be *confidently* uncertain, rather than *unconfidently* certain. In what follows, the practical implications of the CAH for language teachers in the classroom, not for linguists or textbook writers, will be discussed.

#### IMPLICATIONS OF THE MODERATE CAH

The application the CAH to sentence structures generally seems to be limited to a few simple linguistic differences such as the different order of subject-verb-object in English compared to the one of subject-object-verb in Korean. Many other aspects of sentence structure<sup>10</sup> and vocabulary<sup>11</sup> appear to fall into higher domains than the segment-level sound system. The relatively high complexity of syntactic subcomponents often makes it impossible to account for every possible combination of subcomponents. It may be extremely hard to systematically compare the cultural bearings of vocabulary. A wide range of variations in the meaning of a word, in various contexts, seems to make the comparison virtually impossible.

---

<sup>10</sup> Take, for example, complicated rules that govern the position of an adverb in English sentences.

<sup>11</sup> Many of the requirements concerning complements to nouns or verbs in English are not likely to be effectively mastered through interlingual comparisons and explanations. Experience shows that some of them can be efficiently learned through simple memorization and repetition.

On the other hand, the application of the moderate CAH to phonetics and phonology is the most feasible. The fact that there are a relatively limited number of phonetic elements in a language predicts that systematic descriptions are feasible. In this sense, the moderate CAH seems to be helpful for preparing the major portion of this study's pronunciation agenda, specifically a problem-based syllabus. This preparation involves a posteriori analyses of the intra-lingual, as well as inter-lingual, pronunciation problems commonly observed in Korean adult learners of English. One more reason for considering the moderate version of CAH as the framework for analyzing observed pronunciation errors is that this version takes interlanguage rules into consideration, which means encompassing more sources of errors. Moreover, the most criticized contention of CAH (e.g. Pica, 1996) does not appear to belong to the moderate version, but to the strong version, which posits that a given set of linguistic data from the two languages is sufficient for difficulty prediction even with no contact with the speakers.

Interestingly enough, there seems to be a connection between the moderate version of CAH and the hierarchy of difficulty depicted by Prator (1967), one of the researchers that drove the early CAH movement, which leaned more toward the strong version. As reviewed earlier, even in the L2 itself, the most delicate differences will make it *most difficult* for the L2 learners to learn phonetic, grammar, and semantic aspects, according to Oller and Ziahosseiny's (1970) contention. Stockwell and Bowen's hierarchy of FL learning difficulty presents a prediction similar to this contention. This is because Level 5, the highest difficulty-bearing

level, “Split” (Prator in Brown, 1994), appears to describe the characteristic of the delicate differences suggested by Oller and Ziahosseiny. This level of difficulty is supposed to be caused when “[o]ne item in the native language becomes two or more in the target language, requiring the learner to make a new distinction” (p. 196). Take, for example, the sounds that are represented by the following underlined letters: “merge,” “zoo,” and “mirage” (/dʒ/, /z/, and /ʒ/ respectively). Many Korean EFL learners are likely to fail to differentiate these three sounds. The fact that these three words are transcribed into Korean with the same Korean consonant /ㅈ/, which is the Korean equivalent of /dʒ/, for example, “머지,” “주,” and “미라지.”

In addition to the above-discussed accountability of Prator’s hierarchy of difficulty that conforms to the moderate CAH’s unique prediction, there is another rationale for choosing Prator’s model, namely the sufficient number of difficulty levels, which is six. The six levels appear to have been further articulated than the two or three levels that other models offer, for example, the three levels of negative, positive, and zero transfers. Thanks to the superiority in detailing various levels, Prator’s model may be more effective in mapping the dissertation participants’ typical errors, and in investigating whether there is any pattern in the data in terms of difficulty of overcoming them. In what follows, the six levels with examples of Korean as the L1 and English as the L2 are discussed.

The level 0, Transfer, is described as “No difference or contrast is present between the two languages. The learner can simply transfer (positively) a sound, structure, or lexical item from the native language to the target language” (Prator in

Brown, 1994a, p.195). /ʌ/ of Korean and /s/ of English are a pair of examples of this level of hierarchy. This level is especially likely to be applicable to phonemic comparisons. It would not be simple to compare the two sound systems in terms of allophonic difference. For example, the phoneme /ɐ/ in Korean in a coda is the closest to the phoneme /l/ in English. However, [ɾ] and clear /l/ in English can be the closest allophones to /ɐ/ while clear /l/ ([l]), dark /l/ ([ɫ]), and syllabic /l/ ([l̥]) are the allophones of /l/. Therefore, zero transfer might not be the case for allophone-level comparisons even with the same combination, such as /ɐ/ and /l/. This possibility implies the existence of each different hierarchy according to different phonetic levels.

The level 1, Coalescence, is described as “Two items in the native language become coalesced into essentially one item in the target language” (Prator in Brown, 1994a, p.195). For example, both ‘ㅃ’ and ‘ㅍ’ are often transcribed by English speakers as ‘p’ in English. Koreans tend to think ‘ㅃ’ is the same as the letter ‘b’ and ‘ㅍ,’ the letter ‘p.’ This mismatch happens because Koreans tend to unconsciously sense the degree of aspiration to distinguish ‘ㅃ,’ which is voiceless and leniently aspirated, from ‘ㅍ,’ which is also voiceless but strongly aspirated. On the other hand, English speakers tend to unconsciously sense the degree of voicing to distinguish ‘b’ from ‘p,’ paying relatively less attention to the degree of aspiration. Therefore, when Koreans say ‘부산’ and ‘푸산,’ English speakers may tend to write them identically as ‘Pusan’ because they hear no voiced consonants in



either of the two words. In Korean, voiced /b/ and unvoiced /b/ are the two allophones of /ㅂ/.

The level 2, Underdifferentiation, is described as “An item in the native language [that] is absent in the target language. The learner must avoid that item” (Prator in Brown, 1994a, p.196). For example, there is no exact way to write the Korean name /헤/ ([hΛ]) in English. Therefore, Korean people write this name in various ways, ‘Her,’ ‘Hur,’ ‘Heo,’ and ‘Huh,’ only to find that none of these help Americans to pronounce the name correctly. The problems of this level are not more serious to Korean learners of English than level 1.

The level 3, Reinterpretation, is described as occurring when “An item that exists in the native language is given a new shape or distribution” (Prator in Brown, 1994a, p.196). For example, Korean glottal stops seem to exist only in some variations of exclamation sounds while the English glottal stop [ʔ] occur in numerous parts of speech. Most of the time, this level may be regarded as level 4, Overdifferentiation, if Korean learners do not recognize the existence of such sounds in their own native language.

The level 4, Overdifferentiation, is described as happening when “A new item entirely, bearing little if any similarity to the native language item, must be learned” (Prator in Brown, 1994a, p.196). For example, upside down ‘r’ (/ɾ/) sounds in English are often omitted or replaced with /l/ by Korean learners of English. It seems very hard to master the sounds that fall into this level when we

consider the EFL learners' tendency to replace such sounds with fossilized patterns of pronunciation.

The level 5, Split, is described as "One item in the native language [which] becomes two or more in the target language, requiring the learner to make a new distinction. The split is the counterpart of coalescence," (Prator in Brown, 1994a, p.196). For example, as mentioned earlier, the three English sounds /dʒ/, /z/, and /ʒ/ are often recognized as /ㅈ/ in Korean. It may be extremely hard for Koreans to recognize the differences among these English sounds without the help of explicit pronunciation instruction. This level presents almost the same difficulty as the one Texans feel when they need to differentiate the words 'pin' and 'pen.' The chronic mis-transcriptions of imported vocabulary from English in mass media seem to aggravate the difficulty.

In sum, the CAH, especially the moderate version, seems to offer the following teaching implications. First, when prudently employed in an a posteriori manner, CAH may be helpful in diagnosing the source of resistant FL errors. Second, if such analyses of observed L1-rooted errors are accumulated, organized, and shared, monolingual L2 teachers may apply them to seeking breakthroughs for removing fossilized items that the learners have. However, it seems that this approach would be infeasible and ineffective for domains that are more complex than phonology and syntax.

## **AGE ISSUES CONCERNING SECOND LANGUAGE LEARNING**

The research on differences between adult and child learners is important because this dissertation will involve adult EFL learners who seem to possess characteristics that are considerably different from those of child learners. That is, in order to devise a syllabus to help the adults overcome their global pronunciation errors, possible obstacles that adults tend to encounter, as well as their cognitive and motivational advantages need first to be identified.

### **ADULTS' DISADVANTAGES SUGGESTED BY THE CRITICAL PERIOD HYPOTHESIS**

Johnson and Newport's (1989) study on immigrant ESL learners' grammaticality judgment test seems to point to an observation that there is a certain juncture of life after which FL learning becomes relatively ineffective, often referred to as the Critical Period. However, there is also a study that favors language learning past this suggested critical period and that has found cases of nativelike achievement made even by adult learners (Birdsong & Molis, 2001). This investigation, which was a replication study of the aforementioned study by Johnson and Newport, revealed results that were contradictory to the original study. Therefore, it can be said that the debate concerning the effect of age on nativelike L2 attainment still remains in session.

In specific regard to FL pronunciation, Brown (1994) contends that second language learners are not likely to master a nativelike accent once they pass a certain juncture of life located near adolescence. If this is the case, adult language

learners, who have passed their adolescence, are less likely to be good imitators of foreign language sounds. In other words, adult language learners are doomed to be unsuccessful in terms of accent according to this discouraging contention. As a matter of fact, language learners after puberty seem to be insensitive to some delicate L2 sounds. Thus, it may be relatively hard for adult learners to overcome the global pronunciation errors related to these delicate L2 sounds.

Furthermore, the accent issue that the Critical Period Hypothesis raises causes L2 learners to have some misconceptions regarding most viable the age of language learning. Adults, in many cases, hardly expect themselves to be successful language learners, simply because there are few cases of adults acquiring a nativelike accent. Nevertheless, it must be noted that this view seems to be the case only if we imagine an ideal situation where all learner variables other than age of learning are equal. The overall success of a language learner, furthermore, should not be solely measured by the learner's accent according to Brown (1994). Some learners falsely believe that child learners only experience unconscious, automatic language learning with no intentional practice. Nevertheless, child learners are known to consciously practice speaking even their native language (Weir, 1962).

Superficially, Brown's term "The Henry Kissinger Effect"<sup>12</sup> (1994, p. 58) may seem to suggest that second language learners do not necessarily need to attain

---

<sup>12</sup> “. . . the former U.S. Secretary of State whose German accent was so noticeable yet who was clearly more eloquent than the large majority of native speakers of American English.”

a nativelike accent if their goal is to be able to effectively communicate using the L2. Pronunciation skills fluent enough to clearly or articulately speak L2 sounds could therefore be believed to be a sufficient level of L2 pronunciation to replace the concept of nativelike accent. This belief sounds acceptable if the learners do not need to assimilate into the L2 speaking society. For example, imagine a Korean EFL learner who knows how to pronounce the word ‘ventilation’ according to the phonemic transcription (/ˈventəˌleɪʃən/) in his dictionary with a proper word stress and phonemic accuracy. He may well be said to have a clear pronunciation ability for this word because it is likely to be understood aurally by native speakers of English. Moreover, he could also succeed in getting his message across even if he has some slightly inaccurate sounds in the word<sup>13</sup>, probably as the result of L1 interference.

However, when it comes to aural comprehension, merely having a clear pronunciation ability may not guarantee effective or successful communication. Suppose that he is watching a movie and he hears a word sounding like [ˈvenəˌleɪʃən] with an ellipsis of the sound [t] right after the first [n] sound in the word ‘ventilation.’ This pattern of pronunciation can be regarded as nativelike by most of adult EFL learners. The problem is that, without the [t] sound, it might be totally unintelligible to him because it sounds perfectly different from the word he knows. This situation suggests that the hazard of encouraging only clear

---

<sup>13</sup> Take [ventəˈleɪʃən] (Kim, 1999) for an example. This typical Koreanized transcription was generated by Kim’s Rule-Based Korean Accented English Pronunciation Simulator System. The word ‘ventilation’ was chosen by the researcher.

pronunciation could be serious especially for adult learners, who tend to be unable to sense such subtle variation unless explicitly told about it. There would be little need to insist on clear pronunciation if adult learners would jeopardize their aural comprehension abilities by not paying attention to nativelike patterns of L2 pronunciation. In fact, employing the concept of clear pronunciation seems to be effective in relieving the cognitive burden only at the initial stages of L2 pronunciation lecture. Even if an adult learner is unable to imitate such nativelike pronunciation, the language teacher at least needs to advise the student to beware of such hard-to-comprehend instances of realistic pronunciation.

Another problem of adult language learners is that the items that they have learned are often not realized in their communicative repertoires. According to Johnson and Newport (1989), adult learners are superior in phonology and syntax over child learners *only* at the beginning of their studies. This initial advantage seems to originate from adult learners' cognitive superiority in conceptualizing rules and analytical explanations about individual linguistic aspects. Nevertheless, better performance in learning linguistic items does not necessarily equal better performance in using the language. The need to holistically process information that is accumulated through language use may foster children's long-term success in L2 learning and usage.

One possible cause of this long-term success observed in children may possibly be rooted in high degrees of long-term memorization. Brown (1994) suggests that learners possibly tend to maintain what they have "meaningfully"

learned through the systematic forgetting process of detailed inner structure. This systematic process involves grouping of learned items into a bigger meaningful chunk, thus enhancing the efficiency in using them. However, this suggestion may be of little implemental value unless the term “meaningful” is precisely interpreted. Quite a few adult L2 learners, for instance, seem to be successful in maintaining L2 grammatical items in their long-term memory even if they do not show the sufficient competence needed to properly use the learned items in context. Some of them might have retained these linguistic items in their long-term memory just because the grammatical knowledge was highly important or meaningful in order to obtain a high score on their university entrance examination. Thus, for communicative language using, specifically *what* is retained in the learner’s long-term memory can be important as well. For optimal achievement of practical competence, L2 learning processes need to be sufficiently meaningful not for linguistic knowledge, but for contextual language uses.

Some meta-cognitive strategies to compensate for such disadvantage sometimes negatively influence adult learners’ learning. For example, the tendency of some ESL curricula to cover every detail, as if to make the students achieve a nativelike accent regardless of global or local errors, could aggravate the situation. The phonetic and phonological terms used to explain L2 pronunciation aspects in some curricula may confuse language learners, imposing too great a cognitive load on them.

In order to minimize this confusion, the linguistic terminology used in

teaching pronunciation to adult learners should not play a role as the subject to be internalized in the long-term memory of the learners, but as a scaffolding teaching tool in and of itself. Brown (1994, p.84) argues that certain tools, such as “definitions, paradigms, illustrations, or rules,” are believed to be made meaningful to enhance the subsumption process in the early phases of language learning. He adds that these detailed tools are, however, erased from the learner’s memory at later phases of language learning to foster automatic language use.

This argument may suggest that it is desirable to include the following three phases of adult L2 learning activities. First, language teachers need to present definitions, illustrations, and rules to adult L2 learners as temporary tools to facilitate their recognition of their own pronunciation problems. Zampini (1994) also stresses the importance of this kind of organized lecture that involves adult learners’ explicit appreciation of the dissimilarities between the L1 and L2 speech sounds. While lecturing, the instructor should make evaluative comments on these instrumental devices so that the students know that the devices are not the main contents to internalize. For example, the comment “Don’t worry about the phonetic terms! You don’t have to memorize them. They’re only needed temporarily if you don’t figure out where the sounds are generated. As you can see on this chart, the sounds categorized as ‘alveolar’ are produced right here,” will prevent the students from being overwhelmed by the complicated terms. Then, instructions on explicit strategies to effectively eliminate these problems need to be delivered to help learners to correctly produce such delicate L2 sounds. Last,



learning activities should include a sufficient number of opportunities for students to consolidate what they learned. These chances help to secure automatic and efficient language use as well as long-term retention. So far, the cognitive aspects of adult learners' disadvantages have been discussed. In what follows, a major affective facet of their disadvantages is reviewed.

#### **ADULTS' DISADVANTAGES CAUSED BY AFFECTIVE FILTER**

Krashen (1981) argues that the reason adult learners are not expected to succeed in long-term nativelike achievement is that their affective filters are often too high. This argument suggests the existence of mental blocks that prevent nativelike retention of learned items. Again, this tendency may only be the case given that other learner variables are equal. There seems to be a positive correlation between Korean adult students' low level of confidence and their tendency to save face especially in oral L2 communication. Affective filters may explain this correlation. Brown (1994) comments on the characteristics of adolescence that have emotional influence on language learning. The conceptualized selves shaped by the language experience of learners around puberty, according to his comments, tend to become too afraid to take risks due to the concurrent changes in their bodies, emotions, and cognitive abilities. This commentary implies that the stability of ego and identity can be the major drives of affective filter mechanisms. The apprehension of losing face, especially observed in Korean adult learners when speaking L2 poorly, can be perceived as dangerous to

the stability of their egos and identities. According to Schumann (1994), learners can be greatly influenced by high affective filters especially when they are poorly motivated, unconfident, and highly anxious.

Affective filters seem to be more important to adult language learners than to child learners. Scherer (1984) asserts that the emotional structure accumulated in a person by affective experiences seems to assess whether an incident complies with the person's value system about the society and self. In other words, children, who tend to have a relatively simple and malleable value system, may be less overwhelmed by the negative result of this filtering process. Some Korean people believe that it is not desirable to be seen as showy or flaunting. This social value may sometimes prevent Korean adults from actively participating in language classroom activities even if they are fluent enough to do so. Another salient learner variable, the lack of oral communication experience in the L2 due to unbalanced EFL curricula, can cause some adult learners to needlessly shape their self-images as poor at L2 speaking. These negative self-images can, in turn, inhibit the learners from lowering their affective filters and taking desirable risks in communicating with the L2.

#### **ADULTS' COGNITIVE SUPERIORITY**

Adult learners seem to have cognitive superiority over child learners. Hilles and Sutton (2001) contend that adult students have the advantage of a great

deal of life experience and cognitive maturity available for their learning. Life experience, including prior learned items, may provide a number of anchors to which newly learned L2 items can be linked if adult learners make the use of suitable language learning strategies. Adult learners are thought to be able to compensate for their relatively weak imitating ability by relying on their cognitive maturity, which enables efficient recognition of their own L2 sound errors. Therefore, language teachers must play the roles of guides to help adult language learners explore and exploit existing linguistic knowledge. Language teachers who have knowledge of the learners' L1 or similar L2 learning experiences, especially similar interlanguage, may easily guide the learners to make the best of the existing experience.

Brown (1994b) states that adult learners are sometimes expected to handle linguistic matters that are not directly related to their current circumstances. Adults' superiority in understanding abstract and conceptual elements can sometimes facilitate the understanding of certain L2 concepts. This superiority can allow language teachers to effectively use meta-linguistic terms in order to enhance learner uptake. Brown's statement implies that adequately employed meta-linguistic terms may foster systematic forgetting and subsumption by enhancing learners' conceptualization of L2 items. The meta-linguistic terms used in L2 learning classroom should not be overly complex. They should be fairly simple so as to avoid too great a cognitive load that may be harmful to practical language use.

## ADULTS' HIGH MOTIVATION

The other significant aspect of superiority adult learners possess is their advantage in terms of motivation. Gardner and MacIntyre (1991) present experiment results that show instrumentally motivated, as well as integratively motivated, subjects have more desirable learning outcomes. EFL learning situations are often unable to provide opportunities to satisfy integrative motivation to learn English. However, even in EFL settings, various types of instrumental motivation are available to adult L2 learners. The need to learn English to prepare for studying in English-speaking countries or the need to polish one's English conversation skills to get a job in a foreign company are examples. The limitation of this kind of motivation is that learning diminishes once the instrumental motivation disappears.

In a similar vein, there are also some encouraging observations such as the claim that adult students' life experience and cognitive maturity may be beneficial to their learning (Hilles & Sutton, 2001). Moreover, Brown claims that adults can even deal with language remote from a "here and now" context (1994). In addition, adult learners are assumed to be able to take advantage of their instrumental motivation for learning as well as their integrative motivation (Gardner & MacIntyre, 1991).

## **ROLES OF NON-NATIVE ENGLISH-SPEAKING TEACHERS IN THE CLASSROOM**

Issues concerning the roles of non-native English-speaking teachers (NNEST), whether superior or weak, are regarded to be important because this dissertation is aimed at oral English teaching in EFL settings, where there are a relatively large number of NNESTs, compared to the number of native English-speaking teachers (NESTs).

### **NEGATIVE SELF-IMAGES OF NNESTs**

Reves and Medgyes's (1994) survey discloses that non-native English-speaking teachers (NNESTs) have the most difficulty in choosing appropriate words to use. The NNESTs' vocabulary may often be limited to bookish words. The lack of experience in out-of-class language usage may cause the NNESTs' weaknesses in colloquial expressions. In fact, NNESTs seem to have difficulty in grammaticality judgment even though they are superior to NESTs when it comes to explaining grammar rules.

According to Reves and Medgyes (1994), non-native English teachers perceive that spoken English fluency involves the ability to converse promptly and logically in English at a desirable talking speed. The NNESTs tend to believe that they do not possess such abilities. These inabilities are maybe caused by the NNESTs' lack of contextualized expressions to use. There is a positive correlation between the NNESTs' concept of teaching qualification and the length of time they spent in English-speaking countries, according to Reves and Medgyes (1994).

There is also a positive correlation between the NNESTs' concept of teaching qualification and the frequency of the NNESTs' contact with native speakers of English, according to the same survey results. The reason for the lack of promptness in speaking might be that they have had relatively fewer opportunities for meaningful subsumption, which is achieved through repeated language uses and is thought to be essential for long-term retention. For effective EFL education, it may be crucial for NNESTs to compensate for the weaknesses discussed so far. Examining the roles of native English-speaking teachers might present the ideas of what NNESTs need to add to their teaching repertoire.

#### **NEST'S ROLES THAT TEND TO BE UNAVAILABLE TO NNESTs**

The study conducted by Samimy and Brutt-Griffler (1999) suggests that native English-speaking teachers (NESTs) are likely to provide the following benefits. Native speakers can play the roles of model speakers. They are capable of informal and smoother oral communication with fewer errors. They are able to use genuine language. NNESTs can reinforce their lecture with model utterance media if it is impossible for them to be genuine model speakers of English. As language teachers, NESTs seem to possess insights to cope with the delicacy of the learners' L2. Most of the time, they employ curricula focused on conversation, not on practice for language tests. Native speakers are believed to possess various methodology, approaches, and skills for classroom conduct. However, with

respect to these pedagogical instruments, NNESTs are not necessarily inferior to NESTs.

#### **ADVANTAGES OF NNESTs**

Canagarajah (1999, p. 91) states that “My [non-native English-speaking] teacher instilled in me his own curiosity toward the language, the ability to intuit linguistic rules from observation of actual usage, a metalinguistic awareness of the system behind languages, and the ability of creatively negotiate meaning with speakers and texts” (brackets by Canagarajah). Curiosity is one of the strong sources of intrinsic motivation to learn. NNESTs may have abilities to explain linguistic rules more explicitly than NESTs. NNESTs seem to have valuable metalinguistic awareness of the behind-the-scene system that is already systematically forgotten in most NESTs’ mind. NNESTs may have developed and accumulated a number of communication strategies through their learning experiences, which are similar to the ones of their EFL learners.

Samimy and Brutt-Griffler (1999) argue that achieving the goals of the L2 classroom relies on all the learner variables, situational constraints, and various aspects of the teacher, not solely on the teacher variables. Variables other than the nativeness of a language teacher for successful teaching seem to exist. NNESTs are sometimes more likely to offer successful teaching, when supported by authentic materials, to students that need a cure for global errors. NNESTs can play the roles of language teaching professionals with more pedagogical knowledge than

most non-professional native teachers of English have. However, the arguments presented by Samimy and Brutt-Griffler seem to be weak because the survey is limited to non-native MATESOL (Master of Arts in Teaching English to Speakers of Other Languages) students. Student-side reactions to NNESTs classroom performance should be considered to eliminate this weakness. Theoretical or empirical supports are needed to make these arguments stronger.

Cook (1999) contends that the perfect language abilities of native speakers do not seem to be achieved by the L2 learners, overpowering them and thus making them discouraged. With few exceptions, NESTs are hardly expected to play the same roles of advisors as senior EFL learners. NNESTs may be the models of successful EFL learners, who share experiences and concerns similar to the learners'. According to this contention, NNESTs might be superior in teaching low-level adult students, who often have high affective filters.

According to Cook's (1999) argument, the most reasonable way to describe English for ESL learners is to depict it as a second language to them, not as a first language of a native speaker. If English is described the same way as it is described as the L1 of the learners, it could be quite different from the description of English as the L2 of the learners. EFL learners do not intend to be English native speakers, but to be speakers of English as their L2. To some extent, L2 learners' role models are not likely to be English native speakers, but successful NNESTs. Cook (1999) also states that the voice onset times measured in the plosive sounds of the L1 sound system would be subtly shifted to the equivalents of the L2 due to the



learning process of the L2. This statement points to the existence of idiosyncratic learner language that is different both from English as the native language of the native speakers of English and Korean as the native language of the learners. This phenomenon should not be treated as a matter of correct or incorrect pronunciation, but as a matter of how clear and close it is to the sounds of the L2. It seems to be relatively easy for NNESTs to recognize the learners' interlanguage aspects. Language teachers need to value these aspects to help the students to build their confidence with improvement. This confidence is likely to help them continue to study the L2 on their own.

#### **APPLICABILITY OF SOCIO-CONSTRUCTIVISTIC VIEWS TO PRONUNCIATION LEARNING**

Superficially, perspectives that socio-constructivists take do not seem to be related to foreign language pronunciation learning. This might be partially because issues of foreign language phonology often involve discussions on rapid psychomotor controls using human speaking organs. This tendency can falsely make pronunciation learning be perceived as rather closely tied to behavioral views of learning. However, a closer look into some relevant socio-constructivistic (SC) arguments may well yield some fruitful considerations concerning pronunciation.

Socio-constructivists attempt to view learning and learners at various angles. That is, unlike the *behavioral* standpoint of learning, which often considers learners as passive knowledge storers, the constructivistic portion of the SC perspectives depicts learners as being capable of actively constructing, *meaningful* knowledge

structures. The other salient aspect of the SC considerations, the social portion, seems to emphasize that learning is a socially negotiated process, rather than an individual, isolated process.

Nevertheless, it must to be noted that the counterparts, with which learners negotiate their learning process or knowledge, do not necessarily need to be other human beings. Learning can also take place through interaction even with social artifacts, such as books or other semiotic devices (Wertsch, 1991), not to mention with teachers or peers. The ignorance of this position, amplified by the lack of the awareness that the SC approach is a kind of framework rather than a teaching technique, can lead to associating the SC framework merely with conversational language learning or group discussion activities.

#### **ORTHOGRAPHY AS A SOCIAL ARTIFACT**

English orthography is one of the keys to understanding how foreign accent items have come to exist, at least, in the Korean EFL environment. As pointed out in the first chapter, many instances of such orthography originated problems appear to stem from the discrepancy between Korean and English orthography. That is, certain English spelling rules perceived by Korean FL learners are often difficult to generalize into other similar spelling contexts as Korean orthography is regarded as relatively transparent, while English orthography is not as transparent (see Chapter 1 for details).

Therefore, a further question can be asked. Why do Korean EFL learners

tend to rely on English orthography even in oral communication situations, where they are not required to write in the L2? Aren't they able to simply focus on aural model input provided by native or non-native English-speaking instructors? Notwithstanding the FL perception models such as Perceptual Assimilation Model (Best, 2001) or Prator's (1971) Hierarchy of FL Learning Difficulty, one can make one additional speculation as to the reason.

The notion of learning through interaction with social artifacts, specifically FL orthography, needs to be taken into consideration. Based on relevant arguments put forth by Vygotsky in Wertsch (1991), the concept of "Conventional Sign" (p. 91), one of Vygotsky's semiotic mediators, appears to be closely related to FL phonics. It should be noted that most FL learning situations tend to involve studying phonics prior to an ample amount of aural exposure to the target language. This sign system formed in FL learners' minds at initial stages of their FL learning can be "included in the process of behavior" and thus "alter the entire flow and structure of mental functions" (p. 91). As FL learners' initial phonology learning is mediated by the semiotic system called phonics, their FL phonology processing, which can be metaphorically regarded as a form of mental function, may often be conducted via FL orthography, rather than through paying direct heed to the phonetic representation of an entire word. In what follows, the Korean EFL learning environment will be reviewed for any inherent disadvantages in terms of FL learning in the framework of the socio-constructivistic suggestions on learning. This discussion of drawbacks will then lead to possible solutions later in this chapter.

## SOCIO-CONSTRUCTIVISTIC REVIEW OF THE EFL LEARNING ENVIRONMENT

According to Bonk and Cunningham (1998), one of the salient aspects of SC learning is its emphasis on utility as well as knowledge negotiation. EFL learners seem to have at least two different kinds of utilities in learning English. First, as Wu (2001) cautioned, large-scale English exams, for example the college entrance examinations, force the first utility, achieving high English scores, upon these learners. Therefore, in terms of test preparation, EFL learning situations may be considered as satisfying these requirements for utility and knowledge negotiation. Another utility of English learning that was reflected in survey results is the development of speaking skills, which Korean learners of English feel are important (Truitt, 1995).

Unfortunately, many EFL learning activities appear to be ineffective in attaining the goal of practical communication. In other words, for this second utility, which is more communicative, some of the Korean EFL settings are likely to be considered as poorly socio-constructivistic, because the knowledge being negotiated in the EFL classroom is not geared to this utility, but rather to the utility of test preparation. Simply put, many Korean high school learners need to spend a considerable amount of time practicing solving English test items as quickly as they can in order to prepare for the college entrance examination. In what follows, some obstacles, which are a result of these circumstances, to developing pronunciation competence, one of the most neglected skills in the Korean EFL setting, will be discussed followed by a discussion on the possibilities of applying

recommendations from SC proponents.

There are at least two sources of obstacles to acquiring productive oral skill, out-of-class environment and in-class situations. Unlike ESL learning settings, it is very hard to find beneficial “systems of living knowledge” (Moll et al., 1993, p. 159), with regard to English learning, at the learners’ homes after school in their EFL settings. Furthermore, if we expand the scope of language learning beyond classroom situations, learning mediated by “accidental circumstances” (Salomon & Perkins, 1998, p. 15) may be viewed as highly beneficial for ESL pronunciation learning, as well as learning that is mediated by intentional design. This benefit seems to be unavailable to EFL learners because they are not in an immersion setting. For the same reason, EFL learners are not expected to receive even the benefits of “low road” learning (Salomon & Perkins, 1998, p. 15), such as practice and automaticity-based learning, outside the classroom, while ESL learners are assumed to acquire pronunciation skills through unconscious practice or automaticity-based input as they interact with other English speakers. The two above-mentioned disadvantages seem to highlight the importance of “high road” in-class instruction, which is intentional and concept-based in-class instruction, in EFL settings concerning pronunciation.

Nevertheless, even classrooms do not appear to be the reliable locus of oral English learning in EFL settings. In addition to the aforementioned problems caused by the college entrance examination or English proficiency score requirements for many Korean companies, there are two more reasons originating in

the classroom. Before we discuss that, however, a problem due to non-native English-speaking teachers in the classroom needs attention. According to Li's (1998) and Medgey's (1994) surveys, NNESTs perceive their oral English skills to be very poor compared to their grammar or reading skills. As former EFL learners, NNESTs are also subject to the lack of a sufficient immersion environment. The second possible problem that NESTs experienced in EFL teaching can be understood when reflected on the recommendations from SC proponents. As stated earlier, Cook (1999) claims that EFL learners tend to be overwhelmed by NESTs' perfection in their control over English, which is seemingly out of these students' reach. They may think they will never be like their NEST, no matter how hard they try.

One SC aspect that seems to shed light on this unexpected situation is the idea of "intersubjectivity" (Goldstein, 1999, p. 648). Noddings (1984, p. 13) argues that intersubjectivity is the "constellation of conditions" that is perceived and shared by the eyes of both sides, namely, students and the teacher. The problem that causes EFL students to be overwhelmed is that it is not easy to expect intersubjectivity, for example, similar English education backgrounds or the same L1, which often interferes with the L2 sound system, to be formed between EFL learners and NESTs. In order to seek ways to alleviate the problems with EFL environments, the following aspects will be reviewed based on the recommendations from SC approaches: 1) EFL pronunciation learning goals, 2) roles of NNESTs facilitating SC FL learning, 3) learning content or class

organization, and 4) roles of computers.

#### **EFL PRONUNCIATION LEARNING GOALS RECONSIDERED**

First, typical EFL pronunciation teaching goals may need to be modified as follows. To be prepared for the convoluted intricacies of the real world (Bonk & Cunningham, 1998) might be one of the high priority aims that EFL learners have to set. Specifically, nativelike pronunciation again needs to be employed in order to avoid fossilized foreign accents and to reduce potential aural comprehension problems when conversing with native English-speakers in authentic settings. Another goal concerning student activities in the classroom is to involve learners into the learning process, rather than to elicit correct answers from the learners (Goldstein, 1999). This goal is expected to be achieved through “cognitive apprenticeship,” as opposed to the conventional information delivery model of the teacher-student relationship. The metaphor of “skilled trades and crafts” that Bonk and Cunningham (1998, p. 36) used to describe this apprenticeship appears to match well with such a psychomotor skill as L2 pronunciation. In addition, it seems plausible to scaffold the learners into stand-alone learning situations, specifically more communicative activities such as conversation activities, gradually ceding control over to the learners.

## ROLES OF NNESTs AS FACILITATORS OF SOCIO-CONSTRUCTIVISTIC FL LEARNING

As for the second SC recommendation, it seems desirable that EFL pronunciation education is guided by teachers rather than by peers. This argument is worth discussing because SC recommendations in general favors negotiation of information and meanings between a learner and a more knowledgeable peer, rather than *teacher-centered* lecture. The reason for considering *teacher-guided* pronunciation education is that teachers need to assume the responsibility to determine appropriate phonology-learning goals for learners, which is hardly expected from more knowledgeable EFL peers (Noddings, 1984).

In addition, EFL teachers should seek ways to sense each learner's potential problem-solving level, which is the upper boundary of the zone of proximal distance (ZPD). It seems probable that teachers who have a greater amount of intersubjectivity, such as NNESTs, would perform this job easily. Situationally contingent guidance (Salomon & Perkins, 1998) and more personalized approaches are also assumed to be provided by NNESTs, as they may have deeper insights into the problems that EFL learners encounter in class. The idea of learning conversations (Goldstein, 1999) suggests that teachers as well as learners have to adapt to each other, unlike the conventional teacher-centered lecture. By doing this, teachers can dynamically perceive the learners' heightened upper ZPD boundaries and sense when and how much they can relinquish control over to their students.



#### CLASS GROUPING & LEARNING CONTENT RECONSIDERED BASED ON SC SUGGESTIONS

The actual effectiveness of heterogeneous grouping or class organization suggested by Erickson (1996) needs to be reconsidered with respect to Korean EFL settings, contrary to his suggestion. It seems that EFL pronunciation classes need to be formed as homogeneously as possible. The first reason for this grouping is that pronunciation appears to be the component that bears the most self-consciousness compared to reading, listening, and writing skills. In addition, Korean adult learners tend to be more afraid to take risks if they consider their peers to have a superior fluency level than they are when they feel confident with their pronunciation skills. Another reason is that heterogeneous grouping in terms of ZPD may hinder widely shared intersubjectivity. For example, high-level students who have not experienced a certain difficulty in pronunciation would be bored when they are asked to perform learning activities to overcome this difficulty, while other lower-level students may find the same activities to be helpful.

As far as teaching contents are concerned, Prawat and Floden's (1994) recommendation that teachers need to employ representations, not only for contents, but also for students, sounds reasonable. One of the most common reasons that EFL learners find pronunciation tutoring to be boring and hardly meaningful may be because EFL instructors normally rely on phonology or phonetics textbooks, or at best unedited audio or video materials. Therefore, it would be beneficial to use authentic materials such as sitcoms, popular movies, or TV commercials that are adequately edited for effective presentation. However, these model pronunciation

clips should be organized properly so that their linguistic utility can match the learners' need and prior experience in pronunciation. Wertsch (1991) has named the use of semiotic mediation as an effective device in positively altering the flow of instruction and the structure of learners' mental function. Representations semiotically mediated by speaking-organ charts or by graphical presentation of vowel articulation places may well be effective in enhancing conceptually oriented learning, which Salomon and Perkins (1989) conceive of as high-road learning.

#### **FRAMEWORK FOR ADULT FL PHONOLOGY LEARNING: NOTICING HYPOTHESIS**

This section will first describe how the Noticing Hypothesis came to exist. Then, detailed description of this hypothesis will follow. After that, Output Hypothesis and processing instruction, the two salient underpinnings of Noticing Hypothesis will also be discussed. Then, the discussion of how they are related to the syllabus of this study will be provided.

#### **EMERGENCE OF NOTICING HYPOTHESIS**

Describing the emergence of the Noticing Hypothesis may date back to the 1980s, when the famous input hypothesis (Krashen, 1981) impacted the second language acquisition (SLA) theories. The input hypothesis claims that comprehensible input ( $i+1$ ), one of the subconscious learning components, is a sufficient, as well as necessary, condition for children and adults' language acquisition. However, research has shown that there was some counter-evidence

to the validity of this influential SLA approach. According to Swain (1988), children who remained in an immersion SLA setting since kindergarten were clearly identifiable as non-native speakers and writers by the end of their elementary school years even though they scored equivalently to native speakers of the target language in listening and reading comprehension. In addition, it has been shown that students, even with extensive immersion and comprehensible input may not acquire a number of L2 morpho-syntactic features without form-focused instruction (Swain, 1988). These findings suggest that unconscious learning of language solely based on comprehensible input might not be sufficient for fuller acquisition of the L2. The Output Hypothesis (Swain, 1993), which underscores the role of output to force learners to process language syntactically, may be understood as one of the few breakthroughs in the study of these problems. Along the same lines as this movement, the Noticing Hypothesis (Schmidt, 1990), which emphasizes the beneficial function of noticing the forms of the L2, has drawn the attention of language practitioners.

#### **RELEVANT CLAIMS OF NOTICING HYPOTHESIS**

According to Schmidt (1995), only the portion of given input that is noticed and processed by the learner's focal attention and awareness becomes intake for learning. A stronger claim of this Noticing Hypothesis argues that noticing is the necessary and sufficient condition for transforming input to intake, based on the analysis of Schmidt's L2 learning diary (Schmidt & Frota, 1986). The need for

noticing is justified by the following argument. While children are capable of keeping an open awareness when learning, adult learners can only afford focal attention to meaning (Schmidt 1990). Thus, if a learning task does not match the learning focus of the task, say learning syntax, adults' implicit learning of forms is hard for adults to achieve. For example, adult learners may face additional difficulty in recall protocol when they are requested to conduct the task of listening to a passage with the learning focus on the *morphemes*, a *syntactic* focus, while they experience no decrease in their recall abilities when they do the same task, but with the learning focus on the *lexical* items (VanPatten, 1990).

The research concerning noticing seems to be twofold. First, studies on facilitators of noticing can be found. According to Izumi (2002), output activities such as note taking is better for noticing these forms than input enhancement that involves underlined complex noun phrases. A couple of studies including VanPatten and Cadierno (1993) highlight processing instruction as effective in enhancing accuracy. Second, there has been a further attempt to investigate the effects of noticing on eventual learning. However, Izumi et al. (1999) failed to prove that noticing certain forms and being able to incorporate them into language use have a significant effect on acquiring the forms. In what follows, the two facilitators of noticing, namely output and processing instruction will be discussed, as input enhancement does not appear to be beneficial.

## **ROLES OF PROCESSING INSTRUCTION AND OUTPUT**

The Output Hypothesis is closely related to the Noticing Hypothesis in that the output facilitates noticing of certain forms of L2. Swain (1995) argues that the output has a noticing/triggering function. The proponents of Output Hypothesis put forward the following arguments as rationale: The output is claimed to have the role of forcing learners to process L2 syntactically (Swain 1988), which involves quite different processes from semantically comprehending the language (VanDijk & Kintsch, 1983). In its weak version, the Output Hypothesis suggests that output may be another facilitator of SLA under some condition, in addition to comprehension and input (Swain & Lapkin 1995). They claim that without the output, learners' language development is likely to slow down. In a stronger position, they argue that even without feedback from the interlocutors of the learners, on occasion, output practices can help the learners notice the gap between their interlanguage and the target language. However, it seems reasonable to posit that, with no output, learners would not push themselves to produce L2 utterances and they would never acquire such facilitative feedback as requests for clarification, which, in turn, nudge them to modify their prior output.

The notion of processing instruction (PI), which means instructions on strategies to process linguistic data, is also related to the Noticing Hypothesis in that this instruction is effective in making learners notice relevant L2 forms. VanPatten (2002) states that PI may assist learners in making form-meaning connections during input processing (IP). Therefore, PI is likely to make the learners better

processors of input, ultimately. PI is based on the analysis of how learners process input without PI. Based on this failure-driven analysis the original input is intended to be manipulated so that learners can afford a better processing strategy away from their ordinary strategies they otherwise tend to stick to. The bases of PI proponents are as follows: VanPatten (1984, 1988) asserts that, in order to pay attention to language forms, automatic meaning processing is required, which is not likely to be afforded by many L2 learners. Thus, without any help, learners may never acquire certain forms, especially the ones that appear to possess little communicative value or diminutive contribution to the overall sentence meaning.

#### **SYLLABUS BASED ON PROCESSING INSTRUCTION AND OUTPUT**

The syllabus of this study adopts both of these two facilitators, output and PI, to help EFL learners to overcome their often-unintelligible accents. The current EFL learning situations emphasize the difficulty of solely relying on the  $i+1$  driven approaches. As an example, Anō (1998) lamented that, even after four years of comprehensible input in Japan, English learners demonstrated almost no attempt to speak English on their own. Korea's situation is not unlike this because the author of this dissertation received good grades with ease throughout six years of formal English education in Korea without having to say sentences any longer than "Thank you!" Even though most of the investigated areas in noticing research pertain to syntax learning so far, the phonemic and phonological aspects of L2 learning are

also expected to get some benefits using these approaches (VanPatten, 1996; Schmidt, 1990).

The need for including both output and PI approaches is as follows: The PI is not intended to include any production activities (VanPatten, 2002). One salient reason for employing output activities in EFL education is equivalent to the reason why explicit pronunciation education is needed. Comprehending utterances is not necessarily expected to provide decent opportunities to practice producing ones. VanDijk and Kintsch (1983) point out that comprehension, in many cases, does not involve skills needed to produce utterances, often observed in native speakers, but rather uses comprehension strategies that are solely based on lexical items needed to *understand* sentences. The reason that output alone is not suitable for this study's treatment is that it is hard, in EFL settings, to encounter fluent L2 speakers who can be expected to provide beneficial feedback such as requests for clarification. Moreover, research shows that instructed learning is more effective in making learners capable of transferring their learned rules to novel L2 items than incidental or enhanced input learning is. This ability to transfer seems valuable in EFL settings because these settings hardly offer a sufficient amount of input for inductive and unconscious generalizations.

The teaching approach of this dissertation consisted of the following components that are modeled after the PI. First, there was structured input (VanPatten, 2002) for learners. That is, unlike "fluid and freer" haphazard output activities warned about by VanPatten and Cadeirno (1993), short authentic video

clips, such as sitcoms, news broadcasts, and TV commercials, which were likely to deliver more focused model sentences, were implemented. These video clips were collected and edited based on the typical problematic English sounds for Korean adult learners and organized under certain categories to be learned. Second, there were not only desirable pronunciation models, but also some heavily foreign accented or unintelligible examples, in order to follow VanPatten's (2002) recommendation for mixed (correct and incorrect) referential activities. The Korean instructors for the experimental groups could present some of these contrastive incorrect examples. Also video clips edited from Korean soap operas, which were used mainly for their Korean loan words imported from English, were also a part of the examples of this sort. The instructions were primarily focused on why certain English sounds were problematic for Korean adult learners as well as what optimal pronunciation would be like. Due to the unique nature of pronunciation learning, the "Affective Structured Input" activities, recommended by VanPatten (2002) for real world connections, was not included in the pronunciation syllabus.

Inspired by the Output Hypothesis, the learners were encouraged to imitate the model pronunciation right after they received the PI described above. Thus, they were expected to recognize their own relatively unintelligible pronunciation patterns. Intensive comparison between the learners' interlanguage and the model utterances, assisted by a number of convenient user interfaces offered by



computerized video playback programs, instantly followed the learners' attempts to modify their output.

#### USE OF AUTHENTIC VIDEO CLIPS FOR TEACHING PRONUNCIATION

Unfortunately, there seem to be very few cases in which videos are used specifically for teaching pronunciation in spite of the fact that one can find a number of cases in which videos, whether digitized or analog, are used for language or language-related education. Davis (1999) reports in her action research that the participants, who are multi-national adult ESL learners, have made improvements ranging from 80% to 95% in isolated English-word pronunciation. Not to mention the unclear accounts of what exactly the improvement range means, this study is not likely to bring forth much implication to the current study for the following reason. The video materials entitled *Perfect English Pronunciation* were used for the learners who had finished their "Survival English" courses. However, these videotapes seem to have been intended only for pronunciation tutorials, not for any authentic pronunciation models. The need for shifting toward realistic video materials, from video contents which are pedagogically created for language learners, is claimed by Katchen (1997). The reason for this move is that these educational video materials mainly consist of speed-reduced and carefully articulated speech samples. In other words, this is because, if they are only familiar to such inauthentic materials, they can often be frustrated when they encounter even simple, but unedited, English utterances. Other areas, which a

number of video-employed language teaching studies have investigated, include socio-linguistic or cultural competencies (Witten, 2000; Herron et al., 2000; ASIA Society, 1995; Chamberlain, 1994).

#### **USE OF HTML FOR IN-CLASS MULTIMEDIA PRESENTATIONS**

The form of learning through in-class lecture, to which this study purports to apply the syllabus, presents contextual properties including the following. First, regular classroom lectures, unlike online classes, require the learners, as well as the instructor, to manage the limited time given by the study's experiment. Second, classroom lectures afford, the learners, the instructor's real-time guidance throughout the learning activities.

Considering these two characteristics of in-class lecture, HTML (Hyper-Text Mark-up Language) seems to be the most cost-effective means to incorporate various kinds of media into the teaching-aid materials for segment-level English pronunciation. The hypertext or hypermedia, which can be implemented through the HTML technology, is suitable for multimedia presentations that have non-sequential structures (Richter, 2003; Conklin, 1987) such as foreign language sound systems. Let alone hypermedia's versatile media-integrating capabilities, the advantage of instant and random access to embedded multimedia clips is expected to greatly assist the instructor in managing his/her class period. For example, by simply clicking a hyperlink, the instructor can immediately switch the learners' attention back and forth, from the lecture notes to the video clips, which exemplify

the relevant notes within one screen. The random access capability can enable the instructor to dynamically react to the students' needs, especially when their needs are unexpected by the instructor before the lecture. For example, the instructor can instantly replay any pronunciation model-videos to clarify or elaborate the learned items that the learners have not fully understood, without bothering to rewind or fast-forward the videotape.

One of the challenges that hypertext technology has newly brought to education professionals' attention, also points to the use of this technology in the classroom, rather than outside of the classroom for learning activities independently conducted by learners. Richter (2003) warns that each learner's pattern of internalizing presented hypertext content can be so idiosyncratic that the learners still need to be escorted by proper instructional guidance even in Internet-mediated education. This view conforms to Laurillard's (1995) claim that broad instructor-side input is needed, in the form of priming the learners for the learning tasks and for follow-up clarification attempts, together with proper monitoring. The second characteristic of the in-class lectures outlined above suggests that the classroom may be one of the optimal loci that can offer the guidance needed to take advantage of the hypertext innovations.

## Chapter 3

### **METHOD**

The following research questions were addressed in this study:

1. Are adult EFL learners better able to aurally recognize English sounds after they receive explicit L2 pronunciation instruction based on their interlanguage and L1?
2. Will they be perceived as having more intelligible accents by native English speakers when they pronounce English words?
3. Are they better able to transfer the learned pronunciation knowledge to novel English words?
4. Do the English sounds classified according to Prator's hierarchy of difficulty, with the consideration of Korean as the L1, show difficulties of pronunciation learning that are similar to the levels suggested by Prator?

### **OVERVIEW OF THE EXPERIMENTAL DESIGN**

This study was aimed at comparing two of the many possible explicit approaches to teaching pronunciation in an EFL setting. One was a rather conventional classroom instruction model that was based on a pronunciation

textbook written by English native speaking ESL practitioners and which was taught by native English-speaking instructors. In this study, this group is called the control group in that the factor of error analysis was controlled for in this group's syllabus. This group's syllabus was, instead, modeled after a typical pronunciation lab class supplemental to an ESL speaking course for Korean graduate students in the United States. The experimental group's syllabus was organized according to an analysis of pronunciation errors observed among Korean adult learners of English.

To achieve the goal stated above, this dissertation involved a mixed method. The quantitative portion of this study consisted of an experimental design with two cells for each of the experimental and control groups as can be seen on Table 1. The qualitative portion that involved several semi-structured in-depth interviews was included for two reasons. First, this study needed to deal with a somewhat rarely explored teaching technique that utilized authentic videos as pronunciation models, not as a medium that delivered cross-cultural information, which had been adopted by quite a few language practitioners. Therefore, purely quantitative methods may not have been suitable for this somewhat exploratory study. Second, qualitative investigation was likely to substantiate the statistical analysis of this study. For a strong hypothesis testing, it may be better to pick only one component from the teaching techniques of this study. However, such a highly experimental design would not help in guiding teachers with the more complex task of teaching pronunciation.

Table 1

*Experimental Design*

Group	Class <sup>14</sup>	No. of Participants	Instructor
Experimental	A-1	19	K1
	B-1	19	
	A-2	19	K2
	B-2	19	
Control	A-4	18	E1
	B-4	18	
	A-6	18	E2
	B-6	18	
Total	8	<i>N</i> = 148	4

**PARTICIPANTS**

At the beginning of this study, the researcher organized eight English pronunciation classes by assigning the 148 participating students at a Korean tertiary institution as shown on Table 1. In this institution, EFL instruction by native English-speakers was generally available. This assignment process was conducted using a spreadsheet's function that generated random numbers. To control for unequal numbers of prior EFL courses taught by English native speakers, only one grade of students, namely freshmen, was chosen for potential participants. For a similar reason, the students who had graduated from any foreign-language-

<sup>14</sup> Actual names of the classes were used for the participants' convenience. The other non-participating classes took another course on English speaking.

augmented high schools in Korea and who had stayed more than a month in English speaking countries were excluded. Prior to their entrance to the university, all the participants had received 12 semesters of EFL courses that were taught by Korean teachers at their middle and high schools in Korea. Similar to other typical EFL students, they were considered as having neither considerable contact with English native speakers nor as having had opportunities to use English for communication outside of their EFL classes. The participants included approximately 10 percent women.<sup>15</sup>

There were four participating teachers who taught the student participants with the experimental and control group's teaching approaches (see Table 2). One of the experimental group's instructors had been teaching English at the university for four semesters at the time of the experiment. He had studied abroad in the United States for a degree right before he started teaching on this campus. He appeared to have almost no Korean accent when he spoke English. The instructor of the other experimental group, the author of this dissertation, had studied as a Ph.D. student in the U.S.A. for four and a half years. His major was teaching English to speakers of other languages (TESOL). The English native-speaking instructors of the two control classes had been teaching for four and five semesters in the same university. Both of them held a degree in one of the humanity majors. The male native-speaking teacher was from the state of Colorado and the female teacher was a Canadian who considered herself an English native speaker.

---

<sup>15</sup> The admission policy of this university has a mandate to accept only 10% female students a year.

## PRONUNCIATION TEACHING APPROACHES

### TEXTBOOK FOR THE CONTROL GROUP

The teaching approach for the control group was centered on a student workbook entitled *Pronunciation for Success* (Meyers & Holt, 1998). This workbook was written by two native English-speaking practitioners. It contained nine units from “The Basics” through “Linking,” which had been organized for a month-long intensive tutorial of the North American English pronunciation patterns. Only four of the entire units were taught to the control group as they were related to segment-level pronunciation. They are as follows: 1) The Basics, 2) Enunciation: Consonants, 3) Enunciation: Vowels, and 4) Word Stress. These four units were intended for 19 days of instruction according to the authors.

This workbook was chosen for the control group as it appeared to have been organized without paying attention to any specific L1 of the English target learners (refer to Appendix C for a sample page). Most of the consonant minimal pairs had been prepared based on their places of articulation, not on expected difficulty for the target learners to encounter when differentiating between them.<sup>16</sup> For example, the bilabial stops /p/ and /b/ were presented together as a minimal pair, the contrast of which is the existence or absence of voicing. This contrast is relatively easy for Koreans to differentiate, even if they tend to rely more on a different criterion, that being strength of aspiration. At the same time, the control group workbook did not

---

<sup>16</sup> The only one exception was the L2 contrast, /ɪ/ and /I/, which had been widely known to be problematic to Japanese learners of English.



include minimal pairs such as /p/ and /f/, the difference between which is even more confusing for Koreans.

The control group workbook offered explicit pronunciation learning activities. It presented a number of speech organ diagrams in the vicinity of each English sound it explained. In addition, explicit explanations about how to articulate the target sounds, such as tongue positions, jaw openness, and lip rounding, were included in the workbook. In terms of learners' L1-related issues, this workbook contained a kind of L1-L2 difference awareness section, which asked the learners to analyze their native language or try to mimic an American tourist trying to say some L1 phrases transcribed in the English alphabet. However, this activity was not provided in consideration of a specific L1. There was, however, a comment on Korean learners' undesirable tendency to insufficiently open their jaws in the tutorial video.

#### **TEXTBOOK FOR THE EXPERIMENTAL GROUP**

The researcher wrote the workbook for the experimental group that contained 33 pages of teaching items. These items included three sections respectively on word stress patterns, consonants of North American English, and vowels of North American English, followed by appendices. Each section consisted of the following three subsections: 1) Eye-Openers, 2) Important Features for Intelligibility, and 3) Advanced Features for Sounding Nativelike. Each subsection, in general, contained example English words, when necessary, with IPA

transcriptions and explanations in Korean (refer to Appendix B). The subsection Eye-Openers consisted of activities such as double-checking whether one's way of saying the English alphabet was correct and a list of words that had the same vowel sounds and which Korean students tend to believe contain more than one vowel sound (see Chapter 1 for actual examples).

The most salient aspect that characterized the experimental group workbook was its basis, typically committed error patterns, on which the pronunciation agendas were organized. This unpublished workbook is similar to the control group workbook in that they both listed almost the same English segment sounds.

However, the experiment workbook contained minimal pairs that were contrasted in consideration of Korean learners' difficulty. These items were also chosen based on Chan, Burtis, and Bereiter's (1997) recommendation for the inclusion of items that are stunning to the learners or incongruent with their prior knowledge. Unlike learning agendas organized stepwise, they argue, these types of conflicting knowledge that are abruptly presented to the learners may be more effective in leading them to knowledge reformation. There were a number of pictorial materials in the appendix that depicted the organs of speech and some computer screenshots and pictures that showed misleading Korean transcriptions of English words. In what follows, the kinds of pronunciation items included in the experimental workbook will be briefly discussed.

Table 2

*Features of Experimental Syllabi*

Group	Instructor's L1	Language Used for Instruction	Basis of Teaching Agenda Organization	Textbook & Teaching Aid
Experimental	Korean	Korean	Analysis of errors observed in Korean EFL learners	In-house workbook & Computerized authentic video clips
Control	English	English	No specific L1 of the learners considered	<i>Pronunciation for Success</i> <sup>17</sup> & Computerized tutorial videos <sup>18</sup>

**AGENDAS BASED ON L1-INTERFERENCE-INDUCED ERRORS:** The experimental group's syllabus included teaching agendas that were garnered from an analysis of the L1 sound system's interference with the L2 sound system. Best's (2001) Perceptual Assimilation Model and Prator's (1967) Hierarchy of L2 Learning Difficulty were used to determine the anticipated difficulty level of certain L2 contrasts that were contained in typically observed pronunciation errors of Korean learners of English (see Chapter 2 for detailed sound sets and examples of this sort).

<sup>17</sup> A workbook that was used in the pronunciation lab classes supplementary to an adult ESL course in the U.S.

<sup>18</sup> The first one of the two tutorial videos for *Pronunciation for Success* was computerized.

**AGENDAS BASED ON INTERLANGUAGE-INDUCED ERRORS:** Another kind of pronunciation item included in the experimental workbook was prepared through the analyses of the interlanguage patterns that the author of this dissertation observed among Korean EFL students. The term interlanguage in this study needs to be confined further since it has been used by language practitioners for referring to a number of similar, but different, entities; “approximative [linguistic] system” (Nemser, 1971), p. 116) or “idiosyncratic dialect” (Corder, 1971), p. 156) to name a couple.

Nemser’s notion of approximative system seems plausible for describing the nature of the term interlanguage in this text. Nemser (p. 116) defines interlanguage, “An approximative system is the deviant linguistic system actually employed by the learner attempting to utilize the target language.” His emphasis on the successive nature of different L2 proficiency stages may even encourage optimistic views about imperfect, but systematic and indispensable, interlanguage items.

The teaching agendas for the experimental group included relevant examples obtained by the *intralingual* analysis of observed errors as well as examples from the *interlingual* analysis. One may imagine that the teaching agenda only focuses on inter-lingual error analyses when they hear of the *interlanguage* analysis, which has been the basis of the teaching item selection. However, *intralingual* diagnoses on EFL pronunciation patterns should not be considered to be less important. For example, a rendition of the word “desperate,” with the underlined part spoken in the

same way as the word “rate” should be pronounced, can nudge language practitioners to reflect it on the future improvements to their pronunciation curriculum.

On the other hand, the concept of interlanguage system employed in this study excludes instances of Corder’s idiosyncratic dialect. This subset of a linguistic system pertains only to one specific speaker of a certain language even though it seems to mean a number of shared patterns because of the term *dialect*. That is, the part *idiosyncratic* can be understood as denoting that it possesses nothing in common with other “social dialects” (Corder, 1971). Therefore, the part *dialect* for Corder appears to neutrally mean a variation of a linguistic system. A few instances of this category were observed when the author organized the experimental group’s agenda. Nevertheless, they were excluded from the agenda because, being observable in only one particular L2 learner, they seemed to bear little educational implication. An example of this sort was actually found while grading this study’s L2 reading-aloud tests. Consistently on the pre- and post- tests, one of the participants pronounced the sequence ‘tt’ in the word ‘pretty’ as the first consonant of the word ‘three’ ([θ]). The fact that this pattern of substitution had not been observed throughout the researcher’s 25 years of English learning and teaching experience convinced the researcher that this interlanguage pattern yields little L2 learning implication.

**AGENDAS BASED ON SOCIO-CULTURALLY-INDUCED ERRORS:** Another important foundation for the pronunciation agenda was relatively non-linguistic and non-education-originated. This basis included socio-cultural sources of imperfect pronunciation patterns such as Korean EFL learners' value-laden attitude towards certain pronunciation patterns and the undesirable influence of imperfect mass-mediated Korean transcriptions of English loan words. The pilot study (see Chapter 1) revealed the possibility that some Korean adult students considered it to be arrogant or showing-off to include, in their English pronunciation repertoire, ample amount of r-coloring or substitution of [t] sound with the English flap sound. The negative socio-cultural connotations that these items can invoke to others might persuade Korean EFL learners not to speak English as natively as possible, even when they are able to do so. The words 'car' and 'party' provide some examples.

The previously mentioned example of "desperate" can also be used to illustrate another kind of socio-culturally induced errors, the influence of mass-mediated transcriptions. This word happened to be the title of a movie that was imported to Korea. The Korean transcription of "Desperate" was "데스퍼레이트" (refer to Appendix A). Most Korean movie theaters hang a huge banner that is several yards long right above its entrance in order to draw their customers' attention. Therefore, the transcribed title in the banner such as the aforementioned problematic one is normally legible in the distance. The more famous the movie, the more widely this awkward pronunciation can be spread into EFL learners minds

as a reliable pronunciation pattern. In addition, similar phenomena may be easily found on the Internet, television, magazines, etc.

#### **PRONUNCIATION LEARNING SITUATION OF THE CURRENT STUDY**

There were four participant teachers who taught the student participants with the experimental and control group syllabi. Each of the four instructors taught two classes in a row using his/her native language. Each instructor was accompanied by a teaching assistant while teaching, who manipulated computerized presentation materials in the classroom.

The experimental group's instruction provided highly nativelike pronunciation models delivered by short computerized authentic video clips such as edited sitcoms, news broadcastings, or TV commercials. These video clips were chosen and edited by the researcher during his stay in the States from his collection of American movies in the format of DVD (digital video disc). He also recorded some cable channel programs mentioned above using his VCR (video cassette recorder). While preparing for the instruction materials for this study, he scanned the video contents for examples that were likely to be effective in illustrating his interlanguage-based teaching items. In addition, he succeeded in locating a few instances of Konglish<sup>19</sup> pronunciation in some Korean soap operas that he rented from a local Korean grocery store. These samples were expected to help the participants realize what Korean accented English words sound like, according to

---

<sup>19</sup> "Konglish" is an informal Korean word that denotes Koreanized English.

VanPatten's recommendation for mixed referential activities (see p. 53). After choosing appropriate video clips, he digitized them into computer video files, the lengths of which ranged from 8 to 30 seconds.

In contrast, the control group's learners relied on the instructor's pronunciation demonstration and the pronunciation tutorial video that came with the textbook, which was also computerized in order to balance the degree of technology implementation in both groups. In the tutorial video, the authors of the control group's workbook introduced, explained, and demonstrated the teaching items of the workbook. They tried to illustrate how their students changed in terms of foreign accents by filming their students' before-and-after performance in speaking English. One of the differences between the experimental and control groups' video clips was that the former presented unmodified speech streams while, in the latter, examples were intentionally enunciated and elongated by the tutors.

The ease of presenting media materials as teaching aids was virtually identical for both the experimental and the control groups. As mentioned above, the teaching assistants were requested to set up each instructor's laptop computer on which necessary videos were saved in advance. During the class, the teaching assistants instantly played pre-designated computerized videos, with ease, according to the instructor's advance and operated relevant audio-visual devices such as the LCD projector or the stereo that was connected to the notebook computer.



In general, as noted earlier, both the experimental and control groups' syllabus explicitly dealt with pronunciation issues in class. However, they were different in that the experimental group's instruction mainly drew on explicit instructions about *possible causes* of the observed *interlanguage patterns*, while the other primarily focused on explicit instructions about *how to correctly pronounce* English sounds. In summary, the independent variable that characterized the current study's design was *method* of English pronunciation teaching. This factor was stratified into two levels of syllabus, curriculum that was organized on the basis of target learners' error analyses and a rather traditional L2-based curriculum. Each teaching approach was characterized by more than one element, the instructor, teaching agenda, and teaching aid.

One could suggest another experimental design in which a native English-speaking instructor delivers pronunciation lectures based on an error-based agenda specifically observed in Korean EFL learners and a native Korean-speaking instructor delivers pronunciation lecture based on the textbook written by non-Korean speakers.

However, this potential experimental design was out of this study's concern for the reason described in what follows. L1- and interlanguage- based teaching agendas were not likely to help most non-Korean instructors very much. The video clips for this problem-oriented approach were organized through the researcher's analysis of Korean learners' consistent pronunciation errors, as described above. For optimal pedagogical benefits, the instructor needs to have a

decent understanding of the learners' L1 specifics in order to connect these video clips' utterances and the students' foreign accents. This L1-specific awareness includes the knowledge of the learners' L1 sound system, insights into their EFL curricula, an understanding of learning situations outside the EFL classroom, etc, which seem relatively unavailable to most monolingual native English -speaking instructors.

For a similar reason, an experimental design, with a native English-speaking instructor using the Korean language as the teaching medium and a native Korean-speaking instructor using the English language as the teaching medium, was excluded. The findings of this dissertation, of course, were not based on each subcomponent of the teaching approaches, but rather on each of the teaching approach as a whole.

There seemed to be two sources of extraneous variance, the variance resulted from the participants' learning effects in two other English courses at the university. First, as school years in Korea start in March, they took two semesters of English listening comprehension courses at the university, namely *Practice in English* and *Basic English Conversation*. The former was taught by native Korean-speaking teachers of English, while the latter was taught by native English-speaking teachers. The goal of the first course was to make the freshmen, who rarely had afforded opportunities to communicate with native English speakers before college, prepared for the ensuing English conversation courses. The second course that they took

could be regarded as a typical EFL conversation course taught by a native English-speaking instructor.

The second external source of English learning was a couple of other English-related courses that the participants were taking during the experiment period, which was the intensive course. The learning effects of these other courses could have been confounding unless properly handled. In order to neutralize this potential nuisance variable, the participants were randomly assigned to experimental and control groups.

As shown in Table 1, each class was comprised of eighteen or nineteen students. The syllabus was scheduled to start at the beginning of the entire intensive course to avoid the influence of the other intensive courses on the pre-tests, which might undermine external validity of this study's results. However, due to personal complications, the pronunciation classes began at the beginning of the 2<sup>nd</sup> week. The syllabus consisted of four 75-minute classes that were taught over one week. The four A classes (refer to Table 1) were taught first. After a fifteen-minute break, the four B classes were taught. Each of the classrooms for the treatment and control groups was equipped with multimedia devices described above with a large screen in the front of the classroom. There were two blackboards in these classrooms, one in the front and the other on the back of the room.

The instruction for the experimental group was conducted in the learners' L1, Korean, while the native English-speaking instructors delivered their lecture in English to the control group. The lecture only included the segment-level features of the American English sound system except for the word stress patterns. The word stress patterns were mentioned during the experiment although they pertained to supra-segmental aspects. The researcher included these stress patterns to illustrate some stress-related variations of segmental sounds (see Chapter 2).

The teaching aid for the experimental group included pronunciation models imported from a CD-ROM dictionary (the Cambridge Dictionary of American English version 1.0, 2000). In addition, some of the teaching agendas were also accompanied by corresponding realistic models edited from unmodified footages, such as movies, situation comedies, TV news, and TV commercials, as discussed earlier. These authentic pronunciation models in the video clips were expected to provide highly nativelike pronunciation patterns that model utterances from the CD dictionary failed to present. For example, the verb "enter" was pronounced, in the CD dictionary, with the /t/ fully aspirated, while in a movie the /t/ was significantly reduced. These authentic examples were presented to help the learners comprehend the word aurally when they hear them in a realistic conversation.

The lesson plan for the control group was organized based on time needed for the introduction, explanations, media presentation, and the recommendations in the textbook *Pronunciation for Success*. For the current study, the video tutorials for the control group that come with this textbook were also digitized into video

files, in order to maintain a balance in terms of media delivery mode. This allowed for enhanced convenience in manipulating the videos when they were used in both groups' classrooms.

## DATA COLLECTION

### QUANTITATIVE DATA TO ADDRESS RESEARCH QUESTION 1

In order to answer the first research question, “Are adult EFL learners better able to aurally recognize English sounds after they receive explicit L2 pronunciation instruction based on their interlanguage and L1?” a multiple-choice listening test was created by the researcher. This test was administered twice as a pre- and post-test in order to measure the changes in the participants' identification abilities before and after the treatment. The participants were shown an English sentence with a missing word, which was marked with parentheses. Then, they heard the same sentence twice with no missing word. The participants were instructed to choose the right word that they heard to fill in the parentheses. The tested minimal pairs were presented within a sentence, rather than as an isolated word. The testees needed to pay attention to the phonetic contrast of each word for the blank in order to identify it properly, as each embedding sentence matched well with any of the given choices. Refer to the following question for an example.

7. At that time, Alice seemed to be obsessed with playing video games, and was very (        ).
- a. zealous      b. jealous      c. 모르겠음 (*I don't know.*)

There were a total of 16 questions of this sort in the test. The participants were presented the same written stem twice in two different questions. For example, they heard the recorded sentence, “At that time, Alice seemed to be obsessed with playing video games, and was very jealous,” after the seventh question above appeared on the computer screen. The correct answer to this question was ‘b.’ In Question 9, the test-takers encountered the same sentence and choices on the screen. This time, however, they heard a different sentence, “At that time, Alice seemed to be obsessed with playing video games, and was very zealous.” If the participants were able to aurally tell the difference between the minimal pair, “jealous” and “zealous,” in other words, if they could acoustically differentiate /z/ from /dʒ/, they would choose ‘a’ for this question.

To enhance the validity of the score, the following three aspects were considered and handled. First, the test-takers might possess too poor an English listening skill to properly understand and follow directions for the test. To eliminate the error variance due to this kind of external factor, all the directions were recorded in Korean by the researcher. Second, the participants might attempt to guess answers to obtain high scores even when they were not able to differentiate them adequately. As minimal pairs were presented as possible choices, guessing answers might yield an expectancy of 50% accuracy even without listening to the recorded sentences. To control for this, the researcher explained, in the directions, that the test scores would have no effect on their grades. In addition, the participants were encouraged to choose the last choice, “I don’t know,” when they

were not able to identify the word they heard. The last concern about the validity was that the participants might try to find any pattern in the order of the paired questions. The 16 questions were shuffled so that the paired questions appeared randomly on the test, in order to handle this potential exterior variance (refer to Appendix D).

In order to grasp a more general picture of how the participants, as learners, perceived the treatment, a pronunciation-learning experience questionnaire was used at the end of the treatment. This survey involved the following areas of inquiry: participants' attitude toward the syllabus, perceived quality or impressions of the classes, effects of students' L1 as a medium of instruction, effects of L1-L2 contrastive approaches, perceived learning outcomes, etc. The survey items on learning outcomes asked about the student participants' self-perceived improvement in recognizing their own foreign accents, in general listening comprehension, and in internalizing how to pronounce L2 sounds, followed by any perceived changes in their L2 confidence (see Appendix E for the survey). Each item was rated using a 5-point Likert Scale.

## **QUANTITATIVE DATA TO ADDRESS RESEARCH QUESTION 2**

Addressing the second research question, "Will they be perceived as having more intelligible accents by native English speakers when they pronounce English words?" involved a pre- and post- test. Each of these pre- and post- tests consisted of two reading-aloud tasks, reading real words and reading artificial words. These

tests included a collection of words that were believed to contain sounds (see Table 3) typically problematic to adult Korean learners of English. These words were chosen based on the pilot study results and the researcher's hypothetical error-analysis data observed in Korean learners through his EFL learning and teaching experience. There were 38 real words (see Appendix F) that were centered on 19 principal sounds. The artificial words were created by the researcher with the help of two native English-speaking instructors (see Appendix G). The student participants were requested to read the test items aloud as colloquially as they could, imagining that they were imitating American English speakers in everyday conversation.

To control for the nuisance factor of performance errors caused by pronunciation slips, they were allowed to instantly reread any words that they felt they had mistakenly pronounced. Therefore, if there was more than one reading of a word, the last one was graded. All the directions for the tests and the surveys were written in Korean. The length of recorded pre- or post- test ranged from one to one and a half minutes per participant. Before the pre-test, the researcher suggested to the participants that the post-test could be entirely different from the pre-test in order to prevent them from memorizing the test items.

### **QUANTITATIVE DATA TO ADDRESS RESEARCH QUESTION 3**

The researcher attempted to address the 3<sup>rd</sup> research question, "Are they better able to transfer the learned pronunciation knowledge to novel English



words?” using reading-aloud pre- and post- tests that contained artificial words, as seen on Table 3. This novel word reading-aloud test consisted of a set of artificially created non-sense words, for example, “mipper” or “javor,” to measure the degree of the learners’ transfer capability of learned pronunciation rules to unrehearsed contexts. These artificial words were created by the researcher with the help of the two native English-speaking instructors who participated in this study. Three of the artificial words, ‘mansiderate,’ ‘plaw,’ and ‘Tarate’ were not analyzed because the grader of this reading-aloud test, a third native English speaker, found them ambiguous in terms of pronunciation.

#### **QUANTITATIVE DATA TO ADDRESS RESEARCH QUESTION 4**

To seek the answer to the fourth research question, “Do the English sounds classified according to Prator’s hierarchy of difficulty, with the consideration of Korean as the L1, show difficulties of pronunciation learning that are similar to the levels suggested by Prator?,” the results of the two previous language tests were reviewed against the levels of difficulty that the researcher assigned to each tested sound based on Prator’s hierarchy.

Table 3

*Reading-Aloud Items on Pre- and Post- Tests*

	Nativelikeness for Real Words		Nativelikeness for Artificial Words	
Word to be read	organ <u>iz</u> ation	wa <u>ter</u>	trap <u>iz</u> ation	na <u>ter</u>
Principal sound	[z]	[r]	[z]	[r]
Typically observed sound	[dʒ]	[tʰ]	[dʒ]	[tʰ]
Source of difficulty	Level 5, Split <sup>20</sup>	Level 5, Split Or Socio-cultural attitude	Level 5, Split	Level 5, Split Or Socio- cultural attitude

**QUALITATIVE DATA**

The results of the quantitative analyses based on these instruments were substantiated using some qualitative data. In spite of the devices for controlling confounding factors mentioned so far, there still might be some validity and reliability issues. To minimize this kind of possibility, a total of seven participants were interviewed after the quantitative analysis was finished. The author of this dissertation tried to elicit, from the interviewees, concrete examples related to their answers.

<sup>20</sup> See Chapter 2 (p. 55) for Prator's Hierarchy of Difficulty.

The interviewees were chosen based on their pre- and post- test results and their answers to the EFL-learning background and pronunciation-learning experience surveys (refer to Appendix H). These semi-structured in-depth interviews took 40 to 50 minutes per interviewee. There could have been bias in the interviewees' responses when interviewed face-to-face, as the researcher might have appeared to be a future professor of the participants. That is, having to reveal their identities and faces to their seniors while interviewing might make them think needlessly about desirable responses, for example. To minimize this negative inclination, the researcher interviewed them over the phone. There was a third-party coordinator who set up the interview schedule and sent a list of the interviewees' cellular phone numbers to the researcher. This coordinator also sent the instruction sheet on the interviews to the potential interviewees on behalf of the researcher. The interviewees were requested not to reveal their names for the purposes of confidentiality while on the phone with the interviewer. The researcher made approximately six hours of international calls to Korea using an affordable calling card from the states. The interviewees were referred to using the pre-designated pronunciation ID numbers, rather than their names or cadet ID numbers. Each interview session was digitally recorded on the computer for the ease of transcription. In addition, it was also tape-recorded to prevent data loss due to possible technological problems.

## **DATA ANALYSIS**

The data analysis processes involved the analyses of the demographic data, word identification test, segment-level pronunciation evaluation, and interviews. Analysis procedures of the acquired data other than the reading-aloud tests were relatively simple. In general, they were processed for descriptive statistics and inferential statistics such as ANOVA and ANCOVA results. The indices of practical importance were calculated for any statistically significant results. The alpha level of 0.10 was chosen for determining statistical significance, as the approaches of this study's teaching pronunciation were perceived to be unexplored (see Chapter 2). The recorded interviews of the chosen participants were transcribed by the researcher.

### **PROCEDURE TO ANALYZE EXPERIMENTAL TEST RESULTS**

The word reading-aloud tests that aimed to measure changes in the students' segment-level pronunciation abilities necessitated the following analysis processes. First, a principal sound, which was the focus of the evaluation, was designated to each of these words, as exemplified in Tables 3 and 4. In other words, the grader paid attention to only one focused principal sound in each word.

The reading-aloud tests were graded by the researcher, whose native language was not English. It took an hour to listen to and grade 13 to 15 recordings. Each participant yielded a set of four recordings, ones for real words on the pre-test, artificial words on the pre-test, real words on the post-test, and

artificial words on the post-test. The participants who missed one or more of the four recordings were excluded for their data would not lead to proper comparison. A total of 94 participants were evaluated on all of these four tests. Therefore, there were a total of 376 recordings, each of which contained approximately 40 words to grade.

The reason that the researcher wanted to include as many as 19 English sounds in a test, despite the complications ensuing from such a large number of data, was as follows: First, even though the pilot study was conducted with Korean adults, the current participants seemed to have a quite dissimilar EFL-learning profile. The pilot study's participants were 30 to 36 years old, while the current participants were all freshmen. At their pre-tertiary schools, these two groups of learners appeared to have faced a considerably different English curricula and English test policies for the college entrance examination system. Thus, the researcher opted to cover a relatively large number of sounds on the test in order not to miss any potential promising teaching items.

This reason conformed to the second reason for focusing more than a few auspicious English sounds. That is to say, as pointed out earlier in the literature review, the areas of manipulating authentic video clips and using students' L1 to teach L2 pronunciation have not been explored very much. This rather probing nature of the current study, coupled with a teaching agenda that involved as many sounds as this study's tested sounds, was expected to shed more light on EFL pronunciation teaching than typical rote research designs were, with only a few

confusing sounds, such as /l/ and /ɪ/ (Ingram & Park, 1998; Takagi, 2002; Takagi & Mann, 1995). Furthermore, when the researcher considered the fact that the current study was a pedagogical investigation on teaching methods, not a linguistic study that focused on specific acoustical properties of certain L2 sounds, he was quite certain that the amount of light shed would be even more substantive.

### VALIDITY ISSUES

The following validity issues of *nativelikeness* were raised in relation to the quantification of the L2 word reading-aloud samples. The most fundamental consideration was about the domain of the measurement. The major domain of investigation was segment-level English sounds, even though what the participants recorded was L2 words, not isolated L2 sounds. The students, for example, could score 3, which signified ‘nativelike,’ by speaking the word “filler” even as [ˈp<sup>h</sup>ɪˌlɪɹ] with the proper high front lax vowel [ɪ], if the principal sound was [ɪ], whether the other sounds were correct or not. On the contrary any student who spoke the same word as [ˈfɪˌɪɹ] was supposed to earn only 1, which meant “intelligible, but foreign-accented,” despite the fact that he/she pronounced everything else natively.

The second validity issue regards its restriction to the specific L1, namely Korean, which the tested L2 learners spoke outside of the class. This means that scores assigned to the participants’ pronunciation instances may bear a different degree of validity in terms of nativelikeness, if the participants’ native language is not Korean. For instance, if a Korean learner of English reads the word ‘e-mail’ as,

then the learner deserves only 1 because he substitutes the L2 sound [t] with a similar sound of the L1 [l]. On the other hand, a Japanese learner of English may acquire 2, which means “imperfect, but close to nativelike” with the same foreign-accented pattern as [imeɪl]. This is because Japanese does not have the phoneme /l/ and the occurrence of this L2 phoneme observed in the Japanese speaker’s utterance points to a remarkable improvement in English pronunciation.

Another validity consideration centers on how close the grader’s nativelikeness judgment is to the target native English-speakers’ intuition. This issue seems to be closely related to what the grader’s English proficiency and experience in the related fields were like. The brief profile of the researcher, who graded the entire pronunciation samples, as an English speaker, was as follows: He used to be an English instructor at one of the college-level schools in Korea. He could be considered as a typical Ph.D. candidate in TESOL (Teaching English to Speakers of Other Languages). Teaching pronunciation was one of his areas of interest and he completed relevant linguistic and pedagogical courses at a large U.S. university. At the time of the current study, he had lived in the States for four and a half years.

#### **MEASURES TO ADDRESS VALIDITY ISSUES**

The intraclass correlation coefficient (ICC) was used for quantitatively measuring how similar the researcher’s nativelikeness judgment was to the native English informant’s. This statistic is more widely used as an index of inter-rater

reliability, which indicates the consistency of the scores given by more than one grader. The researcher hypothesized that if the scores that he assigned and the scores that the other native English speaking grader assigned showed a reliability above a certain threshold, it could be said that the reading-aloud test graded by the researcher measured what the native English-speaker would measure with decent validity, if not as validly as the test results acquired through this American.

Based on this rationale, the researcher asked a native English speaker<sup>21</sup> to grade approximately 10% of randomly chosen data, before embarking on the process of evaluating the entire reading-aloud samples on his own. She possessed some experience in grading segment-level pronunciation samples, as she was the one who graded the speech samples of Korean learners of English when the researcher conducted a pilot study prior to the current study. First, he prepared for the first meeting with this native English-speaking informant by grading some of the recordings in advance. He was able to create grading criteria for the real and artificial word recordings while grading on his own. These criteria contained transcriptions of the principal sounds that were actually observed in the recorded samples and their corresponding scores suggested by the researcher. The researcher and the informant reviewed the criteria together for the appropriateness of the scores assigned to each of the observed pronunciation patterns (see Appendix I for the final criteria). Then they graded a few samples sitting together, followed

---

<sup>21</sup> The female native speaking informant was born in Texas. She was granted an MA in Foreign Language Education. She had a great deal of interest in phonology. She took a number of phonology and phonetics courses for graduate students, which offered competitive training opportunities to handle issues of phonology and phonetics that seemed highly relevant to the area investigated in the current study.



by short discussion sessions of any discrepancies in these pilot scores. He audio-recorded the English speaker's own readings of the words as model utterances on the spot for future reference. This preparation session took roughly two and a half hours.

Then, he handed the randomly chosen 10% of the entire recordings, which numbered 48 samples, to the native English-speaking grader. The computer audio files that contained the samples were marked by a nickname that only the researcher knew, instead of the classifications "pre" or "post" in their file names, in order to prevent the grader from knowing which recordings were taken before or after the treatment. It took four days for the female grader to evaluate them. Meanwhile, the researcher also graded the same samples on his own, according to the criteria that were made complete through his meeting with the American informant. These pilot samples included recordings of real words done by 24 participants and recordings of artificial words also done by the same 24 participants. Of these samples, only the real word recordings were used for calculating the ICC because the native English speaker revealed that she was relatively unsure of her grading results from the artificial tests. She explained that it was harder for her to assign nativelikeness indices to the sounds in the non-sense words, maybe because she had never heard of them.

The pronunciation evaluation was, in general, based on the North American dialect of English, so-called American English. Unlike the unintuitive Likert scale using nine levels of nativelikeness (Munro & Derwing, 1999), the current study

used four grades of nativelikeness, as illustrated in Table 4. First, the score of 0 was given when a learner's reading-aloud word was *unintelligible*. For *intelligible* words with a degree of Korean accent, the raters were supposed to give the score of 1. The score of 2 was designated for clear pronunciation, but with overt enunciation, for example, 'water' with [t<sup>h</sup>] in it. For cases such as the pronunciation of [k<sup>h</sup>aɪ̯] attempted for the word 'car,' which could be regarded as imperfect, but close to nativelike patterns, the score of 2 was also assigned. Finally, the score of 3 was given for words with full nativelikeness, say, [p<sup>h</sup>.ɪɪɪ] for 'pretty.'

Table 4  
*Coding System for Pronunciation Assessment*

Accuracy Index	Description	Example (Principal sound)	Model Utterance	Observed Utterance
0	Unintelligible	th <u>o</u> ught (ɔ)	[əɔt <sup>ʔ</sup> ]	[əoout <sup>ʔ</sup> ]
1	Intelligible but Foreign Accented	z <u>oo</u> (z)	[zu]	[dʒu]
2	Clear but Unnaturally Enunciated	wa <u>t</u> er (r)	[wɔɾɪ]	[wɔt <sup>h</sup> ɪ]
	Imperfect but Close to Nativelike	ca <u>r</u> (ɪ)	[k <sup>h</sup> aɪ]	[k <sup>h</sup> aɪ̯]
3	Nativelike	pre <u>t</u> ty (r)	[p <sup>h</sup> .ɪɪɪ]	

## Chapter 4

### RESULTS AND DISCUSSIONS

#### EFL LEARNING BACKGROUND SURVEY RESULTS

There were 14 female participants of the entire 144 participants who took the EFL learning background survey. The portion of female students in the participating groups was 10%, which was the same as the portion of females in the entire freshmen class. This demographic survey revealed that four participants had graduated from foreign language high schools in Korea. And, one participant reported that he attended a college in Greece, which was established for international students. These five participants were excluded due to their quasi-English-speaking immersion experience, which might potentially skew the results of the study. Data acquired from two more participants, who lived in English speaking countries for 10 months and three and a half years respectively, were also excluded. The demographic results presented in what follows were based on the responses of the 137 respondents who were not excluded.

The earliest age at which members among these 137 participants had started their EFL learning was 5, while the latest age was 14. On average, they had started studying English at 11.4 years of age.<sup>22</sup> They would have been 3<sup>rd</sup> or 4<sup>th</sup> graders at

---

<sup>22</sup> This figure represents the age in the Korean age convention, in which a newly born baby is one year old at the same time he/she is born. Every New Years day adds one year to this age regardless of date of birth in this system.

this juncture of their life. The current study's freshmen seemed to have a noticeably different learner profile, compared to the participants of this study's pilot, who appeared to have started formal English learning when they entered middle school, approximately 2 years later in their life than the initial English learning age of this study's participants.

Table 5

*No. of Semesters of English at Elementary School*

No. of Semesters <sup>23</sup>	8	7	6	5	4	3	2	1	0
No. of Respondents N = 136	2 (1.4%)	0	6 (4.4%)	0	10 (7.3%)	1 (0.7%)	4 (2.9%)	6 (4.4%)	107 (78.6%)

The majority of the 136 respondents who answered No. 5 (see Appendix H for the survey item) of the survey (78.6%) answered that they took no English course at the elementary school level as can be seen on Table 5. Combined with the previous results on their age of initial English learning, it can be said that the majority of the participants tended to have learned English outside of school when they first learned English. Even if it was outside of school settings, it is hard to say that they learned English as they *used* it in real life, due to the fact that the official language in almost the whole of Korea has been almost exclusively Korean. There are some more possible loci of studying English. They may include various

<sup>23</sup> In the Korean School system, students are supposed to take 12 semesters of classes from 1<sup>st</sup> grade to 6<sup>th</sup> grade in the elementary school.

classroom contexts at private English institutes and the ones created in classes by home-learning tutors, the number of which has remarkably increased in South Korea as of late.

The minimum age at which the participants first learned the English alphabet was 5 and the maximum age was 14, similar to the age range of initial English studying. Interestingly, the survey disclosed that they first learned the English alphabet a couple of years earlier than they started receiving formal English education, namely, when they were 9.7 years old on the average. This result suggests that the participants' pronunciation might be formed through a bottom-up cognitive processing because they were expected to have had no significant exposure to authentic L2 utterances until they started their formal English *learning*. ESL learners or native English speakers tend to first become capable of communicating orally in English, due to the comprehensible input they receive even before they learned how to read letters. Unlike these English learners in immersion, the EFL participants could have relied on analytical hypothesis-testing strategies based on their knowledge about English phonics, when they learned to read or say the alphabet in English.

Another notable aspect of the participants' L2 learner profiles that was different from the one of the researcher is their experience of English classes taught by native English-speaking teachers (NESTs). Unlike the participants of the pilot study, who first had a chance to talk to native English speakers at the college level, there were some participants in this study who had had studied English with NESTs

before entering college. As illustrated in Table 6, roughly 24% of the participants took English classes with an NEST as their instructors for at least one semester before they entered college. This means one in every four students had experienced some form of extended contact with English speakers at the pre-college school.

Speaking ability turned out to be the skill that the participants wanted to improve the most at the beginning of the treatment, together with the second most reported desired skill of listening (refer to Table 7). On the other hand, they responded that they felt the least need for their improving writing skills. In general, the participating students could be said to possess relatively high motivation for polishing their English speaking and listening skills.

Table 6

*No. of Semesters of English Taught by an NEST*

No. of Semesters	No. of Respondents			
	Elementary <i>N</i> = 137	Middle <i>N</i> = 137	High <i>N</i> = 137	Total <i>N</i> = 137
6	0	0	4 (2.9%)	5 (3.6%)
5	0	0	0	1 (0.7%)
4	1 (0.7%)	0	1 (0.7%)	1 (0.7%)
3	0	0	1 (0.7%)	0
2	3 (2.1%)	7 (5.1%)	4 (2.9%)	15 (10.9%)
1	2 (1.4%)	4 (2.9%)	10 (7.2%)	10 (7.2%)
0	131 (95.6%)	126 (91.9%)	117 (85.4%)	105 (76.6%)

Table 7

*Priority of English Skills to Improve*<sup>24</sup>

Speaking	Listening	Reading	Writing
<i>N</i> = 137	<i>N</i> = 137	<i>N</i> = 137	<i>N</i> = 137
1.71	1.76	3.01	3.52

The next aspect of the survey results to explore is the patterns of the participants' L2 use. Based on Table 8 and 9, the participants' L2 learning behavior dramatically changed during the English intensive course (EIC). The majority of the participants, ranging from 63.5% to 86.8% of the entire participant sample across the language skills, reported that they spent on average less than 10 minutes a day in practicing English listening, speaking, and writing according to Table 8. Even for English reading, which was the skill that they practiced and learned in the previous semester, most students (86.2%) expended less than one hour per day in practice. On the contrary, at the beginning of this intensive course's 2<sup>nd</sup> week, the amounts of time they spent in listening, speaking, and reading in English increased and over four hours a day was the answer that the majority of students, 52.1%, 44.2%, and 38.1% respectively, gave (See Table 9). This phenomenon could be largely attributable to the English-only policy of the EIC program. The increase in L2 use observed in the participants might, in turn,

<sup>24</sup> Priority ranged from 1 to 4, 1 being the most improvement and 4 being the least improvement desired.

heighten the influence of error variance on the total variance that could result from the treatment approaches, making the results less statistically significant.

Table 8

*Amount of Time per Day Spent in Using English Skills Before EIC*

Amount of Time	No. of Respondents			
	Listening <i>N</i> = 137	Speaking <i>N</i> = 137	Reading <i>N</i> = 137	Writing <i>N</i> = 136
> 4 hr.	0	0	0	0
2 - 4 hr.	0	0	2 (1.4%)	1 (0.7%)
1 - 2 hr.	6 (4.3%)	2 (1.4%)	17 (12.4%)	2 (1.4%)
1/2 - 1 hr.	15 (10.9%)	3 (2.1%)	43 (31.3%)	7 (5.1%)
10 - 30 min.	29 (21.1%)	13 (9.4%)	45 (32.8%)	16 (11.7%)
< 10 min.	87 (63.5%)	119 (86.8%)	30 (21.8%)	110 (80.8%)



Table 9

*Amount of Time per Day Spent in Using English Skills During EIC*

Amount of Time	No. of Respondents			
	Listening <i>N</i> = 136	Speaking <i>N</i> = 136	Reading <i>N</i> = 136	Writing <i>N</i> = 135
> 4 hr.	71 (52.2%)	60 (44.1%)	52 (38.2%)	10 (7.4%)
2 - 4 hr.	27 (19.8%)	24 (17.6%)	41 (30.1%)	13 (9.6%)
1 - 2 hr.	14 (10.2%)	23 (16.9%)	25 (18.3%)	30 (22.2%)
1/2 - 1 hr.	16 (11.7%)	17 (12.5%)	14 (10.2%)	33 (24.4%)
10 - 30 min.	6 (4.4%)	11 (8.0%)	4 (2.9%)	37 (27.4%)
< 10 min.	2 (1.4%)	1 (0.7%)	0	12 (8.8%)

**L2-SOUND IDENTIFICATION TEST RESULTS**

The  $\alpha$  value that was set by the researcher as the basis to determine statistical significance was 0.10. This exploratory value was chosen as there was almost no pronunciation study that involved the experimental syllabus covering a number of L2 contrasts in research. Therefore, any pronunciation items that showed statistical significance are likely to provide the starting points for further research in a more rigorous experimental design.

A total of 110 participants took both the pre- and post- L2 sound identification tests. The entire participant pool consisted of 50 experimental samples and 60 control samples. The means of post-test scores acquired from the experimental and control groups were compared with each other through the analysis of covariance (ANCOVA) with the covariate of pre-test scores using the Univariate General Linear Model function available in SPSS. Another method of handling models with pre-tests such as gain score analysis was considered to be improper for the current study's data because there was not a perfect linear relation between the pre- and post- test scores ( $r < 0.6$ , at the 0.01 level). Prior to ANCOVA analyses, tests of between-subjects effects were conducted in order to check whether there was significant interaction between the independent variable, group, and the covariate, the pre-test scores. Because there was no significant interaction effect between these two factors, ANCOVA could be regarded as valid for the sound identification data.

#### **TOTAL SCORE OF THE IDENTIFICATION TESTS**

In order to statistically compare the mean difference between the post-test total mean scores of the experimental and control groups, an ANCOVA analysis was conducted with the covariate of the pre-test total score. Prior to this ANCOVA analysis, tests of between-subjects effects were conducted in order to check if there was significant interaction between the independent variable, group, and the covariate, the pre-test scores. The interaction term turned out to be non-

significant ( $p = 0.725$ ), ANCOVA was valid for the total score of the identification post-test.

The ANCOVA analysis showed no significant mean difference ( $F_{(1, 107)} = 0.074$ ,  $p = 0.786$ ) between the total scores of the experimental and control groups (see Table 10). Therefore, the null hypothesis that there is no statistically significant mean differences between the experimental and the control group was not rejected. In terms of identification accuracy tested using the minimal pairs embedded in ambiguous sentences, the effects of the experimental and control groups' pronunciation teaching approaches were not significantly different.

Table 10

*ANCOVA Source Table for Identification Total Score*

Source	Sum of Squares	df	Mean Square	F	Sig.	Eta Squared
Corrected Model	193.965 <sup>a</sup>	2	96.983	24.167	0.000	0.311
Intercept	90.841	1	90.841	22.637	0.000	0.175
Pre-Test Total	193.594	1	193.594	48.242	0.000	0.311
Group	0.298	1	0.298	0.074	0.786	0.001
Error	429.390	107	4.013			
Total	14579.000	110				
Corrected Total	623.355	109				

a R Squared = .311 (Adjusted R Squared = .298)

### IMPROVED IDENTIFICATION ACCURACY BY THE CONTROL GROUP SYLLABUS

To investigate whether there were any individual sounds among the 16 tested sounds that showed statistically significant mean differences, the 16 post-test scores were processed using ANCONA. To control for the possibility of Type I error inflation, the Bonferroni correction was applied to the  $\alpha$  level set by the researcher to detect statistically significant results. Therefore, the  $\alpha$  level of 0.10 was adjusted to 0.006 (0.10 divided by 16).

According to the ANCOVA results, there was only one sound that showed significantly different mean scores between the groups. The ability to identify the word 'bid,' which contained /ɪ/, turned out to show a statistically higher mean in the control group ( $p < 0.006$ ) as shown on Table 11. The participants were requested to choose the right answer based on what they heard as follows:

Wasn't it a wonderful (bid)?  
a. bid   b. bead   c. 모르겠음 (I don't know.)

The test result of the word 'bead,' which contained /i/ was also presented on the same table, even though it failed to show a statistically higher mean difference. The question was presented similarly to the participants as follows:

Wasn't it a wonderful (bead)?  
a. bid   b. bead   c. 모르겠음 (I don't know.)

It is interesting to note that the identification scores of the two minimal-paired contrasts within the same sentence yielded different degrees of improvement. This result suggests that the participants might not have relied on a test-taking strategy, for example, to answer, to this question, with the choice, 'bead,' which was the choice directly opposite to 'bid.' If they had relied on such a strategy, the accuracy rates of the two questions above would have been very close to each other, as there were only two choices 'bid' or 'bead' for each of the two contrasted question. Another possibility is that the participants were not able to compare the two questions. Either way, the score obtained from the questions on the word 'bead' was not likely to have been entirely influenced by the confounding factor of test-taking strategy.

A close look into the mean score changes on Table 11 appears to provide a possible reason for the unsymmetrical accuracy improvements between the two contrasts. The identification ability of the control group participants on the word 'bead' showed quite a high accuracy rate of 80% even on the pre-test. Therefore, it may be possible that the post-test score of this item, 92%, was influenced by the ceiling effect, while the post-score, 77%, of the word 'bid' still had room for improvement.

In sum, these test results point to the possibility that the ability to distinguish between these high front lax and tense vowels, /ɪ/ and /i/ respectively, might be more effectively enhanced through the control group's syllabus. According to the

practical importance indices, the group difference was expected to explain 7 to 10% of the total variance of the differentiation score of those two words in the population.

Table 11

*Identification ANCOVA Results Showing Higher Improvements in the Control Group*

Group	Statistic	<u>bid</u>		<u>bead</u>	
		Pre	Post	Pre	Post
Experimental	<i>N</i>	50	50	50	50
	Mean	0.640	0.540	0.720	0.740
	Std. Deviation	0.485	0.503	0.454	0.443
Control	<i>N</i>	60	60	60	60
	Mean	0.483	0.767	0.800	0.917
	Std. Deviation	0.504	0.427	0.403	0.279
Entire	<i>N</i>	110	110	110	110
	Mean	0.555	0.664	0.764	0.836
	Std. Deviation	0.499	0.475	0.427	0.372
	Mean <sub>Exp</sub> minus Mean <sub>Con</sub>	0.157	-0.227	-0.080	-0.177
	<i>F</i> (1, 107)	8.148		5.52	
	<i>P</i>	0.005		0.021	
	<i>R</i> Squared	0.088		0.121	
	Adjusted <i>R</i> Squared	0.071		0.104	

It was interesting to note that the experimental group students' average ability to recognize the word 'bid' deteriorated as the results of the treatment. Based on the changes in the means of each instructor's participants (Table 12), this deterioration resulted from the performance of the classes taught by the second Korean instructor of English. However, both the control group instructors' cells presented improvements that could be regarded as over medium effect sizes through the treatment.

Table 12

*Mean Changes on 'bid' in Each Instructor's Cell*

Group	Instructor		bid		
			Pre	Post	Gain
Experimental	K1 <i>n</i> = 22	Mean	0.636	0.727	0.091
		Std. Deviation	0.492	0.456	
	K2 <i>n</i> = 28	Mean	0.643	0.393	-0.250
		Std. Deviation	0.488	0.497	
Control	E1 <i>n</i> = 31	Mean	0.548	0.806	0.258
		Std. Deviation	0.506	0.402	
	E2 <i>n</i> = 29	Mean	0.414	0.724	0.310
		Std. Deviation	0.501	0.455	

#### IMPROVED IDENTIFICATION ACCURACY BY THE EXPERIMENTAL GROUP SYLLABUS

On the other hand, no sound indicated significantly higher accuracy of the Korean participants' identification abilities in the experimental group at the Bonferroni-adjusted  $\alpha$  level of 0.006. However, it may be worthwhile to examine a few cases that showed  $p$  values smaller than 0.10, even though the results should not be generalized into the population.

The words 'breathe' and 'jealos' arose as the two most improved sounds through the experimental syllabus, when compared to the improvement occurred in the control group, as can be seen on Table 13. Similar to the case of deterioration previously mentioned, the control group showed a decrease in the mean over the experiment concerning /ð/ in the word 'breathe.' This negative gain resulted mainly from approximately 36% of deterioration in the classes taught by the first native English-speaking instructor according to Table 14. It appears to have cancelled off the desirable achievement (21%) achieved by the other native English-speaking teacher. In addition, the experimental cells made either no or a trivial level of improvement, at 0% and 7%, respectively. This descriptive analysis suggests that even the slightly greater accuracy observed in the experimental group can hardly be considered meaningful.

On the other hand, the word 'jealos' offered a more optimistic result in terms of descriptive statistics. That is, according to Table 15, the cell-wise



improvement tendency was generally similar to the group-wise tendency across the four instructor cells

Table 13

*Identification ANCOVA Results Showing Higher Improvements in the Experimental Group*

Group	Statistic	breath <u>e</u>		jealous		zealous	
		Pre	Post	Pre	Post	Pre	Post
Experimental	<i>N</i>	50	50	50	50	50	50
	Mean	0.600	0.640	0.680	0.860	0.680	0.820
	Std. Deviation	0.495	0.485	0.471	0.351	0.471	0.388
Control	<i>N</i>	60	60	60	60	60	60
	Mean	0.467	0.383	0.650	0.683	0.733	0.700
	Std. Deviation	0.503	0.490	0.481	0.469	0.446	0.462
Entire	<i>N</i>	110	110	110	110	110	110
	Mean	0.527	0.500	0.664	0.764	0.709	0.755
	Std. Deviation	0.502	0.502	0.475	0.427	0.456	0.432
	Mean <sub>Exp</sub> minus Mean <sub>Con</sub>	0.133	0.257	0.030	0.177	-0.053	0.120
	<i>F</i> (1, 107)	5.907		4.689		2.573	
	<i>P</i>	0.017		0.033		0.112	
	<i>R</i> Squared	0.133		0.078		0.060	
	Adjusted <i>R</i> Squared	0.117		0.060		0.042	

The identification result of ‘zealous,’ which is often erroneously perceived as ‘jealous’ by Korean learners of English (see Table 16) was considered to be in the

same direction as the improvement observed in 'jealous.' It is notable that the experimental participants made slightly lower improvement in identification accuracy on this FL sound /z/, roughly 14%, than their improvement (roughly 18%) in identification accuracy with the phoneme /dʒ/. This held true even in the same class taught by the same instructor. Again, however, these results should only be considered when describing what happened in the specific group of participants in the study, not when talking about the general tendency of the intended population.

Table 14

*Mean Changes on 'breathe' in Each Instructor's Cell*

Group	Instructor		breathe		
			Pre	Post	Gain
Experimental	K1 <i>n</i> = 22	Mean	0.773	0.773	0.000
		Std. Deviation	0.429	0.429	
	K2 <i>n</i> = 28	Mean	0.464	0.536	0.071
		Std. Deviation	0.508	0.508	
Control	E1 <i>n</i> = 31	Mean	0.516	0.161	-0.355
		Std. Deviation	0.508	0.374	
	E2 <i>n</i> = 29	Mean	0.414	0.621	0.207
		Std. Deviation	0.501	0.494	

Table 15

*Mean Changes on 'jealous' in Each Instructor's Cell*

Group	Instructor		jealous		
			Pre	Post	Gain
Experimental	K1 <i>n</i> = 22	Mean	0.727	0.909	0.182
		Std. Deviation	0.456	0.294	
	K2 <i>n</i> = 28	Mean	0.643	0.821	0.179
		Std. Deviation	0.488	0.390	
Control	E1 <i>n</i> = 31	Mean	0.677	0.710	0.032
		Std. Deviation	0.475	0.461	
	E2 <i>n</i> = 29	Mean	0.621	0.655	0.034
		Std. Deviation	0.494	0.484	

In sum, the identification test result, as a whole, yielded no statistically significant mean difference between the experimental and control groups. The only statistically significant mean difference was observed in the identification task of /ɪ/ as in 'bid' conducted by the control group participants, who were taught by native English-speaking instructors.

Table 16

*Mean Changes on 'zealous' in Each Instructor's Cell*

Group	Instructor		zealous		
			Pre	Post	Gain
Experimental	K1 <i>n</i> = 22	Mean	0.727	0.864	0.136
		Std. Deviation	0.456	0.351	
	K2 <i>n</i> = 28	Mean	0.643	0.786	0.143
		Std. Deviation	0.488	0.418	
Control	E1 <i>n</i> = 31	Mean	0.677	0.742	0.065
		Std. Deviation	0.475	0.445	
	E2 <i>n</i> = 29	Mean	0.793	0.655	-0.138
		Std. Deviation	0.412	0.484	

**L2-WORD READING-ALoud TEST RESULTS****VALIDITY AND RELIABILITY VERIFIED**

Tukey's test of non-additivity was conducted in order to see if the pilot score samples, which were used to examine the researcher's grading validity, met one important assumption of the ones recommended by (Garson, 2006) for valid intra-class correlation coefficients (ICC) in advance, given that other assumptions were generally met. Items on both the pre-test samples ( $n = 6$ ,  $N$  of cases = 204) and post-test samples ( $n = 12$ ,  $N$  of cases = 408) turned out to possess linear relations ( $p$

> 0.05) to the total scores, according to the non-additivity tests, convincing that the ICC was plausible for the samples.

The post-test pilot scores, not the pre-test scores, assigned by the researcher and the American informant were processed first because the informant, unaware of which set were pre- or post- test recordings, happened to start grading the post-test scores first. The first six samples were discarded due to the different degrees to which these two graders felt familiar with grading. That is, because the researcher had graded considerably more of the recorded samples, while the American grader had just started grading them on her own, the first few samples were skewed. As shown in Table 17, the single measure<sup>25</sup> ICC of the post-test scores was 0.8190, which could be deemed to have a very high reliability<sup>26</sup>. Similarly, the 12 participants' pre-test scores also turned out to be very highly reliable (0.761) according to Table 18. According to the ICC results and the researcher's justification previously mentioned in Chapter 3, his pronunciation grading pattern was considered to be as valid as the American informant's.

The test-retest reliability test was conducted in order to investigate how reliable, or in other words, how consistent the researcher's grading was over time. This measure was important because it took more than one exhausting week to grade the entire recorded samples that included over 10,000 words that were read-aloud by the 94 participants in the four reading-aloud tests. ICC was calculated

---

<sup>25</sup> The single measure, rather than the average measure, ICC was considered because the scores were not averaged, say, from more than one grader.

<sup>26</sup> ICCs between 0.7 and 0.9 are considered as having a very high reliability.

between randomly chosen graded samples as *test* items and the scores assigned after a certain period of time by the same researcher to the same sample as the *retest* items. A total of 10 samples ( $N$  of cases = 350) graded twice between a week's term were used to investigate the test-retest reliability of non-sense word grading. As can be seen on Table 19, the ICC for artificial-word grading indicated a very high reliability of 0.8538. The same procedure was applied to a set of real-word random samples ( $n = 5$ ,  $N$  of cases = 170), and the ICC (0.7506) also fell within the boundary of very high reliability as shown on Table 20.

Table 17

*Source Table of Intraclass Correlation Coefficient for Pilot Post-Test Items*

---

Two-way Random Effect Model (Absolute Agreement Definition):

People and Measure Effect Random

---

Single Measure Intraclass Correlation = .8190\*

95.00% C.I.: Lower = .7680 Upper = .8596

$F = 10.1538$   $DF = (203, 203.0)$  Sig. = .0000 (Test Value = .0000 )

Average Measure Intraclass Correlation = .9005

95.00% C.I.: Lower = .8687 Upper = .9245

$F = 10.1538$   $DF = (203, 203.0)$  Sig. = .0000 (Test Value = .0000 )

---

\*: Notice that the same estimator is used whether the interaction effect is present or not.

Table 18

*Source Table of Intraclass Correlation Coefficient for Pilot Pre-Test Items*

Two-way Random Effect Model (Absolute Agreement Definition):

People and Measure Effect Random

---

Single Measure Intraclass Correlation =	.7621*		
95.00% C.I.:	Lower =	.7123	Upper = .8035
$F =$	7.6816	$DF = (407, 407.0)$	Sig. = .0000 (Test Value = .0000 )
Average Measure Intraclass Correlation =	.8650		
95.00% C.I.:	Lower =	.8314	Upper = .8914
$F =$	7.6816	$DF = (407, 407.0)$	Sig. = .0000 (Test Value = .0000 )

---

\*: Notice that the same estimator is used whether the interaction effect is present or not.

Table 19

*Source Table of Test-Retest Reliability (ICC) for Artificial-Word Grading*

Two-way Random Effect Model (Absolute Agreement Definition):

People and Measure Effect Random

---

Single Measure Intraclass Correlation =	.8538*		
95.00% C.I.:	Lower =	.8226	Upper = .8798
$F =$	12.6667	$DF = ( 349, 349.0)$	Sig. = .0000 (Test Value = .0000 )
Average Measure Intraclass Correlation =	.9211		
95.00% C.I.:	Lower =	.9027	Upper = .9361
$F =$	12.6667	$DF = ( 349, 349.0)$	Sig. = .0000 (Test Value = .0000 )

---

\*: Notice that the same estimator is used whether the interaction effect is present or not.

Table 20

*Source Table of Test-Retest Reliability (ICC) for Real-Word Grading*

---

Two-way Random Effect Model (Absolute Agreement Definition):

People and Measure Effect Random

---

Single Measure Intraclass Correlation =	.7506*		
95.00% C.I.:	Lower =	.6766	Upper = .8095
<i>F</i> = 6.9842	<i>DF</i> = ( 169, 169.0)	Sig. = .0000	(Test Value = .0000 )
Average Measure Intraclass Correlation =	.8575		
95.00% C.I.:	Lower =	.8071	Upper = .8947
<i>F</i> = 6.9842	<i>DF</i> = ( 169, 169.0)	Sig. = .0000	(Test Value = .0000 )

---

\*: Notice that the same estimator is used whether the interaction effect is present or not.

### OVERVIEW OF THE READING-ALoud TEST RESULTS

A total of 93 participants succeeded in recording their voices on all the four required reading-aloud tests. For proper comparison, all the recorded samples of the participants who missed any of the pre-real-word, pre-artificial-word, post-real-word, and post-artificial-word recordings were excluded. The entire participant pool consisted of 41 sets of experimental samples and 52 sets of control samples. The graded reading-aloud scores were processed through the analysis of covariance



(ANCOVA) with the covariate of pre-test scores using the Univariate General Linear Model function available in SPSS.

Unfortunately 7 sounds out of the 18 originally tested sounds were regarded as invalid by the researcher, who was the grader of the pronunciation tests. The reason for this perceived invalidity is that the words that contained those 7 sounds involved a great deal of repetition when they were heard and graded. For example, the grader had to listen repeatedly to the words such as ‘feyer’ and ‘their’ a number of times, in order to determine a proper accuracy index. It should be noted that the grading procedure could be considered as decently straightforward because the grader was equipped with the well-articulated grading criteria examined and negotiated with the help of the American informant. Besides, the range of possible scores was only 3, which was felt to be fairly simple and intuitive, and these levels score were concretely pre-designated to the specific pronunciation patterns typically observable in Korean learners. The grader, in addition, needed to focus on only one sound in each word.

There were two more courses of action put into place as insurance. The grader recorded his own readings of the tested words and let the informant grade them without telling her that they were recorded by the grader. Out of 118 points, he earned 115 points (97.5%) as his total score. There were three less than perfectly pronounced words ‘their,’ ‘zipper,’ and ‘mirage,’ containing [ð], [ɪ], and [ʒ], respectively. The American informant gave an accuracy index of 2, which meant “imperfect but close to nativelike,” for these three words spoken by the grader and

she marked 3, “nativelike,” for the remaining 35 words recorded by the grader. The reasons for the imperfect pronunciation patterns of the NNEST in these three words were the voice-onset time ([ð]) and impression ([ɪ] and [ʒ]), according to the informant. Based on these results, the words that involved the three sounds as their principal sound were excluded. Therefore, the researcher’s pronunciation patterns of the 35 remaining words were verified as nativelike or free of foreign accent. The other measure of insurance was the model pronunciation of the read-aloud words recorded by the American informant. The researcher relied on these model recordings when he faced recording samples that were uneasy to score, by comparing the model and the uncertain samples.

The possible reasons that the grader felt indeterminate of what sound he heard even after several instant repetitions seemed to be three fold. First, even with the reliability and validity measures and the calibration sessions, only a relatively inferior ear for nativelikeness judgment was available to the grader, a non-native English-speaking teacher (NNEST), possibly due to his L1 background. After the researcher finished grading the entire collection of recordings, he double-checked how certain and confident he was when he was grading each of the tested words, by recollecting the experience he had had during the grading sessions. He determined that he was not very confident in the scores given to the words ‘filler,’ ‘meter,’ and ‘seal,’ which contained the lax and tense high front vowels of English. Except for a few samples of this sort that were regarded as highly nativelike, many

of interlanguage patterns were graded with little confidence. Thus, he excluded these words from the analysis.

Second, unlike in natural face-to-face communication situations, the NNEST grader was not able to rely on some non-acoustic clues such as the movement of lips and teeth when grading [v] or [ð] as in 'feyer' or 'their,' because the words were only audio-recorded, not video-taped. For instances of pronunciation that deserved 3 or 0 point, the researcher's judgment could be made with the fullest confidence. However, there seemed to be more room for error variance in the scores assigned to sound patterns that might potentially fall between 3 and 1, which often involved idiosyncratic sounds that were perceived not to belong to the L1 nor the L2 sound system. That is to say, the indecision that the grader had to experience might be caused not only by his insufficient intuition on pronunciation, but also by the innate nature of idiosyncrasy that various stages of interlanguage presented. The fact that the native English-speaking informant had to spend an even longer time in repeating such words, than the NNEST grader until she reached a decision, suggests the plausibility of the speculation on the reasons for the indecision.

Thirdly, there were three words that showed considerable discrepancy in terms of IPA transcriptions between the NNEST and the American informant. They were 'flying,' 'mammoth,' and 'summer' that contained the sounds [aɪ], [m], and [m] respectively. The NNEST and the informant agreed on their grading with each other when they heard perfectly enunciated samples or definitely incorrect

instances of those words. However, they thought they heard different sounds when they faced some interlanguage instances of the three words. For example, for the word ‘flying’ involved two ensuing [ɪ] sounds. In this case, the NNEST believed he definitely heard the imperfect pronunciation [flaɪɪŋ], while the informant thought she heard the acceptable enunciation [flaɪɪŋ] for the same sample.

This overly sensitive judgment of the researcher can be explained as a hypercorrection tendency that the researcher formed through his EFL learning. The words ‘mammoth’ and ‘summer,’ which contained two ensuing consonant letters, points to an interesting insight into L2 nativelikeness. It has been quite widely appreciated, by Korean EFL teachers of English, that even though these two words have two *ms*, they should be spoken with only one *m*, for example as [ˈsʌm.ɪ]. Similarly, the researcher believe that English teachers need to explicitly advise Korean students not to replace two ensuing English vowels, such as [ɪi] as in ‘employee,’ with a single vowel [i] (refer to Appendix J for an example). In order to detect Korean EFL speakers’ tendencies to say these words with two [m] sounds in them ([ˈsʌm.m.ɪ]), the researcher explained this phenomenon to the American informant using the two above-mentioned examples. However, she seemed far less sensitive to the two same consonants, which were a phonologically unobserved, but orthographically possible, combination, in English, when she was grading these words with the researcher.

This tendency was perceived by the researcher as also attributable to the two graders' different L1 backgrounds. In Korean, both ['sʌm̩ɾ] and ['sʌm.m̩ɾ] can be spoken even with no pause between the two [m] sounds, while most adult English speakers do not tend to be as sensitive to this distinction as Koreans are. English speakers' tendency to contract double consonants into one consonant sound, as observed in "I hate taking a bath in a *hot tub*" ([hʌt<sup>h</sup>ʌb]) or even "I live on *First Street*" ([fɪst.ɪt<sup>ɪ</sup>]), confirms this hypothesis. In fact, it can be said that perceiving the distinction between single consonants and double consonants, such as ['sʌm̩ɾ] and ['sʌm.m̩ɾ], may be unnecessary in English because the presence or lack of double consonants does not convey meaning. 'Summer' pronounced with two /m/s does not differ in meaning from 'summer' pronounced with one /m/. Therefore, whether or not they hear two same consonants next to each other, native English speakers would neither consciously nor unconsciously pay attention to the difference. The researcher let four other native English-speakers listen to the minimal pair, "서머" (['sʌm̩ɾ]) and "섬머" (['sʌm.m̩ɾ]). They responded that they heard the same word twice.

Setting aside the discussion of the distinction between ['sʌm̩ɾ] and ['sʌm.m̩ɾ], the words in the test containing such double consonant letters, 'summer' or 'mammoth,' appeared to present little educational implication when included in the reading-aloud test. This is because the researcher was needlessly sensitive to the

possible Korean-accented variations of the words, which could otherwise be realized by NESTs as acceptable.

The researcher originally created the artificial words that were comparable to their corresponding real words. For instance, the non-sense word ‘nesco’ was invented in order to investigate whether the participants became able to apply what they learned about the word ‘disco.’ With the help of the two NESTs who participated in the current study, these non-sense words were double-checked if they actually appeared to be English and whether the designated principal sounds were likely to be pronounced by English speakers, as desired for the current study. However, the artificial word ‘plaw’, which was meant to be compared with the unpaired word ‘coleslaw’ on Table 21, was perceived as potentially being pronounced either as [p<sup>h</sup>lɔ] or [p<sup>h</sup>lɑu] by the Texan informant. This was how the non-sense word ‘plaw’ was removed from the list of words to be analyzed. For its ambiguity in pronunciation, the real word ‘robot’ ([<sup>1</sup>.ɪoubat<sup>2</sup>], [<sup>1</sup>.ɪoubət<sup>2</sup>], or [<sup>1</sup>.ɪoubət<sup>2</sup>]), which was originally devised to be the counterpart to ‘kot,’ was also discarded.

Regardless of how insensitive the native English-speaking graders were to such double vowels or consonants, the remaining 11 tested sounds included in the 19 words were perceived to have been graded by the researcher with concrete judgment and were processed for ANCOVA.

When there was more than one word per principal sound, the scores of them were averaged. However, similar to the argument on the ineligibility of the identification and reading-aloud tests as reliable scales, these average scores might not exhaustively reflect the students' performance on the focused sound in all the possible contexts<sup>27</sup>. For example, in order to test [kw], contrastive contexts such as the one in 'quarter<sup>28</sup>,' which occurred before a mid back vowel, might make the comparison more exhaustive, compared to the currently tested word 'quiz,' which was placed before a high front vowel.

In what follows the ANCOVA results on the 5 vowels and 6 consonants are presented. A total of 22 mean comparisons of these 11 tested sounds were made using the 2 kinds of words, real and artificial words (see Table 21). It should be noted that each presented *p* value is supposed to concern only the inference made on the possible post-test mean differences between the two imaginary populations, given that they are treated with the experimental and control treatment in the current study. Therefore, the discussions on the gain scores, which are related to improvements, and the partial Eta squares, which regard the ratio of the variance due to the group membership to the total variance, need to be considered only as descriptive explanations, not as inferential predictions.

---

<sup>27</sup> For example, in order to test [kw], there might be more contrastive contexts such as [kw] in 'earthquake,' which occurred before a diphthong or mid vowel, compared to 'quiz,' which was placed before a high vowel.

<sup>28</sup> The pilot study (Huh, 2003) showed greater gain scores in the words 'quarter' (+1.5), 'question' (+0.9), and 'quiz' (+0.8), compared to 'quake' (+0.3), indicating that different contexts might yield different syllabus effects even with the same principal sound.

Table 21

*Analyzed Sounds and Words in the Word Reading-Aloud Tests*

	Principal Sound	Real Word	Artificial Word
Vowel	[i]	well-be <u>ing</u>	ke <u>eing</u>
	[e]	la <u>bel</u>	na <u>bel</u>
		pro <u>pane</u>	tro <u>pane</u>
	[o]	coco <u>a</u>	pl <u>oat</u>
		disco <u>o</u>	nesco <u>o</u>
	[ɔ]	coles <u>law</u>	
		tho <u>ught</u>	pho <u>ught</u>
	[ɑ]	jo <u>gging</u>	mo <u>gging</u>
			ko <u>t</u>
	Consonant	[f]	f <u>ashion</u>
		f <u>ile</u>	f <u>ime</u>
[t]		e-ma <u>il</u>	la <u>il</u>
		sm <u>ile</u>	sn <u>ile</u>
[z]		organi <u>zation</u>	trapi <u>zation</u>
		zo <u>o</u>	zo <u>ul</u>
[r]		pre <u>tt</u> y	dre <u>at</u> y
		wa <u>t</u> er	na <u>t</u> er
[ɹ]		ca <u>r</u>	na <u>r</u>
		ru <u>n</u> ning	ru <u>p</u> ping
[kw]	qu <u>iz</u>	qu <u>eet</u>	

**TOTAL SCORE OF THE READING-ALLOUD TESTS**

Prior to ANCOVA analyses, tests of between-subjects effects were conducted in order to check whether there was significant interaction between the independent variable, group, and the covariate, pre-test score. Because there was no significant interaction effect between these two factors (group X pre-test:  $p = 0.646$  for real-word total score;  $p = 0.659$  for artificial-word total score), ANCOVA was valid for the reading-aloud data.



Table 22

*Source Table for the Real-Word Reading-Aloud Total Score*

Tests of Between-Subjects Effects							
Dependent Variable: POSREATO							
Source	Sum of Squares	df	Mean Square	F	Sig.	Eta <sup>2</sup>	
Corrected Model	748.146 <sup>a</sup>	2	374.073	38.352	0.000	0.460	
Intercept	185.264	1	185.264	18.994	0.000	0.174	
PREREATO	504.428	1	504.428	51.717	0.000	0.365	
GRPCODE	184.509	1	184.510	18.917	0.000	0.174	
Error	877.827	90	9.754				
Total	38486.750	93					
Corrected Total	1625.973	92					

a R Squared = .460 (Adjusted R Squared = .448)

The reading-aloud test results indicated that the experimental participants achieved a statistically significant higher mean on both real-word and artificial-word tests ( $p < 0.0001$ ) as can be seen on Table 22 and Table 23. 45% of the entire variance of the real-word reading-aloud total score turned out to be accounted for by the group membership, experimental or control approach to teaching pronunciation. Similarly, on the non-sense word production test, 56% of the total variance was explained by the different teaching approaches.

Table 23

*Source Table for the Artificial-Word Reading-Aloud Total Score*

Tests of Between-Subjects Effects						
Dependent Variable: POSREATO						
Source	Sum of Squares	<i>df</i>	Mean Square	F	Sig.	Eta <sup>2</sup>
Corrected Model	666.978 <sup>a</sup>	2	333.489	60.146	0.000	0.572
Intercept	137.634	1	137.634	24.823	0.000	0.216
PREREATO	401.587	1	401.587	72.427	0.000	0.446
GRPCODE	236.286	1	236.286	42.615	0.000	0.321
Error	499.022	90	5.545			
Total	32995.250	93				
Corrected Total	1166.000	92				

a. R Squared = .572 (Adjusted R Squared = .563)

This result means that we can reject the null hypothesis that the two different teaching approaches are likely to yield no mean difference in terms of L2 production achievement. Therefore, it can be said that the same teaching techniques may lead to a higher mean score in the experimental group at 90 times if they are applied to the intended population at 100 times.

Table 24

*Descriptive Statistics of the Reading-Aloud Total Score*

Types of Words	Group	Mean	Std. Deviation	<i>N</i>
Real	Experimental	21.732	4.521	41
	Control	18.471	3.328	52
	Total	19.909	4.204	93
Artificial	Experimental	20.402	3.536	41
	Control	17.000	2.802	52
	Total	18.500	3.560	93

Table 24 shows the means of the reading-aloud test's total score on the post-test. On the real-word test, the experimental group participants obtained a total score of 21.7 at the average, while the control group participants obtained a total score of 18.5 out of the perfect score 33. They showed a consistent achievement pattern also in terms of the reading-aloud accuracy of the L2 sounds contained in the 11 non-sense words, 20.4 and 17.0, respectively. The experimental groups' statistically higher achievement yielded on the artificial-test suggests that the real-word achievement is hardly considered to have resulted from the memorization effect of the words that are presented to students as pronunciation models.

Table 25

*Estimated Marginal Means of the Reading-Aloud Total Score*

Type of Words	Group	Mean	Accuracy Index	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Real	Experimental	21.501 <sup>a</sup>	1.955	0.489	20.530	22.472
	Control	18.653 <sup>a</sup>	1.696	0.434	17.791	19.515
	Difference	2.848	0.259			
Artificial	Experimental	20.297 <sup>b</sup>	1.845	0.368	19.566	21.028
	Control	17.083 <sup>b</sup>	1.553	0.327	16.434	17.732
	Difference	3.214	0.292			

a Evaluated at covariates appeared in the model: PREREATO = 18.5215.

b Evaluated at covariates appeared in the model: PREREATO = 17.0054.

The estimated marginal mean scores were calculated in order to remove the possible effects of unequal pre-test scores from the observed post-test mean scores. On the real-word post-test the experimental group is likely to attain 21.5, which is 2.0 in terms of the accuracy index that ranges from 0 to 3 (refer to chapter 4 for the descriptions of the accuracy indices), given that their initial pronunciation accuracy is not considerably different according to Table 25. The accuracy index of 2.0 has been defined as “Clear but Unnaturally Enunciated” or “Imperfect but Close to Nativelike.” On the other hand, the control group is likely to attain 18.7, which is 1.7 in terms of the accuracy index. This accuracy index falls between the category

of “Intelligible but Foreign Accented” (Accuracy Index 1) and the aforementioned category of the Accuracy Index 2.

Even in the unlikely event that the experimental group would indicate the lower bound score of 20.5 and the control group would show the upper bound score of 19.5, it can be said that the experimental group’s post-score would not be lower than the control groups’, based on a 95% confidence interval.

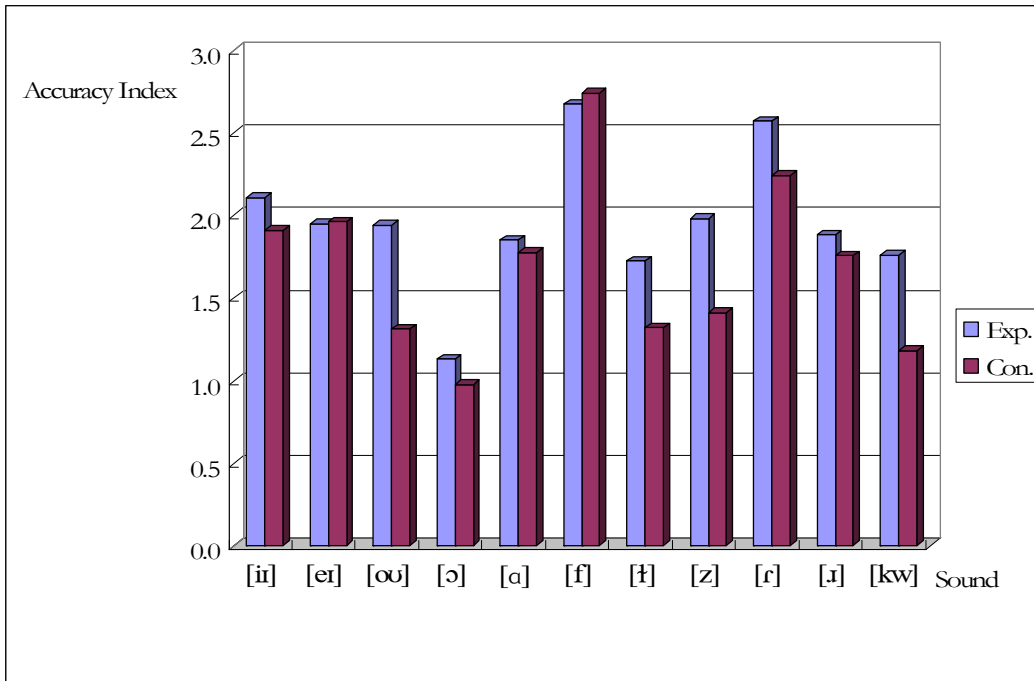
#### **ACHIEVEMENT PATTERNS IN THE INDIVIDUAL SOUNDS**

It is worthwhile to investigate the participants’ achievement patterns concerning the 11 individual sounds as they may shed light on relatively promising segmental English sounds. To do this, ANCOVA tests were conducted for each of the individual sounds. Accordingly, the  $\alpha$  value 0.10 that was set by the researcher as the basis to determine statistical significance needed to be reset to 0.005 (0.10/22).

As shown in Figure 1, in the analyses of the post-test estimated marginal means, the experimental groups showed generally higher means in the post-test scores that were processed in a way that the effect of the pre-test scores was taken into account.

Figure 1

*Estimated Marginal Mean Comparisons of Individual Sounds*



**CLASSIFICATION OF OBSERVED PATTERNS OF INDIVIDUAL SOUNDS**

Generally speaking, the ANCOVA results of the reading-aloud scores were stratified into the following four categories, based on their educational implications. First, some of the sounds revealed practically important as well as highly inferential improvements (see Category A on Table 26). The second category involved improvements that indicated statistically higher means on the real-word test, but not on the artificial-word test. These sorts of results might still be regarded as positive in that the population might show a similar higher mean difference in the

sounds' scores contained at least in those specific words that were tested, as outlined in Category B on Table 26. A second optimistic result concerned relatively high pre- and post- test scores even with no statistically significant mean differences between the two groups as can be seen in Category C on Table 26. Results of this sort deserve some attention because it was likely that the participants had generally achieved a certain desirable level of accuracy before the treatment. As presented in Category D, the fourth category depicted cases in which a certain significant ANCOVA *p* value was observed, but it hardly bore meaningful teaching implications. This category also included a case that presented no gain in one group's real word-readings while the other group's real-word gain was negative, and there was a statistically significant higher gain observed on the artificial word test. In other words, the artificial words were more accurately pronounced on the post-test than on the pre-test despite the fact that there was no improved knowledge to transfer.

#### **SOUNDS WITH SIGNIFICANTLY HIGHER MEANS AND POSSIBLE KNOWLEDGE TRANSFER**

Observed improvements were categorized as this kind (see Category A on Table 26) when there were statistically significant mean differences both on the real and artificial word tests consistently favoring one of the two groups' syllabi. This category was considered to be the most desirable result in terms of its effect size and transfer of the knowledge that was formed through the syllabus. As seen from

Table 27, out of the 11 tested sounds, two sounds turned out to show statistically significant mean differences, both on the real and artificial word-reading tests.

Table 26

*Interpretation of Reading-Aloud Test Results*

Category	Type of Words	Mean Change <sup>29</sup>		<i>p</i> Value	Teaching Implication
		EXP	CONT		
A	Real	++	+ or 0	significant	desirable ++, transfer
	Artificial	++	+ or 0	significant	
B	Real	++	+ or 0	significant	desirable +, no transfer
	Artificial	+	+	insignificant	
C	Real	0	0	insignificant	unproblematic with high means, problematic when not
	Artificial	0	0	insignificant	
D	Real	0	-	insignificant	questionable
	Artificial	+	0	significant	

These relatively salient improvements were observed in the sounds [z], and [kw], in favor of the experimental group's syllabus. As seen on Table 27, the *p* value detected in the means of these two sounds on both the real and artificial word-tests were close to 0, except for artificial words containing /kw/, the *p* value of which case was 0.001. As all the real-word and artificial-word cases yielded *p* value smaller than the Bonferroni-adjusted  $\alpha$  value, 0.005, the null hypotheses set up for these two sounds were discarded. In other words, we can infer that the same

<sup>29</sup> The notations '+-' represent a positive and negative gain observed in the pre- and post- test comparisons respectively. The double positive sign, '+ +,' denotes that the detected gain was greater than in the ones marked with '+-' The mean changes with '0' showed almost no gain.



treatment, if administered to the defined population, might yield relatively high mean scores as the result of the experimental group's syllabus, compared to the scores resulting from the control group's syllabus at 90 times out of 100.

Table 27

*Consistent ANCOVA Results of the Real-Word and Artificial-Word Reading-Aloud Tests*

Principal Sound	Type of Words	Experimental			Control			ANCOVA	
		Pre	Post	Gain ( $d^{30}$ )	Pre	Post	Gain (d)	$p$	Eta2
[z]	Real Mean	1.35	1.96	0.62	1.44	1.43	-0.01	0.000	0.22
	STD	0.49	0.65	(96%)	0.57	0.51	(-2%)		
	Artificial Mean	1.32	1.81	0.49	1.39	1.29	-0.10	0.000	0.18
	STD	0.56	0.74	(67%)	0.59	0.56	(-17%)		
[kw]	Real Mean	1.32	1.81	0.49	1.19	1.15	-0.04	0.000	0.16
	STD	0.72	0.99	(49%)	0.56	0.46	(-8%)		
	Artificial Mean	1.17	1.81	0.64	1.10	1.21	0.12	0.001	0.11
	STD	0.67	0.99	(64%)	0.45	0.64	(18%)		

The sound [z] showed, in the experimental samples, the most improvement in terms of effect size, 96% of the standard deviation (STD). The independent variable with the two different pronunciation teaching approaches explained 16% to 22% of the total variance in the scores of the real word tests, as for the nativelikeness of [z], and [kw], as shown on Table 27.

The correlation analyses between the real and artificial word reading scores seems to support the optimistic ANCOVA results of the sounds [kw] and [z]. The

<sup>30</sup> ' $d$ ' means the effect size.

scores that the experimental participants acquired from the tests of the real word 'quiz' and the non-sense word 'queet' yielded a statistically significant correlation coefficient of 0.56 (at the 0.01 level, 2-tailed). Based on this coefficient, 31% ( $= 0.56^2$ ) of the total variance is explained by the consistency each participant showed in getting the two types of word reading scores. The average score of the sound [z] showed a smaller coefficient, 0.38 (at the 0.05 level, 2-tailed), which means only 14.4% of the total variance is explained by the correlation. This correlation seems to have resulted mainly from the improvement of the sound in the word initial positions as in 'zoo' and 'zoul,' ( $r = 0.34$  at the 0.01 level, 2-tailed), rather than from the improvement of the same fricative in the word-medial positions as in 'organization' and 'trapization.' This is because the word-medial /z/ showed no statistically significant correlation. Based on these correlation analysis results and the ANCOVA results of the individual real words, the sounds [z] in 'zoo' and [kw] in 'quiz' showed the most consistent improvement results favoring the experimental group's teaching approach.

#### **SOUNDS WITH SIGNIFICANTLY HIGHER MEANS BUT NO KNOWLEDGE TRANSFER**

As seen from Table 28, the achievement suggested by the statistical significance observed in the real-word test result of [ou] was not confirmed by its corresponding artificial word-test result. The improvement of producing [ou] in the tested real words indicated a relatively strong effect size, 94% of the STD, compared to the improvement achieved by the control group, 18% of the STD.

This result was statistically significant. However, the  $p$  value obtained through the ANCOVA calculation of the artificial-word scores was too small to be statistically confirmed ( $p = 0.05 > \alpha = 0.005$ ). Therefore, the null hypothesis set up for the novel-word production test was not rejected. It is probable that the experimental participants have achieved a higher mean score than the mean score that control group achieved on the real-word test because they had become more familiar with the words given to them as pronunciation models than the control group. At best, the experimental participants are likely to be perceived as being more able to produce the diphthong [ou] natively when they pronounce the tested words than the control group would be.

Table 28

*Inconsistent ANCOVA Results of the Real-Word and Artificial-Word Reading-Aloud Tests*

Principal Sound	Type of Words		Experimental			Control			ANCOVA	
			Pre	Post	Gain (d)	Pre	Post	Gain (d)	$p$	Eta <sup>2</sup>
[ou]	Real	Mean	1.30	1.95	0.65	1.21	1.30	0.09	0.000	0.24
		STD	0.40	0.70	(94%)	0.46	0.47	(18%)		
	Artificial	Mean	1.09	1.46	0.38	1.16	1.21	0.05		
		STD	0.57	0.71	(53%)	0.52	0.68	(7%)		

### SOUNDS WITH QUESTIONABLE GAIN PATTERNS

Table 29 presents a questionable ANCOVA result. The pattern observed in [ʔ] may generally be characterized with considerable deterioration in the control

group's performance. If we were not to deny the validity of the grading processes, there might be two other sources of degradation in the mean scores. First, the nativelikeness of the control group's pronunciation patterns may have actually degraded. The pilot study, which had no control group, also showed a similar tendency of deterioration for some of the tested words over the treatment administered by the researcher. Out of the 238 real words that were spoken by the four Korean participants, principal sounds in 29 (12.2%) words turned out to be worsened on the post-test (Huh, 2003). One of the common features between the pilot test treatment and the current study's treatment was that they were conducted for a relatively short amount of time, which was only four classroom meetings, excluding other L2 exposure outside of the classroom. In the current study, 6 out of the 7 participants reported in the interview about the current study's treatment that they had little time to digest or internalize what was taught. The other participant also revealed, "Compared to the contents given to learn, the given class period was not felt to be too insufficient." This was also confirmed by the instructor's own comments on the class time.

The participants might have noticed that their pronunciation patterns were different from the model pattern of their native English-speaking instructor in the course of the intensive pronunciation lectures conducted in English. Therefore, there is a possibility that the students' perceived imperfect self-efficacy might have shaken the stability or confidence of the foreign accented pronunciation that they had been comfortable with before the treatment. If the instructor did not properly

sense and handle the participants' ZPD, possibly as a result of insufficient time, the EFL students may have failed to internalize the acceptable pronunciation patterns delivered by the instructor. The shaken self-efficacy coupled with insufficient treatment duration to overcome the confusion could have impacted the students' performance on the post-test. There was a comment made by one of the interviewees that supports this speculation. Two of the four NEST classes performed relatively poorly on the tests compared to the other two classes that were taught by the other NEST. Participant A-07 from one of these two low-performance classes stated, "... the instructor let the students repeat after him over and over again, sometimes with no [effective] instructions on how to pronounce the words." Based on this remark, it appears probable that the NEST may have often skipped explicit instructions on how to pronounce the sounds or words and relied mainly on repetition, which may have saved him a considerable amount of time. In turn, it can be inferred that the four sessions of the pronunciation tutorials were too short for them to *implicitly* internalize pronunciation patterns provided through the *repetition*.

The second discussion of potential reasons for the deterioration observed in the control group relates the grading criteria. While grading these sounds, the researcher noticed some positive changes in the participants' renditions of these two sounds. For example, samples of 'mail' that were pronounced as '이메이어' ([imeɪə]) on the post-test, compared to the previously most observed pattern of '이메일' ([imeil]). This change could be perceived as a desirable interlanguage

variation that was hardly achievable with no explicit instruction. However, without the tip of the tongue touching the gum ridge<sup>31</sup>, the researcher determined that the interlanguage pattern [imeɪə] only deserved 0, due to its drastic substitution of [ɬ] with [ə]. The group membership explained merely 8% of the total variances observed in the real-word reading score concerning [ɬ].

Table 29

*ANCOVA Result With Questionable Gain Patterns*

Principal Sound	Type of Words	Experimental			Control			ANCOVA		
		Pre	Post	Gain (d)	Pre	Post	Gain (d)	<i>p</i>	Eta <sup>2</sup>	
[ɬ]	Real	Mean	1.29	1.68	0.39	1.51	1.38	-0.14	0.007	0.08
		STD	0.84	0.83	(47%)	0.80	0.77	(-18%)		
	Artificial	Mean	0.96	1.36	0.39	1.12	0.96	-0.15	0.002	0.11
		STD	0.54	0.79	(50%)	0.67	0.60	(-25%)		

**SOUNDS WITH NO SIGNIFICANT MEAN DIFFERENCES**

The experimental and control groups' syllabi turned out to have no differential effects on the learning of the sounds [eɪ], [iɪ], [ɑ] and [ɪ] as indicated in Table 30. For these cases, we should not reject the null hypothesis. Nevertheless, it was a relief to discover that the participants' final accuracy indices of these sounds reached somewhere near or above the accuracy index of 2.0, which denoted "imperfect but close to nativelike," regardless of the group they belonged to.

<sup>31</sup> This aspect of articulation was relatively easy to detect while grading.

Table 30

*ANCOVA Results With No Significant Mean Differences*

Principal Sound	Type of Words		Experimental			Control			ANCOVA	
			Pre	Post	Gain (d)	Pre	Post	Gain (d)	<i>p</i>	Eta <sup>2</sup>
[eɪ]	Real	Mean	1.88	1.96	0.08	1.85	1.95	0.11	0.945	< 0.01
		STD	0.92	0.97	(9%)	0.84	0.88	(12%)		
	Artificial	Mean	1.99	2.29	0.30	2.14	2.22	0.07	0.100	0.03
		STD	0.86	0.84	(36%)	0.79	0.88	(8%)		
[ɑ]	Real	Mean	1.76	1.93	0.17	1.37	1.69	0.33	0.692	< 0.01
		STD	0.96	1.00	(17%)	0.74	0.85	(38%)		
	Artificial	Mean	1.45	1.44	-0.01	1.06	0.94	-0.12	0.030	0.05
		STD	0.74	0.89	(-1%)	0.71	0.71	(-16%)		
[ɪɪ]	Real	Mean	1.76	2.07	0.31	1.89	1.94	0.06	0.255	0.02
		STD	0.91	0.95	(33%)	0.90	0.87	(7%)		
	Artificial	Mean	1.71	2.24	0.53	1.79	2.10	0.31	0.324	0.01
		STD	0.87	0.91	(59%)	0.80	0.75	(41%)		
[ɪ]	Real	Mean	1.88	1.89	0.01	1.85	1.75	-0.10	0.330	0.01
		STD	0.78	0.69	(2%)	0.57	0.70	(-14%)		
	Artificial	Mean	1.60	1.94	0.34	1.63	1.61	-0.02	0.006	0.08
		STD	0.64	0.69	(49%)	0.63	0.62	(-3%)		

It should also be noted that the real-word reading-aloud results concerning the rest of the tested sounds, [ɔ], [r], and [f], indicated that ANCOVA analysis was invalid for these cases because there were statistically significant interactions between the group membership and the pre-test score (see Table 31). Refer to Appendix K for the estimated marginal means and ANCOVA interpretations for each of the reading-aloud items.

Table 31

*Descriptive Statistics of the Sounds Showing Significant Group X Pre-Test Interaction*

Principal Sound	Type of Words		Experimental			Control			ANCOVA	
			Pre	Post	Gain (d)	Pre	Post	Gain (d)	<i>p</i>	Eta <sup>2</sup>
[ɔ]	Real	Mean	0.94	1.12	0.18	0.98	0.98	0.00	0.009 <sup>32</sup>	Interaction
		STD	0.50	0.52	(35%)	0.57	0.40	(0%)		
	Artificial	Mean	0.83	1.00	0.17	0.71	0.81	0.10	0.14	0.02
		STD	0.63	0.66	(26%)	0.57	0.56	(17%)		
[r]	Real	Mean	2.52	2.60	0.07	2.33	2.18	-0.14	0.025	Interaction
		STD	0.53	0.48	(15%)	0.63	0.78	(-18%)		
	Artificial	Mean	2.21	2.30	0.09	2.02	2.01	-0.01	0.03	0.05
		STD	0.42	0.50	(18%)	0.44	0.51	(-2%)		
[f]	Real	Mean	2.81	2.75	-0.06	2.64	2.71	0.07	0.033	Interaction
		STD	0.51	0.54	(-11%)	0.59	0.56	(12%)		
	Artificial	Mean	2.84	2.85	0.00	2.78	2.74	-0.04	0.512	< 0.01
		STD	0.44	0.49	(1%)	0.53	0.60	(-7%)		

**RESULTS OF THE LANGUAGE TESTS BASED ON DIFFICULTY LEVELS****RELATIONSHIP BETWEEN IDENTIFICATION RESULTS AND DIFFICULTY LEVELS**

Concerning the last research question, the pre-test, post-test, and gain score of the aural identification test were analyzed using descriptive statistics. The

<sup>32</sup> The *p* values specified as “Interaction” indicate the significance of the interaction term between the factor Group Membership and Pre-Test Score. For the artificial-word test scores, the *p* values with squared Etas are the significance of the valid ANCOVA tests.



tested sounds were classified, by the researcher, into Levels 0, 3, 4, and 5, which were suggested by Prator (1971) as shown on Table 32. This judgment was solely based on the researcher's understanding of Korean and English. Therefore, there could be different classification results. There were no items pertaining to Level 1, Coalescence, and Level 2, underdifferentiation, in the identification test.

Table 32

*Identification Items Categorized Based on Prator's Levels of Difficulty*

Difficulty Level	Principal Sound	Tested Word
0, Transfer		
	[dʒ]	jealous
	[oʊ]	flow
	[ʌ]	cut
	[s]	pass
	[d]	breed
3, Reinterpretation		
	[l]	load
	[r]	road
4, Overdifferentiation		
	[ð]	breathe
	[ə]	path
5, Split		
	[ɪ]	bid
	[i]	bead
	[ɔ]	caught, flaw
	[z]	zealous, mazard
	[ʒ]	measure

The pre-test scores of the participants were believed to reflect their initial identification abilities as the learning outcomes that resulted from the EFL courses in their pre-tertiary schools. According to Table 33, the identification tasks for the words classified into Level 5, Split, before the current study's treatment indicated the lowest accuracy, 58.4%, as predicted by the Hierarchy of Difficulty. Levels 4, 3, and 0 yielded higher means of accuracy ranging from 72.7% to 77.5% on the pre-test than the accuracy shown in the Level 5 sounds.

Table 33  
*Identification Pre-Test Means and Levels of Difficulty*

Instructor	Level 5	Level 4	Level 3	Level 0
K1	0.545	0.864	0.818	0.827
K2	0.587	0.714	0.732	0.707
Exp. Subtotal	0.569	0.780	0.770	0.760
E1	0.585	0.726	0.694	0.845
E2	0.611	0.672	0.690	0.724
Con. Subtotal	0.598	0.700	0.692	0.787
Total	0.584	0.736	0.727	0.775

On the identification post-test, the participants showed accuracy results that more smoothly affirmed the predictions based on Prator's Hierarchy of Difficulty than the pattern that the pre-test accuracy showed. As the levels of difficulty decreased, the post-test identification accuracy increased as can be seen on Table 34.

However, in terms of accuracy gain, Levels 4, 3, and 0 yielded a rather puzzling result (see Table 35). In that particular sample, for the participants of this

study, accuracy improvement was considered to be rather arbitrary with respect to the expected levels of difficulty. For example, Level 5 yielded a higher gain score (0.005) than the gain score observed in the Level 3 items (-0.018). It is notable that the experimental group showed a negative gain in the Level 5 items, while the control group showed a negative gain in the Level 3 items. At least, it can be said that the items at Levels 5 and 4 were more difficult to improve than the items at Levels 3 and 0.

Table 34

*Identification Post-Test Means and Levels of Difficulty*

Instructor	Level 5	Level 4	Level 3	Level 0
K1	0.545	0.886	0.818	0.891
K2	0.556	0.696	0.768	0.800
Exp. Subtotal	0.551	0.780	0.790	0.840
E1	0.590	0.565	0.790	0.781
E2	0.655	0.776	0.776	0.848
Con. Subtotal	0.621	0.667	0.783	0.813
Total	0.590	0.718	0.786	0.825

Table 35

*Identification Gain-Score Means and Levels of Difficulty*

Group	Level 5	Level 4	Level 3	Level 0
K1	0.000	0.023	0.000	0.064
K2	-0.031	-0.018	0.036	0.093
Experimental	-0.017	0.000	0.020	0.080
E1	0.005	-0.161	0.097	-0.065
E2	0.044	0.103	0.086	0.124
Control	0.024	-0.033	0.092	0.027
Total	0.005	-0.018	0.059	0.051

**RELATIONSHIP BETWEEN READING-ALoud RESULTS AND DIFFICULTY LEVELS**

There were only two levels of difficulty that were assigned to the analyzed sounds in the reading-aloud test. Of the 11 tested sounds, six were classified into Level 3, Reinterpretation, while the other five sounds were considered to present the difficulty of Level 5, Split as can be seen from Table 36.

As can be seen from Table 37, on the pre-test the experimental group turned out to have performed better in the production of the sounds classified as Level 3 than in the production of Level 5 sounds by approximately a third of the standard deviation. However, this result, which was intuitively plausible for Prator's predictions, was not observed in the control group's performance on the same pre-test. That is, the Level 3 items yielded even a slightly lower accuracy mean (1.64)

in the control group than the accuracy index (1.70) observed in the Level 5 items. As a result the entire 93 student sample performed similarly in the reading-aloud items classified Level 3 and 5, 1.69 and 1.68 respectively.

Table 36

*Reading-Aloud Items Categorized Based on Prator's Levels of Difficulty*

Difficulty Level	Principal Sound	Real Word	Artificial Word
3, Reinterpretation			
	[ɪr]	well-being	keeing
	[eɪ]	label	nabel
		propane	tropane
	[oʊ] <sup>33</sup>	cocoa	ploat
		disco	nesco
	[ɑ]	jogging	mogging
			kot
	[r]	pretty	dreaty
		water	nater
	[kw]	quiz	queet
5, Split			
	[ɔ]	coleslaw	
		thought	phought
	[f]	fashion	fishion
		file	fime
	[t]	e-mail	lail
		smile	snile
	[z]	organization	trapization
		zoo	zoul

<sup>33</sup> Unlike the same sound (Level 0) in the identification test, [oʊ] here in the reading-aloud test was classified into Level 3, Reinterpretation, as the spelling “oa” needed to be reinterpreted to be decoded as [oʊ].

Table 37

*Mean Scores of the Reading-Aloud Pre-Test Categorized Into Difficulty Levels*

Group	Instructor		Level 5	Level 3
Experimental	K1	<i>N</i>	19	19
		Mean	1.68	1.75
		STD	0.35	0.30
	K2	<i>N</i>	23	23
		Mean	1.63	1.76
		STD	0.49	0.38
	Exp. Subtotal	<i>N</i>	42	42
		Mean	1.65	1.76
		STD	0.43	0.34
Control	E1	<i>N</i>	27	27
		Mean	1.71	1.61
		STD	0.38	0.31
	E2	<i>N</i>	24	24
		Mean	1.68	1.67
		STD	0.40	0.30
	Con. Subtotal	<i>N</i>	51	51
		Mean	1.70	1.64
		STD	0.39	0.31
Total	<i>N</i>	93	93	
	Mean	1.68	1.69	
	STD	0.41	0.33	

Table 38

*Mean Scores of the Reading-Aloud Post-Test Categorized Into Difficulty Levels*

Group	Instructor		Level 5	Level 3
Experimental	K1	<i>N</i>	19	19
		Mean	2.02	2.19
		STD	0.39	0.52
	K2	<i>N</i>	23	23
		Mean	1.77	1.94
		STD	0.41	0.44
	Exp. Subtotal	<i>N</i>	42	42
		Mean	1.88	2.05
		STD	0.42	0.49
Control	E1	<i>N</i>	27	27
		Mean	1.69	1.69
		STD	0.35	0.42
	E2	<i>N</i>	24	24
		Mean	1.61	1.73
		STD	0.36	0.31
	Con. Subtotal	<i>N</i>	51	51
		Mean	1.65	1.71
		STD	0.35	0.37
Total	<i>N</i>	93	93	
	Mean	1.76	1.86	
	STD	0.40	0.46	

Nevertheless, on the post-test, both the experimental and the control groups consistently showed a higher accuracy mean in the production of the Level 3 items

than the accuracy mean observed in the Level 5 items. This led to different mean scores, when averaged across the entire sample, of 1.76 and 1.86 for Level 5 and Level 3 respectively (see Table 38). However, the control group, for some reason, still showed a relatively small difference in terms of performance in the two different levels of difficulty.

In terms of the read-aloud accuracy gain score, both the experimental and control groups indicated greater accuracy gains in the Level 3 items than the accuracy gains they achieved in producing the Level 5 items, which led the entire group to obtaining average gains of 0.08 for Level 5 and 0.17 for Level 3 (see Table 39). It should be noted that the control group showed an even larger mean difference, 0.11, between the accuracy indices in the two levels, -0.04 and 0.07, while the control group's gains yielded a smaller difference, 0.06, between the mean gains 0.23 and 0.29, for Level 5 and Level 3, respectively. The control group's performance gain might have been more greatly influenced by the effect of the high difficulty that the Level 5 items imposed than the experimental group's performance gain was affected by Level 5. In other words, in the particular case of the current study, the experimental group's approach might have been more effective in improving the pronunciation of the Level 5 sounds than the control group's approach was in improving the same Level 5 items.



Table 39

*Mean Scores of the Reading-Aloud Gain Categorized Into Difficulty Levels*

Group	Instructor		Level 5	Level 3
Experimental	K1	<i>N</i>	19	19
		Mean	0.34	0.43
		STD	0.42	0.58
	K2	<i>N</i>	23	23
		Mean	0.13	0.18
		STD	0.32	0.34
	Exp. Subtotal	<i>N</i>	42	42
		Mean	0.23	0.29
		STD	0.38	0.48
Control	E1	<i>N</i>	27	27
		Mean	-0.03	0.08
		STD	0.34	0.29
	E2	<i>N</i>	24	24
		Mean	-0.06	0.07
		STD	0.29	0.33
	Con. Subtotal	<i>N</i>	51	51
		Mean	-0.04	0.07
		STD	0.32	0.31
Total	<i>N</i>	93	93	
	Mean	0.08	0.17	
	STD	0.37	0.41	

Interestingly enough, the control group showed degradation in their performance of producing the Level 5 items. This aspect is worth discussing, as both groups' pre-test scores were not extremely different (see Table 37). In spite of

this generally equal initial performance, the control group's accuracy showed a degradation when they read-aloud the Level 5 items on the post test. This negative gain is partly attributable to the unintuitive pre-test result of the control group, which indicated even slightly higher accuracy for Level 5. It seems impossible to determine whether it was the dissimilar prior level of accuracy or some dissimilar characteristics of the two teaching approaches that caused these differential gain patterns in Level 5 and 3. This is because this analysis only involved descriptive statistics.

## FINDINGS AND PEDAGOGICAL IMPLICATIONS

### RESEARCH QUESTIONS ANSWERED

#### IMPROVEMENT IN THE L2-SOUND IDENTIFICATION ABILITY

The effects of the two pronunciation teaching approaches on the improvement in L2-sound identification achieved by the intended population, adult Korean learners of English, were investigated in this study. Research question 1 concerned this kind of effect.

1. Are adult EFL learners better able to aurally recognize English sounds after they receive explicit L2 pronunciation instruction based on their interlanguage and L1?

The experimental group's approach to teaching pronunciation based on error analysis typically observed in EFL learners' English pronunciation turned out not to be more effective than the control group's approach is ( $p = 0.786$ ). This is because, in terms of inferential statistics, the means of the total identification scores observed in the experimental and control groups were not significantly different. Therefore, it can be said that the two different approaches to pronunciation teaching have no

significantly different effect on the L2 identification ability tested using the 16 identification questions in this study.<sup>34</sup>

In terms of individual tested L2 sounds, only the high front lax vowel [ɪ] in the word “bid” showed a statistically higher identification mean score in the control group<sup>35</sup> than the mean score observed in the experimental group.<sup>36</sup> ( $p = 0.005$ ). Therefore, it can be said that the control group’s approach, which involved a native English-speaking instructor and explicit teaching agendas from the workbook *Pronunciation for Success*, may be more effective in enhancing the identification ability of [ɪ] as in “bid” than the experimental group’s approach would be.

However, the experimental group’s relatively high identification outcome observed in the word “jealous,” ( $p = 0.033$ ) compared to the control group’s performance on the same word, deserves some attention.<sup>37</sup> This result was considered to be statistically insignificant in spite of the decent  $p$  value because the  $\alpha$  value that had been set by the researcher was adjusted from 0.10 to 0.006 based on the Bonferroni’s recommendation for Post-Hoc mean comparisons. If there had been a theoretical or experimental basis for choosing this specific word and sound for ANCOVA tests, the Bonferroni adjustment would not have been applied.

---

<sup>34</sup> The estimated marginal means for the experimental and control groups were 11.2 and 11.3 with the standard errors of 0.28 and 0.26, respectively. The perfect total score was 16, 1point per each of the 16 test items.

<sup>35</sup> The estimated marginal mean score was 0.77 with a standard error 0.43, out of 1.00.

<sup>36</sup> The estimated marginal mean score was 0.54 with a standard error 0.50, out of 1.00.

<sup>37</sup> The decent  $p$  value, 0.017, observed in the identification of “breathe” was not of concern here because it appears to have been caused by the control group’s unusually poor performance (-0.355 as the gain) for the same word on the post-test, rather than by the experimental group’s improved performance (see Table 13, p. 137).

Without the Bonferroni adjustment in consideration, the above mentioned  $p$  value, 0.033, can in fact be considered meaningful in many areas of social science that involve 0.05 as the  $\alpha$  value to determine statistical significance. Therefore, this sound might yield a more fruitful outcome if it was investigated in a more planned experimental comparison, rather than the haphazard Post Hoc comparisons that the current study relied on.

In addition, as noted earlier, the experiment was conducted in the middle of an English intensive course that might have involved a number of confounding variables that would make the experimental effect difficult to be statistically detected.

#### **IMPROVEMENT IN THE L2-SOUND PRODUCTION ABILITY**

In this study, the L2-sound production ability was measured in the form of reading-aloud tests. The second research question involved this ability.

2. Will they be perceived as having more intelligible accents by native English speakers when they pronounce English words?

It can be said that the experimental group's approach, which involved Korean instructors of English with error-analysis-based teaching agendas, is more effective in making the population have intelligible accents than the control group's approach would be. Based on the ANCOVA result of the sum of the post-test

scores acquired from the 11 tested sounds, the experimental group showed a statistically significant higher mean difference.<sup>38</sup> The experimental population's estimated marginal mean of the post-test sum score was 21.5, while the control population' was 18.6, out of the perfect score of 33, given that the same reading aloud test is taken by the population. Even if the experimental population's mean of the sum of the 11 scores is expected to be significantly higher, the answer to this research question needs to be considered carefully, as the sum score is not regarded as representing all the possible segment level-pronunciation scores. Based on the Eta squared, the difference of the syllabus, as a whole, seemed to account for about 18% of the total variance.

In terms of individual sounds, the averaged post-test scores of the four sounds [z], [kw] and [ou], among the 11 analyzed sounds, may be the three segmental sounds that may be more effectively enhanced by the experimental group's approach. Similar to the identification test results, it should be noted that mean comparison result of the English dark /l/ ([ɫ]) was determined insignificant in spite of its decent *p* value, 0.007, based on Bonferroni's adjustment for 22 Post Hoc comparisons<sup>39</sup>.

#### **ROLE OF THE ARTIFICIAL-WORD TEST IN L2-PRODUCTION TESTING**

Except for a few non-sense words that are likely to be pronounced in more

---

<sup>38</sup> The observed *p* value was close to 0.000 (refer to Table 22).

<sup>39</sup> Only *p* values smaller than 0.005 (0.10 divided by 22 comparisons) can be regarded as statistically significant following Bonferroni's recommendation.

than one way by English speakers, the artificial word reading tests seem to have played the role of double-checking whether the learned items were internalized so that they became ready for future application to novel contexts. If not, the improved accuracy on the real-word testing was no more than another instance of behavioral conditioning encouraged by relatively frequent opportunities to listen and/or repeat the tested words. As outlined in a section of Chapter 4, which described the interpretation of the ANCOVA tests, the discussions on the possibility of knowledge transfer or on the observed questionable gain patterns would not have been made possible without the artificial-word reading tests.

In general, the artificial-word reading test can be said to have helped detect implicational patterns among the ANCOVA statistics (see Chapter 4), in terms of improvements, which most educators desire, not merely in terms of the deterioration that happened in the comparison group. However, it should be advised that the uses of artificial words need to be restricted only to the purpose of comparison in the course of investigation. They were never intended as examples for the L2 learners to practice. In addition, artificial-word reading tests are likely to substantiate results observed in real-word reading aloud tests when the target L2 has a relatively transparent orthography system rather than when the L2 has an opaque system. This is because, an opaque orthography tends to induce the native speakers to pay more attention to word-specific particularities rather than to generalizable spelling rules.

## L2-SOUND PRODUCING SKILLS TRANSFERABLE TO NOVEL WORDS

One may think the afore-mentioned higher accuracy improvements can be achieved if the two teaching approaches present a different set of example words. For example, the experimental group's teaching agendas happen to have presented more words to the students that were included in the reading-aloud test items than the control group's agendas did. To test this possibility of familiarity-originated mean differences, as opposed to mean differences due to internalized segment pronunciation skills, the third research question was addressed.

3. Are they better able to transfer the learned pronunciation knowledge to novel English words?

With respect to the sum of pronunciation accuracy scores in the 11 sounds, the experimental group's teaching approach may also be considered more effective in enhancing the population's production accuracy of the segments in the artificial words than the control group's approach is. Therefore, it is difficult to attribute the control group's relatively low accuracy rate to the possibility of their potential unfamiliarity with the tested words. In other words, the experimental group's approach may be more effective in making the population able to pronounce the segmental sounds even when they occur in unlearned words.

In terms of transfer of the improvements detected in the individual real-word readings to the artificial words, only the sounds [z] and [kw] can be said to have



indicated desirable knowledge transfer. The case of the  $p$  value, 0.049, obtained from the artificial words containing the diphthong [ou], which indicated a statistically significant mean difference in the real-word reading-aloud test, again was not considered to be statistically significant for it was greater than the Bonferroni-adjusted  $\alpha$  value, 0.005 (refer to Appendix K).

The fact that the artificial-word reading accuracy rate of the segment [t̥] showed a statistically significant higher mean ( $p = 0.002$ ) while the real-word reading accuracy rate did not ( $p: 0.007 > \alpha: 0.005$ ) is worth discussing. One of the possible speculations regarding this result is that the effect of the experimental group's approach on the transfer of the learning of [t̥] to novel-word reading may be greater than the effect of the control group's approach on the transfer ability. This is because the effects of both group's teaching approaches on the improvement in the *novel*-word reading accuracy turned out to be statistically dissimilar while the effects of both group's teaching approaches on the improvement in *real*-word reading accuracy did not. Simply put, the control group population would produce [t̥] as accurately as the experimental group population would, when the sound is contained in real words. However, the experimental group population would be more accurate than the control group population would be, when they pronounce [t̥] occurring in unfamiliar words.

Nevertheless, it appears to be reasonable to postpone such an inference about the aforementioned transfer suggested concerning the production of [t̥]. This is

because the observed significance seems to have been detected partly due to the control group's negative gain, as well as the experimental group's positive gain.<sup>40</sup> Therefore, in order to be certain about the case of [t̚] production, we need further explanations for the observed cases with negative gain. One of the possible explanations is that negative gain scores are not surprising, as a couple of similar cases were also observed in the pilot study (Huh, 2003). Interestingly enough, the same segmental sound, [t̚], yielded negative gain scores when it occurred in the tested words “kill” and “kettle” within the pilot study. In order to examine other possibilities, further research is needed with a more strict experimental design.

#### CAH'S PREDICTIONS COMPARED TO THE OBSERVED L2-SOUND DIFFICULTIES

Prator's Hierarchy of Difficulty with actually observed L2 sound learning difficulties. The last research question concerned this comparison.

4. Do the English sounds classified according to Prator's hierarchy of difficulty, with the consideration of Korean as the L1, show difficulties of pronunciation learning that are similar to the levels suggested by Prator?

Based on the review of the CAH literature, the researcher was able to propose levels of L2 learning difficulty by assigning one level from 1 to 5 as to the nature of trouble that Korean learners of English might face concerning each of the

---

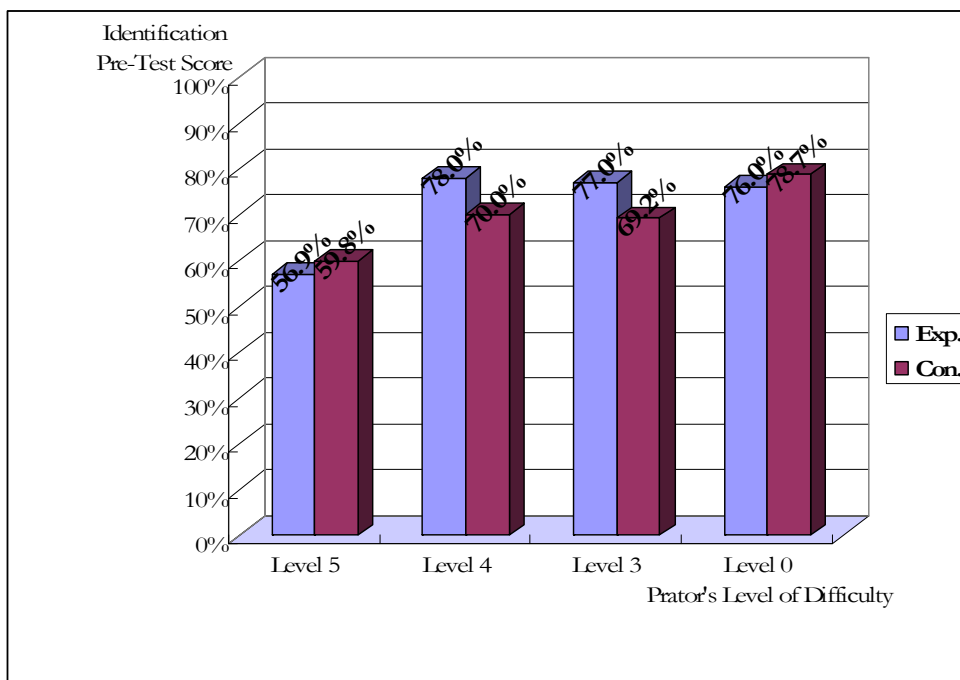
<sup>40</sup> The control group's gain was -0.15, which was approximately one fourth of the standard deviation, while the experimental group's positive gain was 0.39, which was about a half of the standard deviation (Refer to Table 29, p. 166).

tested sounds. Among the sounds that were investigated using the aural differentiation tests, there were four different levels of difficulty, [positive] Transfer (Level 0), Reinterpretation (Level 3), Overdifferentiation (Level 4), and Split (Level 5).

As seen from Figure 2, it turned out that the highest level of Prator’s difficulty was actually the most problematic even with the participants’ years of EFL learning on the pre-test. Other than that, Levels 0, 3, and 4 did not seem to differ from each other to a very large extent in the scores of identification pre-test. The view might be taken that Prator’s Hierarchy of L2 learning difficulty need some more consideration in order to properly explain these observed results.

Figure 2

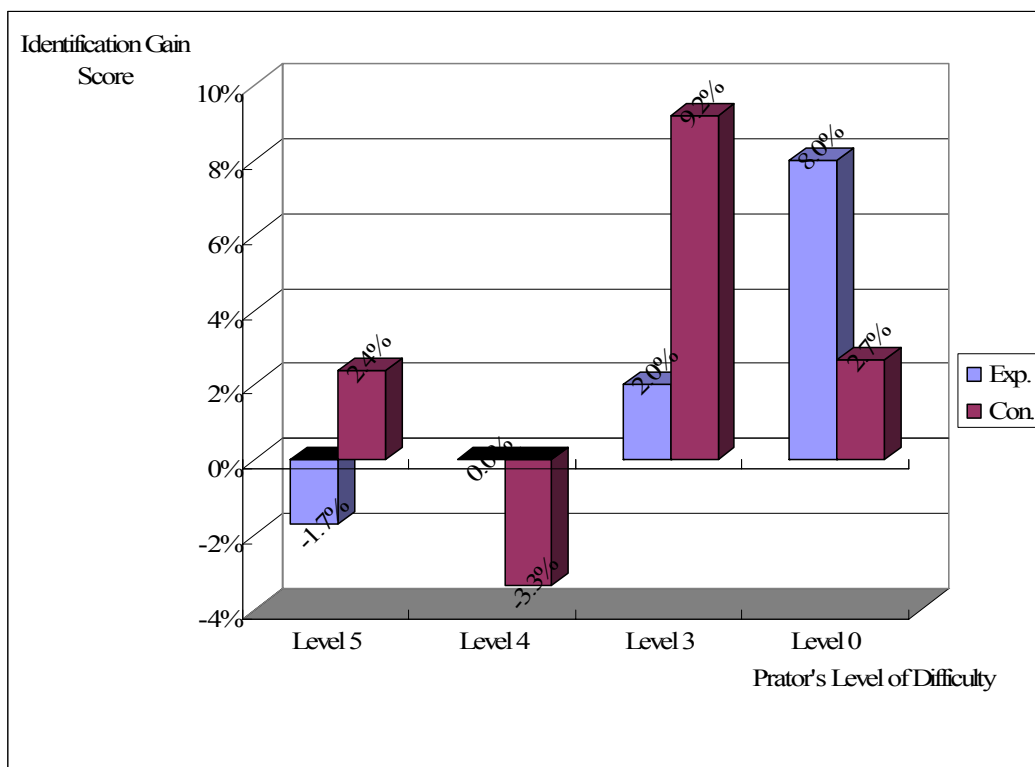
*Identification Pre-Score Comparison Based on the Proposed Levels of Hierarchy*



A similar chart that compares the gain scores acquired from the identification tests may shed more light on the relationship between the observed improvements and the difficulty levels that were assigned by the researcher. Based on Figure 3, in terms of L2 sound identification, it seems that Split (Level 5) and Overdifferentiation (Level 4) instances were the hardest levels to make improvement, while words with sounds in the categories of Positive Transfer (Level 0) and Reinterpretation (Level 3) were relatively easy to enhance.

Figure 3

*Identification Gain-Score Comparison Based on the Proposed Levels of Hierarchy*



Unfortunately, the reading-aloud tests featured only two levels of difficulty, Split (Level 5) and Reinterpretation (Level 3). This limited composition of difficulty levels appears to be accounted for when we review the basis on which the researcher chose the target sounds for the current study. That is, these two levels are the ones that involved L1 items that were also present in the L2 sound system, but with “a new shape or distribution” (Level 3) and with “two or more” variations in the L2 (Level 5), based on Prator in Brown (1994a, p. 196).

The analysis of the reading-aloud pre-test scores suggests that L2 production accuracy rates at Levels 5 and 3 did not differ very much.<sup>41</sup> Therefore, it can be said that the particular sample of Korean adult learners of English showed no considerable difference between their production accuracy rates of Levels 5 and 3, probably as a result of their EFL learning before the current study.

However, in terms of *improvement*, the two levels of difficulty yielded different results. As discussed earlier in the previous chapter, the entire sample ( $N = 93$ ) made slightly dissimilar achievements in reading-aloud sounds at Level 5 and Level 3, 0.08 and 0.17, respectively.

Specifically, the experimental group showed a relatively small deferential score (0.06) between the observed accuracy rates in the sounds classified into the two different levels, compared to the deferential score observed in the control group (0.11). One may suggest that the experimental group’s approach might have been

---

<sup>41</sup> On average, the entire sample ( $N = 93$ ) scored 1.68 for Level 5 and 1.69 for Level 3, out of the perfect score 3.00 (refer to Table 37, p. 174).

more effective in closing the gap between the improvement rates of Level 5 and Level 3.

However, it is hasty to judge that one of the two teaching approaches was more effective than the other as the observed score difference was considered to be so small that it was not certain the observed difference was actually due to the treatment effects. As suggested similarly in the discussion of the CAH predictions in relation to the observed identification difficulty, the researcher considers some more factors affecting L2 learning difficulty other than the mere discussions on the existence or absence of certain sounds in the L2 and L1. They include, at least, different orthography conventions, linguistic contexts in which L2 sounds occur, and socio-linguistic barriers. In what follows, the descriptive data concerning the first two factors will be discussed as candidate factors influencing L2 sound learning difficulty.

#### **MODIFICATIONS CONSIDERED FOR CAH-BASED EXPLANATIONS**

##### **ORTHOGRAPHY AS A FACTOR OF L2-SOUND LEARNING DIFFICULTY**

Orthography, despite its complexity especially in English, is one of the two potent factors that are likely to govern much of interlanguage formation processes. As can be seen on Table 26, the initial performance of the participants on the word ‘disco’ is about 50% greater than the one on the word ‘cocoa.’ Furthermore, this relatively inferior accuracy observed on ‘cocoa’ seemed to be more largely shared by the participants according to its standard deviation (0.37), which is less than a half

of the standard deviation (0.81) on ‘disco.’ This more homogeneously inferior pattern of the word ‘cocoa’ meant that it was prevalently and incorrectly pronounced, for example as [k<sup>h</sup>ok<sup>h</sup>oa], while many more of the same participants were able to pronounce the same sound in ‘disco’ with a higher accuracy.

Clarification remarks on unsystematic orthographic conventions such as ‘thought’ and ‘fought’ may help learners overcome their false rendition, making them stop pronouncing the two letters ‘ou’ one by one as [ou] in these words. On the other hand, some of them might be overwhelmed when they were presented additional examples like ‘flaw’ or ‘broad’ as the words that show different orthographic symbols for the same sound [ɔ]. It appears in Table 26 that the experimental group achieved a greater post-test mean score (1.81), compared to the one of the control group (1.35). This difference might have been made possible through the explicit instructions on the various spelling forms (see A. Eye-Openers in Appendix B) that represent the same diphthong [ou], during the experimental syllabus.

Therefore, it might be insufficient to merely consider the absence or existence of L2 sounds in the L1 and L2, in order to shed fuller light on the issues concerning the specific cases in which Prator’s hierarchy did not show reasonable relationship in the test results. In this sense, orthographic conventions may work beneficially when we consider the following two facts about the English alphabet’s place in EFL learning. First, many learners in Western EFL countries may feel

familiar with the English alphabet even though their L1 is not English, partly thanks to its Roman origination. In this global information-sharing era, the English alphabet must have become well known even to many non-English speakers from Asian countries. The negative aspect of this superficially desirable tendency is that the relationships between the English letters and their suggested sounds may be understood by many EFL learners as a one-to-one correspondence, similar to Spanish's patterns of oral reading. Second, the fact that EFL learners tend to learn the English alphabet first, even before they start learning English in class (refer to the EFL learning background survey results in Chapter 4), not to mention before they have an opportunity to talk to English speakers, is likely to exacerbate the situation.

Table 40

*Case Summaries of the Words That Showed Orthography-Sensitive Performance*

Group		Pre		Post	
		<u>coc<u>o</u>a</u>	<u>disc<u>o</u></u>	<u>coc<u>o</u>a</u>	<u>disc<u>o</u></u>
Experimental	<i>N</i>	42	42	42	42
	Mean	1.00	1.60	1.81	2.10
	Std. Deviation	0.38	0.80	1.09	1.01
Control	<i>N</i>	52	52	52	52
	Mean	0.98	1.44	0.98	1.62
	Std. Deviation	0.37	0.83	0.37	0.89
Total	<i>N</i>	94	94	94	94
	Mean	0.99	1.51	1.35	1.83
	Std. Deviation	0.37	0.81	0.88	0.97



## PHONOLOGICAL CONTEXT AS A FACTOR OF L2-SOUND LEARNING DIFFICULTY

A second additional aspect to be considered when analyzing observed errors may be the context, in which the L2 sound occurs. In the current study, a few such contexts were uncovered. One salient pair of examples is ‘Karate’ and ‘meter’ as shown on Table 27. On the pre-test, the final-tensed sound [i] in ‘Karate’ generally indicated a mean of only one eighth (0.31) of the one on the word ‘meter’ (2.24), both with almost the same STD around 0.60. Surprisingly, the word ‘가라테’ or ‘가라테,’ which must have been pretty familiar to the Korean participants, was pronounced prevalently as [k<sup>h</sup>əreɪt<sup>h</sup>] with no vowel at all at the end of the word, not to mention without the Korean-accented pattern [e].

The context of the occurrence, the word final one, seemed to be the major reason for this poor performance. One possibility is that this situation was probably caused by a false analogy between this word’s context and similar contexts observed in other words such as ‘Jane,’ in which the letter ‘e’ is silent. The accuracy index may have changed from 0.29 to 0.79 (see Table 27) as a result of the instructions for the experimental group’s participants on such context-related erroneous interlanguage. The increased standard deviation from 0.60 to 1.16 appears to suggest that the observed improvement was achieved by only the part of the participants, in contrast to the slight improvements that commonly occurred in most of the participants. There may be more instances of such context-sensitive

levels of difficulty including ones like ‘organization’ and ‘zoo,’ the former of which scored lower than the latter.

Table 41

*Case Summaries of the Words That Showed Context-Sensitive Performance*

Group		Pre		Post	
		Karate	meter	Karate	meter
Experimental	<i>N</i>	42	42	42	42
	Mean	0.29	2.24	0.79	2.24
	Std. Deviation	0.60	0.66	1.16	0.73
Control	<i>N</i>	52	52	52	52
	Mean	0.33	2.25	0.46	2.21
	Std. Deviation	0.58	0.71	0.70	0.78
Total	<i>N</i>	94	94	94	94
	Mean	0.31	2.24	0.61	2.22
	Std. Deviation	0.59	0.68	0.94	0.75

## FINDINGS CONCERNING SOCIO-CONSTRUCTIVIST ASPECTS

### NOTICING APPROACHES REFLECTED ON THE SURVEY RESULTS

The syllabus features of processing instruction (PI) that were proposed by the Noticing Hypothesis included instructions on input processing (IP) and mixed referential activities. In general, the IP features were realized as the explicit explanations that the four instructors provided to the participants. As the syllabi’s

focuses were overcoming Korean accents, the instructors were encouraged to specifically talk about how to process and organize what the students would learn from the teaching agenda (learner uptake), into internalized and transferable knowledge (intake). This approach was also implemented by the two groups' workbooks, which were full of tips and illustrations on how to improve the learners' pronunciation. To the control group, even the pronunciation tutorial video clips may have presented approaches similar to the PI, for that same reason. The only unavoidable difference in the designs of syllabi was that the experimental syllabus offered more of the mixed referential activities, which were mainly fulfilled by a number of Konglish pronunciation patterns and the model utterances in the authentic video clips. There was another type of mixed referential activities that consisted of the contrasts made with the model patterns manifested in authentic videos and the sound files from the computerized dictionary, which were not as highly nativelike as the ones in the video clips

Unfortunately the survey on the participants' pronunciation learning during the experiment failed to disclose statistically significant mean differences in the items relevant to the effects of such PI or mixed reference activities. The score of Quality of Classes (see Appendix L), into which this group of relevant survey items were averaged, did not indicate that the quality of the classes in terms of the Noticing Hypothesis's recommendations was significantly higher in one group.

Among the 11 survey items three items showed, nevertheless, statistical significance. Item 19 that asked whether the instructors were looking professional

in the relevant area indicated a higher mean score in the NNEST ( $p = 0.006$ ). This may have been due to the NNESTs' meta-learning instructions, in Korean, on English pronunciation.

On the other hand, Item 18, which investigated the participants' impressions on the degree of how helpful the instructors' explanations were, pointed to a higher mean score in the control group instructors' conducts ( $p = 0.027$ ). At least, the NESTs may not have given the impression that they were incompetent teachers to the participants.

It was also embarrassing to find out that the control group's participants reported that they thought the transcriptions on the experimental group's materials to be helpful in Item 22. This was surprising because the transcriptions for the control group were not written in the IPA symbols, but with a convention that the researcher found unfamiliar. This issue was easily solved by the semi-structured interview with some of the participants. It seemed that, probably due to changes in the National Curricula, secondary school students did not need the IPA symbols when they studied English as much as they had before. They answered that they did not learn to read the IPA symbols at school any more, even though these symbols were the only device that written dictionaries use to describe how to pronounce their entries. Some of them, nevertheless, responded that they were able to read along the IPA transcriptions to a degree even though they did not usually rely on these transcriptions as offered in bilingual dictionaries.

As for the mixed referential examples provided to the experimental group's students, they answered, as an average, 3.23 out of 5.0 (64.5%) regarding the effectiveness of the comparison between the highly nativelike examples and overly articulated items from the dictionary, as can be seen from Item 23. The other mixed referential agendas, contrasts and comparisons between the Korean and English sound systems marked 3.54 (70.9%). It may be said that these mixed references were not perceived as considerably ineffective by the student participants.

#### **IMPORTANCE OF INSTRUCTOR'S ACTUAL CLASSROOM CONDUCT**

It was interesting to note that the participants in the two classes each taught by K1 and E2 generally performed better on the current study's language tests, despite the fact that these students were the ones who responded that they thought the class duration of 75 minutes was too short, as shown in Item 7 in Appendix L. One of the two other instructors, K2 was the one who received the lowest score on Item 27, which asked them about the instructors' sensitivity to the learners' ZPD and about the his/her ensuing adjustments to the flow of lecture to better meet the students' needs (see Appendix E for actual survey items).

Although unproven by inferential statistics, the factor of students' perceived appropriateness of the class duration seemed to present further issues to consider in order to improve the learning environment. The survey items 8, 9, 10, and 13 appeared to have a certain relationship with the item about perceived class duration. As revealed on the result table in Appendix L, the scores of these items given by the

students in K2 and E1 's classes, the duration of which was deemed relatively suitable by the students, were generally lower than the other two instructors' scores. It might be carefully argued that the effectiveness perceived by students tend to be influenced by the factor of perceived suitability of class duration. This is because it seems reasonable that the intended materials may not be perceived as effective if there is not enough time during the class to cover them, for there would also be less time to incorporate the audio-visual materials into the instructor's PI.

Whether the perceived sufficient class duration was detrimental or not, the degree to which the participants perceive the prepared materials as helpful to them might also depend on some factors of the instructors as well as the materials' face values, according to the above-made argument. Probably the notion of instructor's caring suggested by (Goldstein, 1999) may have something in common with the skill of making the best of the class duration by adjusting the flow or modality of instruction dynamically based on the sensed ZPD that the students seem to fall in to.

## **FINDINGS CONCERNING EDUCATIONAL IMPLICATIONS**

### **UTILITY OF EFL LEARNERS' L1 AND INTERLANGUAGE**

As discussed so far, language practitioners' awareness of EFL learners' L1, interlanguage, and socio-cultural background is likely to provide some insightful a posteriori explanations on the problems observed in the learners. In order to make the best of the analysis of the observed errors, this understanding needs to be

applied properly based on the priorities of teaching set by each language practitioner. Moreover, by taking, into consideration, even a superficially arbitrary-looking interlanguage instance, NNESTs, or maybe also NESTs, could acquire insights that may effectively help the learners. For example, when an NEST discovered that his/her Korean student arbitrarily pronounced the words ‘go’, ‘know,’ ‘sew,’ and ‘aloe’ as [go], [nou], [seu], and [aloe], it might be a nice chance for the teacher to learn that the EFL student was suffering from the stabilized hasty generalization that the student mistakenly made about English phonics.

Similar to what was warned against in the review of literature, too much attention to these linguistic and analytical aspects of the L1 and L2 may lead language practitioners to the same predicament that the proponents of the strong CAH version had to face. Therefore, analyses of the L1 and interlanguage patterns may need to be used for a posteriori problem solving, rather than for a priori prediction.

#### NEED FOR CAREFUL EVALUATION OF OBSERVED IMPERFECT INTERLANGUAGE

It was thought-provoking to give the lowest accuracy index 0 to the participants who changed their pronunciation of the word ‘e-mail’ and tried to say it close ([ĩmerɔ]) to what they believed heard or learned, but not close enough to speak it as [ĩmerɪ]. The problem of [ĩmerɔ] was that it was articulated without the proper contact of the tip of the tongue on the gum ridge. As for the general impression of

the authentic pronunciation of the sound [t̚], this less-than-perfect rendition may sound closer to the natively one to some English learners<sup>42</sup>, rather than the more Korean-accented one [ĩmeɪl] with the clear ‘l’ at the end of it. Due to its critically larger possibility of being unintelligible, ‘이메이어’ received 0, while the more Konglish-sounding samples ironically received 1. This was how some of the deterioration cases of [t̚] were caused while grading.

However, when we consider the transient nature of this unfairly treated case of interlanguage, we may need to regard it as more promising because it seems to be much harder to change [ĩmeɪl] to [ĩmeɪt̚] than to change [ĩmeɪə] to [ĩmeɪt̚]. The reason that the latter case may be more feasible is that this interlanguage pattern shows that the learner must have already overcome their insensitivity to the natively version, unlike the former case, in which the over-generalized phonological translation of [t̚] into ‘ㄹ’ may still be blocking the students’ noticing. Therefore, it may be more reasonable to provide compliments to the students who showed the rendition without [t̚] or [l], for their meaningful change and cheer them up to go on to let their tongue touch the gum ridge.

---

<sup>42</sup> Both of the researcher’s daughters, who learned English for the first time in the States as a 1<sup>st</sup> and 2<sup>nd</sup> grader, showed similar tendency when they relied on code-switching to say such words with [t̚], as if there was no consonant at the end of the word ‘school,’ for example “우리 스쿠어가 ([skuə]) 좋아요,” not “우리 스킨이 ([skul]) 좋아요.” The latter seems to have been definitely a more widely observable instance of Korean accent.



### POSSIBILITY OF NNESTs' HYPERCORRECTION

The researcher happened to notice some of the possible hyper-correction instances that are often incorrectly justified by NNESTs like him in the course of zeroing-in his grading criteria, similar to the sound [m] in 'summer' and 'mammoth,' which was discarded during the reading-aloud test analysis. These undesirable *teacher* language (as contrasted to the term '*learner* language') items need to be carefully investigated as to whether or not they are really correct with the help of NESTs, in order to prevent the spread of another kind of foreign accent pattern. They may include vowel clusters of the two same vowels [ɪɪ] as in 'flying' or 'studying,' and diphthongs such as [ou] as in 'disco.' The researcher does not insist on the plausibility of the Korean transcriptions such as '플라잉,' '스터딩,' or '디스코.' Rather, he is warning against some potential undesirable situations based on his embarrassment when he learned about these cases while working with the American informants. Such situations may include the ones in which those sound clusters and diphthong were overly pronounced by EFL learners with so much elongation and articulation that the words' primary stress patterns might be misunderstood by English speakers when they heard them.

The observed sounds [ɪɪ], [ou] and [m], which seemed typically prone to hyper-correction, appeared to fall in the category of Level 3, Reinterpretation, or Level 2, Underdifferentiation. This is because the Korean sound system has virtually the

same items, as exemplified by the words ‘이이,’ ‘죽마고우,’ and ‘맘마,’ but they either have different shapes or distribution, or they are absent in the L2.

#### TEACHER EDUCATION FOR NNESTs

Not unrelated to the arguments previously made about the implication of instructor’s efficacy for students’ perceptions about the course materials, the need for appropriate teacher education deserves due attention. Not to mention the undesirable possibility of hyper-correction, there seem to be a number of agendas for English teacher education. Even NESTs might be assisted by the L1- and interlanguage- teaching items, if the preparation of them can be made cost-effective.

Similar to what was outlined in (Medgyes, 2001), the current study’s participants did not trust the pronunciation patterns of the NNESTs despite the fact that they considered the experimental group’s Korean instructors as more professional in appearance. Item 20, which concerns this issue, was one of the few survey items that indicated significant ANOVA results. The experimental participants perceived reliability of their NNESTs’ pronunciation patterns as significantly ( $p = 0.002$ ) lower than the one perceived by the other group of students concerning the native English-speaking teachers. The teaching items and materials designed for the current study could be one of the possible agendas to assist EFL NNESTs in Korea in improving their own accuracy. Without being familiar with the input processing instructions such as the ones included in the current study’s experimental syllabus, NNESTs might look even unprofessional. This negative

perception may make it harder for the Korean instructors of English to make the best out of the strengths that NNESTs were claimed to afford in the review of literature.

#### **NEED FOR TIMELY UPDATE OF TEACHING AGENDAS**

Unlike what was observed in the pilot test's *older* participants, the *freshmen* participants did not show as many observations of the unnatural pronunciation of the English flap sound, as the older learners tended to exhibit. For example, two of the pilot study's participants revealed that they felt they should not speak the words, for example, 'water' or 'party' properly using the flap ([r]), even though they knew how to say this linguistically easy sound (Transfer of word initial '≡') properly in these words. The researcher often observed EFL learners, even teachers, who believed that they should say [t<sup>h</sup>] or [d] instead, just because they were not very proficient in English or due to the apprehension that they might look as if they were boasting. The researcher, unlike this older tendency, noticed that most of the freshmen uttered the word 'water' natively and their initial average for it was decently high (2.44 out of 3.00).

No matter what caused this kind of change, it might be worthwhile to try to collect updated information on the target students' initial capabilities. If a teacher plans to spend a considerable amount of time for such cases as the flap sound based on his/her own English learning experience, he/she may need to double-check if they still need to do it when they teach today's EFL students.

## UTILITY OF EXPLICIT PRONUNCIATION SYLLABI

The researcher believes that most of his English competence and enhanced pronunciation accuracy were made possible by his experience of participating in conversations with English speakers. The pronunciation curricula designed for the current study, therefore, were not intended as a mainstream course that solely deals with explicit pronunciation matters. The syllabi, instead, are thought to be more suitable for a relatively short preparatory courses or a semester-long lab class supplementary to a regular English course, for the purposes of scaffolding. If the learners could heighten their confidence and efficacy in segment-level English listening and speaking, by perfecting their interlanguage items that used to be somewhat prone to unintelligibility, the scaffolding goals of these treatment agendas may be assumed as decently achieved. However, in the less normal situation of teaching student teachers or current teachers, if needed, the syllabi might take the form of a full semester course.

## Chapter 6

### CONCLUSIONS

#### SUMMARY OF FINDINGS

This study attempted to compare the effects of two different kinds of approaches to teaching English pronunciation to Korean adult EFL learners. The two syllabi were organized so as to last only four 75minute long classes as L2 pronunciation lectures are typically conducted as a supplementary tutorial to a main conversation course rather than as a separate semester-long L2 pronunciation course. The study was conducted in the middle of an English intensive course that lasted five weeks.

Based on the findings concerning identification abilities, the two explicit approaches to teaching pronunciation may have no significantly different effect on enhancing the population's L2 sound identification ability. The ability to identify English high front lax vowel [ɪ] may be more effectively enhanced through the control group's approach, which involves a native English-speaking teacher and teaching agendas for general L2 learners aided by video tutorials that a published pronunciation-learning workbook provides.

However, the findings as to L2 producing abilities suggest that the experimental group's approach, which involved a Korean teacher of English and teaching agendas based on error-analysis commonly observed in Korean EFL

learners aided by authentic video clips that provide model utterances, is more likely to improve the population's reading-aloud accuracy. The experimental group's approach is more likely to enable the population to transfer their pronunciation skills that they are expected to attain in the course of the four-day explicit learning, to the production of the same sounds in unfamiliar words. The production accuracy rates for the individual sounds [z] and [kw] may be enhanced more effectively by the experimental group's approach.

This study has descriptively compared various L2 sounds that were assigned different levels of difficulty by the researcher based on Prator's Hierarchy of L2 Learning Difficulty. Based on the non-inferential mean comparisons, the Level 5 sounds [ɪ], [i], [ɔ], [z], and [ʒ] were considered to be more difficult for the Korean freshmen, prior to the explicit pronunciation learning, to identify than were the L2 sounds that were assigned Levels 4, 3, and 0. These three levels were not regarded as having imposed considerably different degrees of identification difficulty on the same sample of students. In terms of improvement, Level 5 indicated a distinctly lower gain score than the identification gain scores that Levels 3 and 0 showed.

Concerning the sample students' production of the L2 sounds, the Level 5 and Level 3 sounds showed almost no difference in terms of prior reading-aloud accuracy. When it comes to the reading-aloud gain score, the Level 5 sounds [ɔ], [f], [ʃ], and [z] turned out to be slightly more difficult to improve than the Level 3 sounds (refer to Table 36 for detailed ramification of the levels). Based on two

cases of descriptive mean comparisons, this study suggests that other factors, such as complexity of L2 orthography and linguistic contexts in which L2 sounds occur, need to be added to the CAH's considerations, which only concerned the presence or absence of certain L1 and L2 sounds. Also, socio-cultural values concerning certain patterns that L2 learners avoid when speaking, which are likely to be shared by the L2 learners, were considered to be potential non-linguistic obstacles to desirable L2 pronunciation by the researcher. However, it must be noted that these descriptive data should be confirmed by stricter experimental designs in order to be generalized.

#### **LIMITATIONS OF THE STUDY**

The researcher was able to recognize the following two areas of limitations among the many possible aspects that are subject to improvement. They are the areas of limitations of the measurement instruments and the limitations in terms of external validity.

#### **LIMITATIONS OF THE MEASUREMENT INSTRUMENTS**

First of all, the validity of the reading-aloud test, one of the two language testing instruments employed in this study, bears weaknesses. Even with the justification of the reasons that made the researcher discard some of the sound samples that he felt unconfident of while grading, there may be a possibility that the grading procedure was biased by extraneous factors, at least including the NNEST's

pronunciation accent whatsoever and his potential enthusiasm toward the experimental syllabus. Moreover, as the pronunciation samples were evaluated by one grader, reliability issues may not have been properly secured, compared to a better case, in which more than one grader produces averaged scores.

Next, the reading-aloud test items were administered in a linguistic testing context rather than a realistic communication context. That is, the results of the tests conducted in the situation, in which the participants read each of the single words, needs to be generalized into similar individual-word reading contexts. As suggested by the dissertation committee, reading more than one word as a phrase, which may enhance the perceived authenticity of the test items, can be one of the solutions to this problem given that the test's experimental validity is still secured.

The distribution of the accuracy scores in the reading-aloud test grading criteria is not free from limitations. This is because the criteria may not be very objective and consistent across the entire tested sounds. The researcher was not able to avoid this limitation, as he also had to consider the external validity of the test results and the seriousness of the foreign-accented patterns toward the intelligibility of the spoken words. As pointed out in Chapter 3, therefore, the same English sounds may need to be provided with different criteria if they are intended for L2 learners other than Korean learners of English.

Another weakness of the measurement instruments is that the students may have been influenced by the pronunciation patterns of their colleagues adjacent to



them while they were recording the test items. Although they were wearing an ear-covering headphone, other participants' voices may have been heard through the microphone because the microphones were very sensitive. This is probable because the grader was able to notice the readings of participants other than the graded participant while listening to the recorded samples.

#### **LIMITATIONS IN TERMS OF EXTERNAL VALIDITY**

There seem to be some limitations regarding the generalizability of the current study's findings, including the following two. First, the sound pool included in the proficiency tests of this study is not exhaustive, even though it covered a wider range of sounds than the sounds tested by some previous research such as [1] and [1]. The currently tested sounds were chosen based on the degree of potential problems, for Korean adult learners, expected by the researcher. There might be some more sounds that may reveal different insights on English segment-level pronunciation.

Second, the fact that the participants were taking an English intensive course right before the pronunciation treatment seems to raise a considerable external validity issue. This is so because typical populations other than the participants may not have any opportunity to take such intensive course. This means that there might be some more sounds that are potentially problematic, but not detected, in terms of improvement even with the designed treatments. Similarly, the participants' dramatic increase in time spent in using English during the treatment,

as outlined in Chapter 4, may make them dissimilar to the intended population, ordinary Korean college learners of English.

#### **SUGGESTIONS FOR FUTURE RESEARCH**

Based on the findings and the limitations debriefed above, the researcher suggests the following future research. First, a study involving a more contextualized reading-aloud test would yield more fruitful findings. For example, the reading-aloud test can be organized with everyday reading texts or conversation scripts in easy language that would minimize potential biases due to the unequal levels of L2 reading abilities or vocabulary.

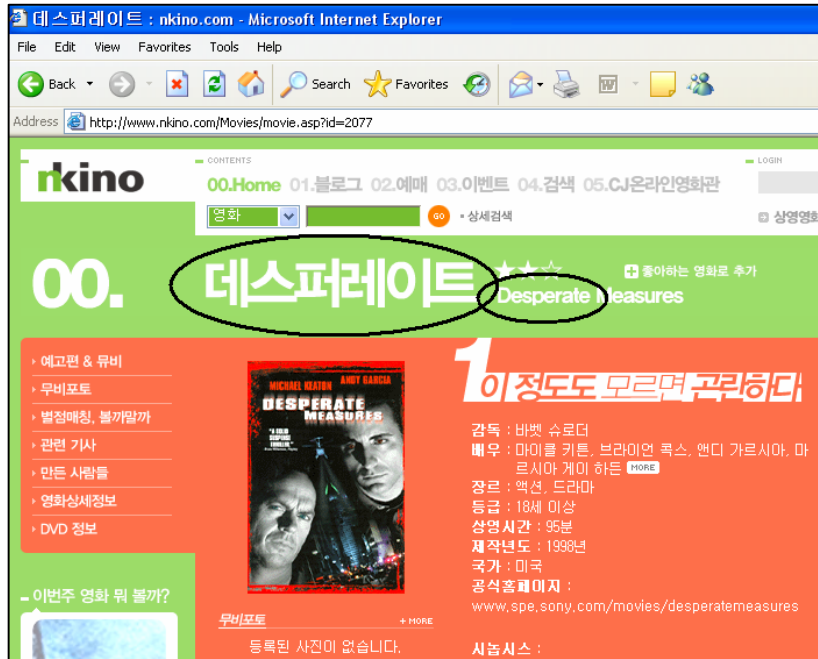
Next, based on the potential characteristics of particular sounds revealed in this study, one can devise a study that focuses on a smaller number of sounds in a more experimental setting, in order to more robustly test hypotheses suggested in this study. For example, the categories of problematic English fricatives or English tense vowels, as long as they are studied within pedagogical boundaries, appear to be two of the possible meaningful investigation agendas.

Thirdly, it may seem worthwhile to investigate the longitudinal effects of the experimental and control syllabi on the participants' learning outcomes. As the interviewee suggested that they needed more time to internalize what they recognized in the pronunciation lessons, extensive caring for the students' changing interlanguage patterns may lead to higher quality English competence. Reading texts in the area of the students' interest, for example, Military English materials or

passages for foreign aviators written in English, may make the learning experience more meaningful and useful.

Last, a study on the possible hypercorrection cases that was suggested concerning the discrepancy between NNEST's and NEST's perception of the same foreign-accented sounds, as in “summer” and “flying” (refer to Chapter 4) appears to be worth conducting. This study may be fruitful the interlanguage rule that was considered to cause the hypercorrection tendency seems to be shared by many Korean teachers of English. In addition, there were a few cases, in which native English-speaking informants were not sensitive to the distinction that the hypercorrection may emphasize, in the course of the current study.

1. The Movie Title “Desperate” and Its Problematic Korean Transcription



2. Poster of the Imported Movie “Saw” With Its Misleadingly Transcribed Title



<sup>43</sup> Pay attention to circled parts for the misleading transcriptions.

Appendix B, Sample Workbook Page for the Experimental Group

Vowels of North American English

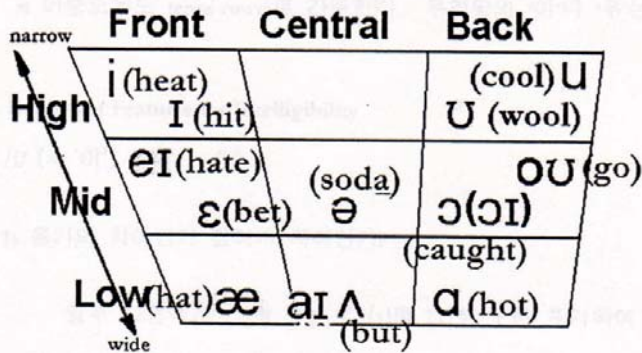
A. Eye-Openers

1. "know," "no," "sew," "goat," "alone," "code," and "cocoa."

How many kinds of vowels?

2. 모음 조성 위치 (Place of Articulation):

1) 미국영어 모음 위치도와 그 예들 → Vowel Chart ㉠-11



2) 미국영어 모음과 한국어 모음 비교 → Vowel Comparison Chart ㉠-12

3) 미국인들이 단모음으로 소리내기 곤란한 모음들: 외래어 중 그 소리가 남아 있는 경우를 제외하고는 대부분 아래와 같이 바뀐다.

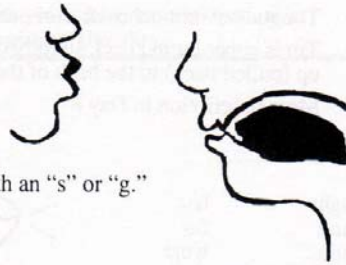
/a/ → /ə/ ('aluminum,' 'asthma') ㉠, but, /aɪ/는 가능: 'part' 'cart'

/e/ → /i/ ('Panske,' 'Karate'): final tensing 현상

/o/ → /ou/ ('Brazo'): final tensing 현상

### Sh and Zh

- Teeth touching or almost touching
- Air vibrating through the teeth
- Lips are outrounded in a TRUMPET shape



sh (no voice) / zh (voice)

"zh" is always in the middle or end of a word and spelled with an "s" or "g."

shashashasha.....zhazhazhazhazha

mesh

measure

cash

casual

bash

beige (don't put an e sound at the end)

gash

garage (don't put an e sound at the end)

fashion

occasion

Sh / Zh sounds

She surely made a fashion statement with her questionable shoes.

We usually measure the pleasure of our leisure.

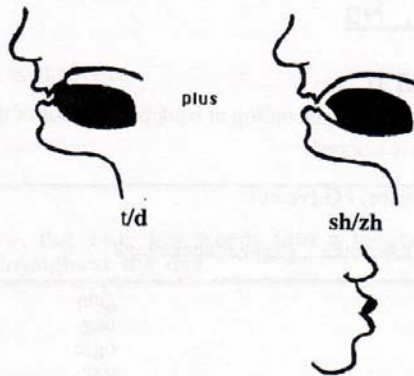
We casually measured the beige garage for the Asian company.



Add your own "sh" or "zh - sound" sentences below. Put your key words into a tongue twister and practice that sentence often throughout the day.

### Ch and J

- Tip of the tongue on the alveolar ridge
- Air is stopped and then released.
- Lips are outrounded in a TRUMPET shape



Ch (no voice) / J (voice)

Ch is a combination of T + Sh

J is a combination of D + Zh

chew

Jew

which

wedge

choose

juice

matches

magic

nature

injure

The church bells chime which makes the children chuckle.

Joe's job is just a joke.

The judge ordered a major judgment against the juvenile.



Add your own "ch" or "j" sentences below. Put your key words into a tongue twister and practice that sentence often throughout the day.

## Appendix D, English Sound Identification Test

이 테스트의 결과는 여러분이 EIC 기간 중에 참여하게 될 영어 발음 교육의 기초 자료로 사용될 예정입니다. 이 테스트의 점수는 본 EIC 성적은 물론이고 생도 학과 어떤 과목의 성적에도 영향을 미치지 않습니다. 이 테스트는 여러분의 영어 청취 혹은 발음 능력을 진단하여 EIC 발음 수업간 여러분의 영어 발음 향상을 위해 사용되어집니다. 그러므로 답을 모르겠는 경우 추측으로 답을 선택하지 마시고 “모르겠음”을 선택하시기 바랍니다. 테스트를 마치면 테스트 결과가 컴퓨터에 자동으로 저장됩니다. 테스트 중에 혹시 테스트 실시와 관련된 문제가 발생하면 손을 들어 알려 주시기 바랍니다.

지금부터 컴퓨터 스크린에 떠 있는 프로그램 창에 필요한 정보를 입력하세요.

문제는 총 16 문제입니다. 카세트에서 읽어 주는 문장을 듣고 컴퓨터 화면에 표시된 문장과 비교하여 괄호 안에 들어가야 할 단어에 해당하는 답을 a, b, c, d, 혹은 e 중에 골라 확인 버튼을 누르시면 답이 저장됩니다. 다시 한 번 강조하지만 답을 모를 경우, 추측하여 선택하지 말고, “모르겠음”에 해당되는 알파벳을 선택하여 입력하시기 바랍니다.

지금부터 테스트를 시작합니다.

1. Wasn't it a wonderful ( )?

a. bid<sup>44</sup> b. bead c. 모르겠음

2. This time, I don't want to take that ( ).

a. road b. load c. 모르겠음

---

<sup>44</sup> Underlined choices are the right answer to each question.

3. I am so sorry that they have all been ( ).

a. caught b. cut c. 모르겠음

4. I don't understand why it's a bad idea to let them naturally ( ).

a. breathe b. breed c. 모르겠음

5. Wasn't it a wonderful ( )?

a. bid b. bead c. 모르겠음

6. They've been hiding its ( ).

a. flows b. flaws c. 모르겠음

7. At that time, Alice seemed to be obsessed with playing video games, and was very ( ).

a. zealous b. jealous c. 모르겠음

8. Can you spell the word "( )?"

a. measured b. meazard c. masured d. mazard e. 모르겠음

9. At that time, Alice seemed to be obsessed with playing video games, and was very ( ).

a. zealous b. jealous c. 모르겠음

10. This time, I don't want to take that ( ).

a. road b. load c. 모르겠음



11. They've been hiding its ( ).

a. flow    b. flaw    c. 모르겠음

12. I am so sorry that they have all been ( ).

a. caught    b. cut    c. 모르겠음

13. Yesterday, he's finally got a ( ).

a. pass    b. path    c. 모르겠음

14. Can you spell the word "( )?"

a. measured    b. meazard    c. masured    d. mazard    e. 모르겠음

15. Yesterday, he's finally got a ( ).

a. pass    b. path    c. 모르겠음

16. I don't understand why it's a bad idea to let them naturally ( ).

a. breathe    b. breed    c. 모르겠음

### 연구 참여 간 언어 학습 경험 설문<sup>45</sup>

발음 ID No.:

*금번 영어 발음 수업 기간 중 당신의 영어 학습 관련 경험에 대해 생각해 보고 다음에 주어진 문장들에 당신이 얼마나 동의하는지 빈 칸에 적으시오. (1번은 가장 적게 동의하는 경우를 의미하고, 5번은 가장 많이 동의하는 경우를 의미합니다.)*

1. 나의 영어 발음 교육 전 영어 말하기 실력은 우수했다. ----- ( )
2. 나의 영어 발음 교육 전 영어 듣기 실력은 우수했다. ----- ( )
3. 나의 영어 발음 교육 전 영어 쓰기 실력은 우수했다. ----- ( )
4. 나의 영어 발음 교육 전 영어 읽기 실력은 우수했다. ----- ( )
5. 나는 영어 집중 교육 (발음 수업 제외한 나머지)에 열심히 임했다. ----- ( )
6. 나는 4일간 실시한 영어 발음 교육에 열심히 임했다. ----- ( )
7. 영어 발음 특강 기간 (70분씩 4회)은 수업 내용에 비해 짧았다. ----- ( )
8. 영어 발음 특강용 교재(Workbook)는 평상시 한국인으로서 발음하기 어려운 영어 소리들의 발음상 차이나 유사점을 인식하는데 전반적으로 도움이 되었다. ----- ( )
9. 영어 발음 특강용 교재(Workbook)에 제시된 설명은 평상시 한국인으로서 발음하기 어려운 영어 소리들의 발음상 차이나 유사점을 인식하는데 도움이 되었다. ----- ( )
10. 영어 발음 특강용 비디오 자료는 평상시 한국인으로서 발음하기 어려웠던 영어 소리들의 발음상 차이나 유사점을 인식하는데 도움이 되었다. ----- ( )
11. 영어 발음 특강용 오디오 자료는 평상시 한국인으로서 발음하기 어려웠던 영어 소리들의 발음상 차이나 유사점을 인식하는데 도움이 되었다. ----- ( )
12. 영어 발음 특강용 그림 및 사진 자료 (자음/모음 차트, 발성기관 도식, 기타 사진 자료)는 평상시 한국인으로서 발음하기 어려웠던 영어 소리들의 발음상 차이나 유사점을 인식하는데 도움이 되었다. ----- ( )

---

<sup>45</sup> This Korean version is the actual survey taken by the participants.

13. 영어 발음 특강 수업 중 예로서 제시된 영어 단어들은 평상시 한국인으로서 ---- ( )  
발음하기 어려웠던 영어 소리들의 발음상 차이나 유사점을 인식하는데 도움이  
되었다.
14. 영어 집중 교육 이전에 영어 발음과 관련하여 자신의 문제점이 무엇인지 ---- ( )  
구체적으로 알고 있었다.
15. 영어 발음 특강을 통해 영어 발음과 관련하여 자신의 문제점이 무엇인지 ---- ( )  
구체적으로 알게 되었다.
16. 영어 발음 특강을 통해 이전 보다 영어 청취 능력이 향상 되었다. ---- ( )
17. 영어 발음 특강을 통해 이전에 발음이 잘 안 되던 영어 소리를 정확히 낼 수 ---- ( )  
있는 방법을 알게 되었다.
18. 영어 발음 특강 간에 강사의 설명은 이해가 잘 되었다. ---- ( )
19. 영어 발음 특강 강사는 교육 내용에 대한 전문지식이 풍부했다. ---- ( )
20. 영어 발음 특강 강사의 영어 발음은 미국 영어 원어민처럼 정확하다고 신뢰할 ---- ( )  
수 있었다.
21. (한국인 강사 교반에만 해당) 영어 발음 특강 수업을 국어로 진행해서 영어 ---- ( )  
발음에 대한 이해와 향상에 도움이 되었다.
22. 영어 발음 특강 중 제시된 영어 발음 기호 (한국인 강사 교반은 Workbook, ---- ( )  
원어민 강사 교반은 비디오 자료에 제시된 발음기호)는 해당 영어 소리를  
이해하는데 도움이 되었다.
23. (한국인 강사 교반에만 해당) 컴퓨터 사전에서 발췌된 영어 발음 오디오 ---- ( )  
자료는 비디오로 제공된 발음 자료들보다 덜 자연스러운 것도 있었다. 이 두  
가지 종류의 발음법을 대조 제시한 것이 자연스러운 (natural) 영어 발음을  
배우는데 도움이 되었다.
24. (한국인 강사 교반에만 해당) 한국인 강사가 한국어와 영어의 소리를 ---- ( )  
비교/대조하여 설명한 내용은 한국인으로서 발음하기 어려웠던 영어 발음  
향상에 도움이 되었다.
25. (영어 원어민 강사 교반에만 해당) 영어 발음에 대한 설명을 영어로 진행 한 ---- ( )  
것은 영어 발음에 대한 이해와 향상에 도움이 되었다.
26. (영어 원어민 강사 교반에만 해당) 원어민 강사가 강의를 진행하기 위해 ---- ( )  
사용하는 영어 자체를 수업시간에 지속적으로 접했던 것이 (설명의 이해  
측면이 아니고 영어를 계속 들음으로서 오는 영어 발음 input 측면) 영어 발음  
향상에 도움이 되었다.

27. 영어 발음 특강 강사는 학생들의 영어 발음상 어려움을 잘 이해하여 적절히 ---- ( )  
수업 진행에 반영하였다. (예, 발음 요령을 이해 못 하는 경우, 이를 인지하고  
다른 방법으로 설명을 한다든지)
28. 영어 발음 특강을 통해 배운 내용들이 다른 영어 집중 교육 수업에 도움이 ---- ( )  
되었다.
29. 영어 발음 특강을 통해 영어 발음에 자신감이 생겼다. ---- ( )
30. 영어 발음 특강 수업에서 소개된 내용들은 그 이전의 영어 수업 (직전 EIC, ---- ( )  
혹은 대학 입학 전 영어 수업)에서 들어 보지 못한 생소한 내용이 있었다.

**대단히 감사합니다.**

## Language Learning Experience Questionnaire<sup>46</sup>

Pronunciation ID No.:

*Read each of the following statements and write your answers in the blanks based on how you agree on the statement concerning your language learning experience through the English pronunciation classes. (1 means that you least agree, 5 means that you most agree.)*

1. Before the current English pronunciation classes, I had an excellent English speaking ability. ----- ( )
2. Before the current English pronunciation classes, I had an excellent English listening ability. ----- ( )
3. Before the current English pronunciation classes, I had an excellent English writing ability. ----- ( )
4. Before the current English pronunciation classes, I had an excellent English reading ability. ----- ( )
5. I worked hard in the English intensive classes other than in the pronunciation classes. ----- ( )
6. I worked hard to improve my English pronunciation in the pronunciation classes. ----- ( )
7. The total length of the pronunciation classes (70 min., 4 times) was short considering the contents of the classes. ----- ( )
8. The English pronunciation workbook used in the pronunciation classes was helpful in general in recognizing differences and similarities between English sounds that are thought to be hard for Koreans to pronounce. ----- ( )
9. The explanations presented in the workbook were helpful in general in recognizing differences and similarities between English sounds that are thought to be hard for Koreans to pronounce. ----- ( )
10. The video materials presented in the pronunciation classes were helpful in general in recognizing differences and similarities between English sounds that are thought to be hard for Koreans to pronounce. ----- ( )

---

<sup>46</sup> This English version is provided for the reader's convenience.

11. The audio materials presented in the pronunciation classes were helpful in general in recognizing differences and similarities between English sounds that are thought to be hard for Koreans to pronounce. ----- ( )
12. The pictures and photographs (consonant/vowel charts, pronunciation organ illustrations, and other photographs) presented in the pronunciation classes were helpful in general in recognizing differences and similarities between English sounds that are thought to be hard for Koreans to pronounce. ----- ( )
13. The English words presented as examples in the pronunciation classes, were helpful in general in recognizing differences and similarities between English sounds that are thought to be hard for Koreans to pronounce. ----- ( )
14. Before the English Intensive Course, I knew specifically what problems I had concerning English pronunciation. ----- ( )
15. Through the pronunciation classes, I came to know what problems I had concerning English pronunciation. ----- ( )
16. Through the pronunciation classes, my English listening comprehension has improved. ----- ( )
17. Through the pronunciation classes, I came to know how to pronounce English sounds that were hard for me to say before the classes. ----- ( )
18. The oral explanations given to you by your instructor were easy to understand. ----- ( )
19. Your instructor seemed to have an abundant amount of professional knowledge on the taught area. ----- ( )
20. The pattern of your instructor's English pronunciation was considered to be nativelike and thus reliable. ----- ( )
21. (Only for the students taught by Korean instructors) Conducting English pronunciation classes in Korean was helpful for understanding the English sound system and for improving my pronunciation. ----- ( )
22. The IPA (International Phonetic Alphabet) symbols presented in the pronunciation classes (in the workbook for Korean-instructor classes, in the video tutorials for the American-instructor classes) were helpful for understanding the associated English sounds. ----- ( )
23. (Only for the students taught by Korean instructors) Some of the audio materials from the computer English dictionary sounded less natural than the materials from the video clips. Contrasting these two patterns of English pronunciation was helpful for learning natural English pronunciation patterns. ----- ( )

24. (Only for the students taught by Korean instructors) The explanations given by your Korean instructor of English on the comparison and contrast between Korean and English sounds were helpful for improving your pronunciation of the English sounds that are thought to be hard for Koreans to pronounce. ----- ( )
25. (Only for the students taught by native English-speaking instructors) Giving explanations on English pronunciation in English was helpful for understanding the English sound system and for improving my pronunciation. ----- ( )
26. (Only for the students taught by native English-speaking instructors), Being exposed to English utterances (as the language input that resulted from extensively listening to the instructor's utterances as model pronunciation instances) that was spoken to conduct the lecture was helpful for improving my English pronunciation ability. ----- ( )
27. Your pronunciation instructor understood the students' difficulty in English pronunciation and properly reflected it on his/her classroom conduct. (For example, when a student did not understand how to pronounce certain sounds, the instructor noticed this problem and tried to explain about it in a different way.) ----- ( )
28. The items that you learned from the pronunciation classes were helpful when you took other courses in the English intensive program. ----- ( )
29. I became confident in English pronunciation through the English pronunciation classes. ----- ( )
30. They taught some new items in the pronunciation classes about which I had not heard while I was taking English courses before this pronunciation classes (for example, in other English Intensive course classes before the pronunciation classes, or in pre-college English classes). ----- ( )

***Thank you very much!***

## Appendix F, Real-Word Reading-Aloud Test

### 발음 진단 평가 A<sup>47</sup>

※ 이 평가지에는 아무 내용도 표시하지 마세요. 다른 학생들이 반복해서 사용합니다.

다음에 제시되는 단어를 미국인들이 회화에서 발음하듯 자연스럽게 말해 주세요. 영어 강사들이 스펠링 시험을 볼 때처럼 과도하게 또박또박 말할 필요는 없습니다. 만약, 각 번호의 단어를 읽고 나서, 실수로 자신이 평상시 말하는 것과 다르게 잘 못 읽었다고 판단되면 그냥 바로 이어서 한 번을 다시 읽어도 됩니다. 이 경우 맨 마지막에 읽은 것을 진단평가에 사용하겠습니다. 한 단어 녹음이 끝나면 "다음 단어로 진행"을 클릭해서 다음 단어로 넘어가세요.

\* 발음교반 학생들은 자신의 발음 교반 ID No.를 영어로 두 번 말해 녹음하세요.

"예, A-09 ==> A nineteen"

이제부터 다음 단어를 읽어 나갑니다. 단어당 간격은 1-2초간 둡니다.

- |              |                  |
|--------------|------------------|
| 1. bath      | 20. meter        |
| 2. bathing   | 21. mirage       |
| 3. car       | 22. organization |
| 4. cocoa     | 23. pretty       |
| 5. cole slaw | 24. propane      |
| 6. desperate | 25. quiz         |
| 7. disco     | 26. robot        |
| 8. e-mail    | 27. running      |
| 9. faith     | 28. seal         |
| 10. fashion  | 29. smile        |
| 11. fever    | 30. soda         |
| 12. file     | 31. summer       |

<sup>47</sup> The underlined part of each word in the list indicates the segmental sound that was checked for accuracy. Actual testing sheets given to the participants did not have any of the underlined parts shown in this appendix.



13. filler

14. flying

15. jogging

16. Karate

17. label

18. leisure

19. mammoth

32. their

33. thought

34. Venice

35. water

36. well-being

37. zipper

38. zoo

평가가 끝났습니다. 감사합니다.

Appendix G, Artificial-Word Reading-Aloud Test

발음 진단 평가 B

※ 이 평가지에는 아무 내용도 표시하지 마세요. 다른 학생들이 반복해서 사용합니다.

다음에 제시되는 단어를 미국인들이 회화에서 발음하듯 자연스럽게 말해 주세요. 영어 강사들이 스펠링 시험을 볼 때처럼 과도하게 또박또박 말할 필요는 없습니다. 만약, 각 번호의 단어를 읽고 나서, 실수로 자신이 평상시 말하는 것과 다르게 잘 못 읽었다고 판단되면 그냥 바로 이어서 한 번을 다시 읽어도 됩니다. 이 경우 맨 마지막에 읽은 것을 진단평가에 사용하겠습니다. 한 단어 녹음이 끝나면 "다음 단어로 진행"을 클릭해서 다음 단어로 넘어가세요.

\* 발음교반 학생들은 자신의 발음 교반 ID No.를 영어로 두 번 말해 녹음하세요.

"예, A-09 ==> A nineteen"

이제부터 다음 단어를 읽어 나갑니다. 단어당 간격은 1-2초간 둡니다.

- |                     |                       |
|---------------------|-----------------------|
| 1. bam <u>m</u> ith | 21. na <u>b</u> el    |
| 2. be <u>e</u> l    | 22. na <u>r</u>       |
| 3. blo <u>s</u> ure | 23. na <u>t</u> er    |
| 4. bly <u>i</u> ng  | 24. ne <u>a</u> l     |
| 5. d <u>i</u> ller  | 25. nes <u>c</u> o    |
| 6. dreat <u>y</u>   | 26. nira <u>g</u> e   |
| 7. <u>f</u> ime     | 27. pho <u>u</u> ght  |
| 8. <u>f</u> ishion  | 28. pla <u>w</u>      |
| 9. fum <u>m</u> el  | 29. ploa <u>t</u>     |
| 10. ja <u>v</u> or  | 30. que <u>t</u>      |
| 11. kee <u>i</u> ng | 31. ru <u>p</u> ping  |
| 12. ko <u>t</u>     | 32. simple <u>t</u> e |
| 13. lai <u>l</u>    | 33. sni <u>l</u> e    |

14. math
15. mathing
16. mansiderate
17. meather
18. mipper
19. moda
20. mogging

34. Tarate
35. trapization
36. tropane
37. Valice
38. wath
39. zol

평가가 끝났습니다. 감사합니다.

영어 학습 경험 설문<sup>48</sup>

발음 ID No.: \_\_\_\_\_

성별: 남 / 여

아래 질문의 빈칸에 답해 주시기 바랍니다. '/' 로 구분되어진 경우 본인에 해당하는 답에 동그라미를 그려주세요.

1. 영어권 국가에서 거주해 본 경험은? ( ) 년 ( ) 개월, 국가명: ( ), 거주 시작 시 나이: ( )
2. 외국어 고등학교나 국제학교 등 일반 수업에 영어를 사용하는 학교 재학 여부는? 있다 / 없다, 학교명 ( )
3. 처음 영어 알파벳을 배운 시기는? ( ) 세, ( )년, 유치원 이전 / 유치원 / 초등학교 / 중학교 / 고등학교. ( ) 학년
4. 영어를 처음 배우기 시작한 시기는? ( ) 세, ( )년, 유치원 이전 / 유치원 / 초등학교 / 중학교 / 고등학교. ( ) 학년
5. 학교에서 영어수업을 받은 것은 몇 학기인가요? 초등학교 ( ) 학기, 중학교 ( ) 학기, 고등학교 ( ) 학기, 대학교 ( ) 학기
6. 위에 적은 영어수업을 받은 학기 중 영어원어민이 가르쳤던 학기는 몇 학기인가요? 초등학교 ( ) 학기, 중학교 ( ) 학기, 고등학교 ( ) 학기, 대학교 ( ) 학기
7. 다음 영어 능력 중 본인이 발전시키고 싶은 우선순위를 적어 주세요. (1 = 최우선으로 희망하는 분야, 4 = 가장 우선순위상 뒤에 있는 분야)  
말하기: ( ), 듣기: ( ), 쓰기: ( ), 읽기: ( )
8. 하루에 영어를 듣는 총 시간은? 10분 미만 / 10분 이상 - 30분 미만 / 30분 이상 - 1시간 미만 / 1시간 이상 - 2시간 미만 / 2시간 이상 - 4시간 미만 / 4시간 이상
9. 하루에 영어를 말하는 총 시간은? 10분 미만 / 10분 이상 - 30분 미만 / 30분 이상 - 1시간 미만 / 1시간 이상 - 2시간 미만 / 2시간 이상 - 4시간 미만 / 4시간 이상
10. 하루에 영어를 읽은 총 시간은? 10분 미만 / 10분 이상 - 30분 미만 / 30분 이상 - 1시간 미만 / 1시간 이상 - 2시간 미만 / 2시간 이상 - 4시간 미만 / 4시간 이상
11. 하루에 영어로 글을 쓰는 총 시간은? 10분 미만 / 10분 이상 - 30분 미만 / 30분 이상 - 1시간 미만 / 1시간 이상 - 2시간 미만 / 2시간 이상 - 4시간 미만 / 4시간 이상

<sup>48</sup> This Korean version was the actual survey that the participants took. See next page for the same survey translated into English.

## English Education Background Survey<sup>49</sup>

**Pronunciation Course ID No.:** \_\_\_\_\_

**Gender:** Male / Female

*Fill in the parentheses with appropriate answers to the following questions. Draw a circle around your answer if divided by '/'.*

1. How long have you lived in an English-speaking country? ( ) year(s) ( ) month(s), name of the country: ( ), your age when you started living there
2. Have you attended special schools, such as foreign language high schools or international schools, where English was the medium of instruction? yes / no, name of the school: ( )
3. At what age, in what year, at what grade did you first learn English alphabet? at the age of ( ), in (19 ), prior to pre-K / pre-K / elementary / junior high / senior high in the grade of ( )
4. When did you start learning English at school? at the age of ( ), in (19 ), prior to pre-K / pre-K / elementary / junior high / senior high in the grade of ( )
5. How many semesters have you taken English courses in Korea? elementary school: ( ) semesters, middle school: ( ) semesters, high school: ( ) semesters, college: ( ) semesters
6. Out of the total semesters above, how many semesters were taught by native speakers of English? elementary school: ( ) semesters, middle school: ( ) semesters, high school: ( ) semesters, college: ( ) semesters
7. Prioritize the skills of English that you would like to improve using numbers 1 through 4 (1 being the most you would like, 4 being the least you would like)  
speaking ( ), listening ( ), writing ( ), reading ( )
8. How many hours per day do you hear English? less than 10 min. / 10 min. - 30 min. / 30 min. - 1 hour / 1 hour - 2 hrs. / 2 hrs. - 4 hrs. / more than 4 hrs.  
How many hours per day do you speak English? less than 10 min. / 10 min. - 30 min. / 30 min. - 1 hour / 1 hour - 2 hrs. / 2 hrs. - 4 hrs. / more than 4 hrs.  
How many hours per day do you read in English? less than 10 min. / 10 min. - 30 min. / 30 min. - 1 hour / 1 hour - 2 hrs. / 2 hrs. - 4 hrs. / more than 4 hrs.  
How many hours per day do you write in English? less than 10 min. / 10 min. - 30 min. / 30 min. - 1 hour / 1 hour - 2 hrs. / 2 hrs. - 4 hrs. / more than 4 hrs.

---

<sup>49</sup> The original survey, which was written in Korean, has been translated into this English version for readers' convenience.

Appendix I, Grading Sheets for the Reading-Aloud Tests

1. Grading Sheet for Real Words

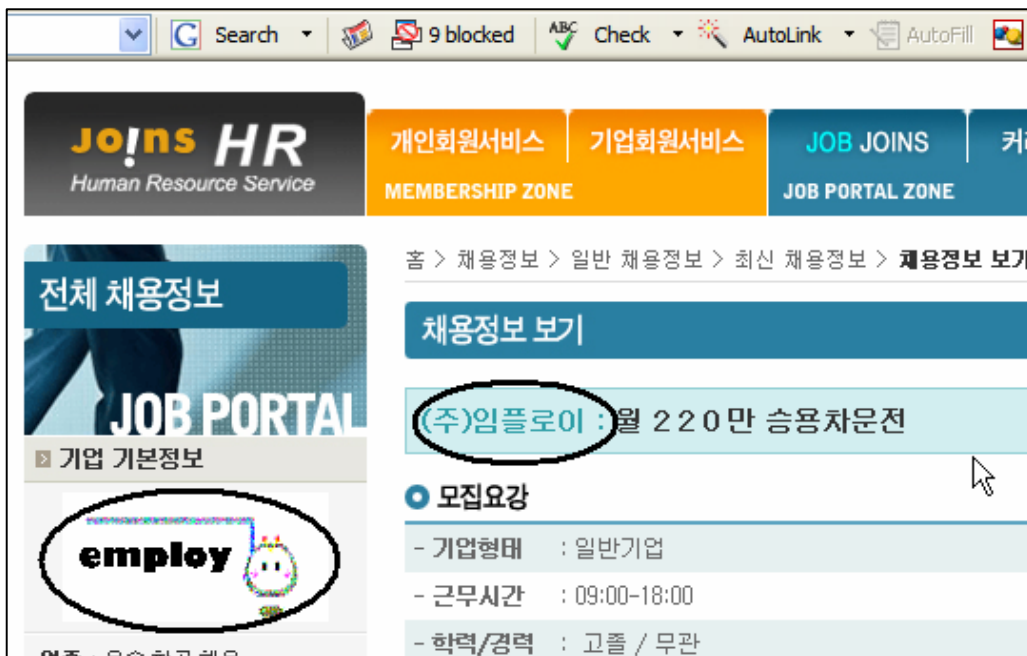
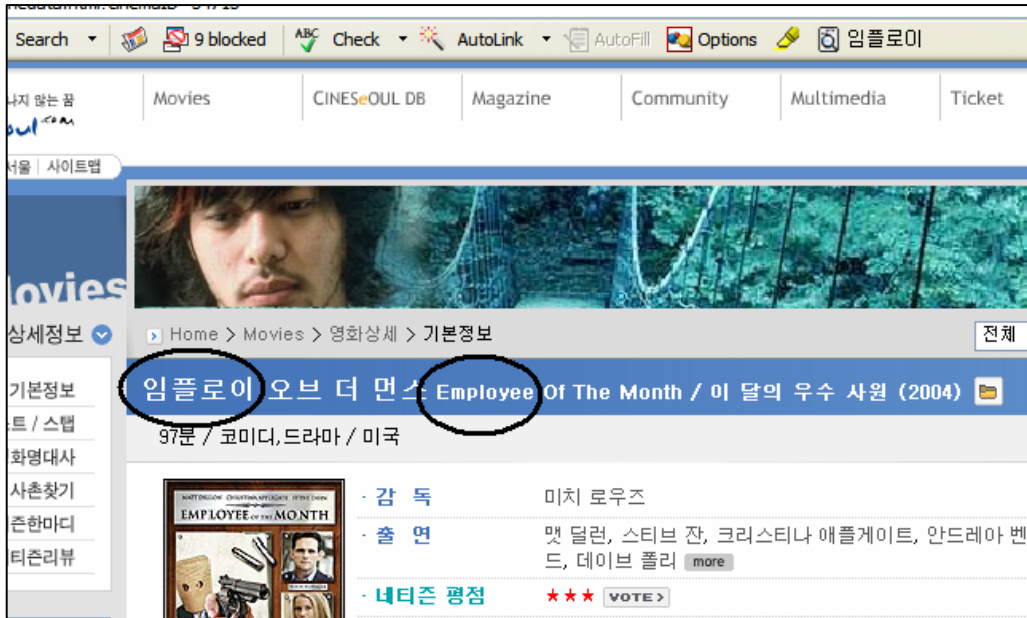
Words	Grading Criteria	A-01 James	A-02 James	A-03 <sup>50</sup> James
1. <u>ba</u> th	θ:3 s:1 t <sup>h</sup> :0			
2. ba <u>th</u> ing	ð:3 d:1 θ:0 s:0			
3. <u>ca</u> r	r:3 ɪ:2 l:1 Ø:1			
4. <u>co</u> coa	ou:3 oa:1			
5. <u>co</u> leslaw	ɔ:3 ou:1 au:1 o:1			
6. <u>de</u> sperate	r:3 er:1 e:0			
7. <u>dis</u> co	ou:3 o:1 a:0			
8. e- <u>ma</u> il	t:3 l:1			
9. <u>fa</u> ith	θ:3 s:1 t <sup>h</sup> :0			
10. <u>f</u> ashion	f:3 p:1			
11. <u>fe</u> ver	v:3 b:1			
12. <u>fi</u> le	f:3 p:1			
13. <u>fi</u> ller	r:3 i:1			
14. <u>jo</u> gging	ɑ:3 a:2 o:1 ʌ:0 ə:0			
15. <u>Ka</u> rate	i:3 r:2 e:1 ε:1			
16. <u>la</u> bel	er:3 a:1 e, ε:0			
17. <u>le</u> isure	ʒ:3 dʒ:1 s,ʃ:0			
18. <u>me</u> ter	i:3 r:2 e:1 ε:1			
19. <u>mi</u> rage	ʒ:3 dʒ:1 z:0 s,ʃ:0			
20. <u>orga</u> nization	z:3 s:2 dʒ:1			
21. <u>pre</u> tt <u>y</u>	r:3 d:2 t:2			
22. <u>prop</u> ane	er:3 a:1 e:1 ε:1			
23. <u>qu</u> iz	kw:3 k <i>with unrounded</i> u:1			
24. <u>ru</u> nn <u>ing</u>	r:3 r:1 l:0			
25. <u>se</u> al	i:3 r:1			
26. <u>sm</u> ile	t:3 l:1			
27. <u>so</u> da	ə:3 a:1			
28. <u>the</u> ir	ð:3 d:1			
29. <u>thou</u> ght	ɔ:3 o:1 ou:0 au:0			
30. <u>V</u> enice	v:3 b:1			
31. <u>wa</u> ter	r:3 t:2 d:2			
32. well- <u>be</u> ing	ir:3 ɪ:2 ii:2 i, r:1			
33. <u>zi</u> pper	r:3 i:1			
34. <u>z</u> oo	z:3 s:2 dʒ:1			

<sup>50</sup> Actual grading sheets listed 10 pronunciation ID codes per page.

## 2. Grading Sheet for Artificial Words

Words	Grading Guides	A-01 Chris	A-02 Chris	A-03 Chris
1. <u>bee</u> l	i:3 r:2 ε:1			
2. blo <u>su</u> re	ʒ:3 dʒ:1 s,ʃ:0			
3. <u>dill</u> er	r:3 i:1			
4. <u>dreat</u> y	r:3 d:2 t:2			
5. <u>fime</u>	f:3 p:1			
6. <u>fish</u> ion	f:3 p:1			
7. <u>javor</u>	v:3 b:1			
8. <u>keci</u> ng	ir:3 r:2 ii:2 i, r:1			
9. <u>ko</u> t	ɑ:3 a:2 o:1			
10. <u>lai</u> l	t:3 l:1 Ø:1			
11. <u>maith</u>	θ:3 s:1 t <sup>h</sup> :0			
12. <u>maithi</u> ng	ð:3 d:1 θ:0 s:0			
13. <u>meath</u> er	ð:3 d:1 θ:0 s:0			
14. <u>mip</u> per	r:3 i:1			
15. <u>moda</u>	ə:3 a:1			
16. <u>mog</u> ging	ɑ:3 o:1 ʌ:0 ə:0			
17. <u>nabel</u>	er:3 a:1 e, ε:0			
18. <u>nar</u>	r:3 ɪl:2 l:1 Ø:1			
19. <u>nater</u>	r:3 t:2 d:2			
20. <u>neal</u>	i:3 r:1			
21. <u>nesco</u>	ou:3 o:1 a:0			
22. <u>nirage</u>	ʒ:3 dʒ:1 z:0 s,ʃ:0			
23. <u>phoug</u> ht	ɔ:3 o:1 ou:0 au:0			
24. <u>ploa</u> t	ou:3 o:1 a:0			
25. <u>queet</u>	kw:3 k <i>with unrounded</i> u:1			
26. <u>rapping</u>	r:3 r:1 l:0			
27. <u>simple</u> te	i:3 r:2 e:1 ε:1			
28. <u>snile</u>	t:3 l:1 Ø:1			
29. <u>trapiza</u> tion	z:3 s:2 dʒ:1			
30. <u>tropane</u>	er:3 a:1 e:1 ε:1			
31. <u>Valice</u>	v:3 b:1			
32. <u>wath</u>	θ:3 s:1 t <sup>h</sup> :0			
33. <u>zoul</u>	z:3 s:2 dʒ:1			

Appendix J, Example of Vowel Contraction<sup>51</sup> in Korean Words Loaned From English



<sup>51</sup> Due to this tendency, the two different English words 'employ' and 'employee' are often transcribed identically in Korean.



Appendix K, Estimated Marginal Means and Interpretations of Reading-Aloud Results

		ANCOVA					
Principal Sound & Interpretation	Composition of Words		Exp.	Cont.	Covariate	<i>p</i>	Eta <sup>2</sup>
[z]	Real	Marginal Mean	1.985	1.416	1.399	0.000	0.22
Category A <sup>52</sup>		Std. Error	0.083	0.075			
Desirable ++	Artificial	Marginal Mean	1.801	1.272	1.355	0.000	0.18
Transfer <sup>53</sup>		Std. Error	0.088	0.078			
[kw]	Real	Marginal Mean	1.768	1.188	1.255	0.000	0.16
Category A		Std. Error	0.103	0.093			
Desirable ++	Artificial	Marginal Mean	1.758	1.230	1.129	0.001	0.11
Transfer		Std. Error	0.118	0.105			
[ou]	Real	Marginal Mean	1.938	1.310	1.250	0.000	0.24
Category B		Std. Error	0.088	0.079			
Desirable +	Artificial	Marginal Mean	1.450	1.193	1.129	0.049	0.04
No Transfer		Std. Error	0.096	0.086			
[t]	Real	Marginal Mean	1.736	1.328	1.410	0.007	0.08
Category C		Std. Error	0.109	0.098			
Desirable +	Artificial	Marginal Mean	1.377	0.933	1.048	0.002	0.11
Weak Transfer		Std. Error	0.101	0.090			
[er]	Real	Marginal Mean	1.925	1.962	1.862	0.945	< 0.01
Category C		Std. Error	0.111	0.100			
Unproblematic <sup>54</sup>	Artificial	Marginal Mean	2.323	2.082	2.075	0.100	0.03
		Std. Error	0.109	0.097			
[ɑ]	Real	Marginal Mean	1.838	1.765	1.543	0.692	< 0.01
Category C		Std. Error	0.134	0.120			
Unproblematic	Artificial	Marginal Mean	1.369	1.017	1.231	0.030	0.05
		Std. Error	0.117	0.103			

<sup>52</sup> Refer to Table 26 (p. 160) for detailed descriptions of the categories of teaching implications based on the interpretations of the ANCOVA results.

<sup>53</sup> ‘Transfer’ means the transfer of pronunciation skills attained through the experiment to the pronunciation of novel words.

<sup>54</sup> An observed result is considered to be unproblematic given that the marginal mean is determined as high.

Appendix K, Estimated Marginal Means and Interpretations of Reading-Aloud Results – Cont'd

Principal Sound & Interpretation	Composition of Words		ANCOVA				
			Exp.	Cont.	Covariate	<i>p</i>	Eta <sup>2</sup>
[iɪ] Category C Unproblematic	Real	Marginal Mean	2.108	1.913	1.830	0.225	0.02
		Std. Error	0.119	0.107			
	Artificial	Marginal Mean	2.238	2.082	1.753	0.324	0.01
		Std. Error	0.118	0.105			
[ɪ]	Real	Marginal Mean	1.884	1.757	1.862	0.330	0.01
		Std. Error	0.096	0.086			
	Artificial	Marginal Mean	1.957	1.601	1.613	0.006	0.08
		Std. Error	0.095	0.084			
[ɔ]	Real	Marginal Mean	Invalid <sup>55</sup>	Invalid	Invalid	Invalid	Invalid
		Std. Error	Invalid	Invalid			
	Artificial	Marginal Mean	0.977	0.845	0.763	0.140	0.02
		Std. Error	0.067	0.059			
[ɹ]	Real	Marginal Mean	Invalid	Invalid	Invalid	Invalid	Invalid
		Std. Error	Invalid	Invalid			
	Artificial	Marginal Mean	2.268	2.038	2.102	0.030	0.05
		Std. Error	0.077	0.068			
[f]	Real	Marginal Mean	Invalid	Invalid	Invalid	Invalid	Invalid
		Std. Error	Invalid	Invalid			
	Artificial	Marginal Mean	2.823	2.755	2.807	0.512	< 0.01
		Std. Error	0.077	0.068			

<sup>55</sup> Results were marked as “Invalid” in case that there was a statistically significant interaction between Group and Pre-Test Score.

Appendix L, Results of the Survey on Pronunciation Learning During the Experiment

Pronunciation Learning Experience Survey Summary

Survey Item	Grand				Section A				Section B			
	Huh	Lee	Joel	Jen	Huh	Lee	Joel	Jen	Huh	Lee	Joel	Jen
<b>Attitude</b>												
5 . Worked Hard in Entire English Courses	3.00	0.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	0.00	0.00	0.00
6 . Worked Hard in Pronunciation Classes	3.00	0.00	4.00	4.00	4.00	3.00	3.00	4.00	5.00	0.00	0.00	0.00
Average	3.00	0.00	4.00	4.00	4.00	3.50	3.50	4.00	4.50	0.00	0.00	0.00
<b>Quality of Classes</b>												
8 . Workbook, Helpful	4.00	0.00	3.00	4.00	5.00	5.00	3.00	5.00	4.00	0.00	0.00	0.00
9 . Explanations in Workbook, Helpful	3.00	0.00	3.00	3.00	5.00	4.00	3.00	4.00	4.00	0.00	0.00	0.00
10 . Video Clips, Helpful	5.00	0.00	2.00	4.00	5.00	5.00	4.00	4.00	2.00	0.00	0.00	0.00
11 . Audio Clips, Helpful	5.00	0.00	3.00	4.00	5.00	4.00	4.00	5.00	4.00	0.00	0.00	0.00
12 . Pictorial Materials, Helpful	4.00	0.00	1.00	4.00	5.00	5.00	4.00	3.00	4.00	0.00	0.00	0.00
13 . Effective Example Word Choice	4.00	0.00	2.00	4.00	3.00	5.00	5.00	4.00	5.00	0.00	0.00	0.00
18 . Instructor's Explanation, Helpful	3.00	0.00	3.00	4.00	5.00	5.00	5.00	4.00	4.00	0.00	0.00	0.00
19 . Professional Knowledge Looking	5.00	0.00	5.00	4.00	3.00	4.00	5.00	5.00	5.00	0.00	0.00	0.00
22 . IPA Symbol Transcriptions, Helpful	4.00	0.00	2.00	3.00	2.00	2.00	3.00	4.00	2.00	0.00	0.00	0.00
27 . Proper ZPD Management	4.00	0.00	1.00	4.00	5.00	4.00	0.00	4.00	3.00	0.00	0.00	0.00
7 . Appropriate Length of Class (7 reverse)	2.00	6.00	3.00	3.00	3.00	2.00	3.00	2.00	2.00	6.00	6.00	6.00
Average	3.91	0.55	2.55	3.73	4.18	4.09	3.55	4.00	3.55	0.55	0.55	0.55
<b>L1 or L2 as Instruction Media</b>												
21 . Korean Helpful	4.00	0.00	3.00	0.00	0.00	0.00	4.00	5.00	5.00	0.00	0.00	0.00
25 . English Helpful	0.00	0.00	0.00	4.00	4.00	3.00	0.00	0.00	0.00	0.00	0.00	0.00
<b>Trustable Pronunciation of Instructor</b>												
20 . Nativelike Pronunciation	4.00	0.00	3.00	4.00	3.00	4.00	5.00	4.00	5.00	0.00	0.00	0.00
<b>Learning Outcome</b>												
15 . Degree of Problem Recognition	4.00	0.00	2.00	4.00	4.00	5.00	4.00	5.00	3.00	0.00	0.00	0.00
16 . Improved LC	3.00	0.00	3.00	4.00	4.00	4.00	3.00	3.00	2.00	0.00	0.00	0.00
17 . Learned How to Pronounce Precisely	3.00	0.00	1.00	4.00	4.00	5.00	4.00	3.00	4.00	0.00	0.00	0.00
28 . Helpful for Other English Classes	4.00	0.00	2.00	3.00	4.00	4.00	4.00	3.00	3.00	0.00	0.00	0.00
29 . Confidence in Pronunciation	4.00	0.00	3.00	3.00	3.00	5.00	4.00	4.00	4.00	0.00	0.00	0.00
Average	3.60	0.00	2.20	3.60	3.80	4.60	3.80	3.60	3.20	0.00	0.00	0.00
<b>Problem Recognition Gain</b>												
14 . Degree, Before	4.00	0.00	1.00	2.00	2.00	4.00	3.00	4.00	3.00	0.00	0.00	0.00
15 . Degree, After	4.00	0.00	2.00	4.00	4.00	5.00	4.00	5.00	3.00	0.00	0.00	0.00
Difference	0.00	0.00	1.00	2.00	2.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
<b>Contative Approach</b>												
23 . Nativelike vs Enunciative, Helpful	3.00	0.00	1.00	0.00	0.00	0.00	5.00	4.00	3.00	0.00	0.00	0.00
24 . Korean vs English Sound Systems, Helpful	4.00	0.00	1.00	0.00	0.00	0.00	3.00	5.00	4.00	0.00	0.00	0.00
Average	3.50	0.00	1.00	0.00	0.00	0.00	4.00	4.50	3.50	0.00	0.00	0.00
<b>Trustability of Instructor's Pronunciation</b>												
20 . Reliable Pronunciation	4.00	0.00	3.00	4.00	3.00	4.00	5.00	4.00	5.00	0.00	0.00	0.00
<b>New Items</b>												
30 . Brand New Learning Items	4.00	0.00	5.00	3.00	5.00	5.00	5.00	5.00	5.00	0.00	0.00	0.00
<b>Language Input</b>												
26 . Helpful English (L2) Input from Instructor	0.00	0.00	0.00	4.00	4.00	5.00	0.00	0.00	0.00	0.00	0.00	0.00

## BIBLIOGRAPHY

- Acton, W. (1984). Changing fossilized pronunciation. *TESOL Quarterly*, 18(1), 71.
- Adamson, D. (1984). Monitoring and the Monitor Model: Labov versus Krashen. *WATESOL Working Papers*(1), 21-29.
- Ano, K. (1998). An experimental study of Output Hypothesis: The role of Output in second language learning. *Journal of Japan-Korea Association of Applied Linguistics*, 1, 63-72.
- Asia Society. (1995). *Tune in Japan: Approaching culture through television. A teacher's guide*. New York: Asia Society Education Department.
- Best, C. T. (1995). Learning to perceive the sound pattern of English. In C. Rowe-Collier, L. P. Lipsitt & H. Hayne (Eds.), *Advances in infancy research* (Vol. 9, pp. 217-304). Norwood: Ablex.
- Best, C. T., McRoberts, G. W., & 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.
- Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, 44(2), 235-249.

- Bonk, C. J., & Cunningham, D. J. (1998). Searching for learner-centered, constructivist, and sociocultural components of collaborative educational learning tools. In C. J. Bonk & K. S. King (Eds.), *Electronic collaborators: Learner-centered technologies for literacy, apprenticeship, and discourse*. Mahwah, NJ: Erlbaum.
- Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, *101*, 2299-2310.
- Brown, H. D. (1994). *Principles of language learning and teaching* (3rd ed.). Eaglewood Cliffs, NJ: Prentice Hall Regents.
- Brown, H. D. (1994). *Teaching by principles: An interactive approach to language pedagogy*. Upper Saddle River, NJ.: Prentice Hall Regents.
- Canagarajah, A. S. (1999). On EFL teachers, awareness, and agency. *ELT Journal*, *53*(3), 207-213.
- Celce-Murcia, M., Brington, D., & Goodwin, J. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages* (1 ed.). New York: Cambridge University Press.
- Chan, C., Burtis, J., & Bereiter, C. (1997). Knowledge building as a mediator of conflict in conceptual change. *Cognition and Instruction*, *15*(1), 1-40.

- Ciske, M. D. (1984). The Monitor Model: More evidence. *WATESOL Working Papers*(1), 30-35.
- Clarke, C. M. (2000). Perceptual adjustment to foreign-accented English. *Journal of the Acoustical Society of America*, 107(5), 2856.
- Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America*, 116(6), 3647-3658.
- Cook, V. (1999). Going beyond the native speaker in language teaching. *TESOL Quarterly*, 33(2), 185-209.
- Corder, S. P. (1971). Idiosyncratic dialects and error analysis. *International Review of Applied Linguistics in Language Teaching*, 9(2), 147-160.
- Crawford, W. W. (1987). The pronunciation monitor. In *Current perspectives on pronunciation*.
- Derwing, T. M., & Rossiter, M. J. (2002). ESL learners' perceptions of their pronunciation needs and strategies. *System*, 30(2), 155-166.
- Dieterich, T. G., & Others, A. (1979). A Linguistic analysis of some English proficiency tests. *TESOL Quarterly*, 13(4), 535-550.
- Erickson, F. (1996). Going for the zone: The social and cognitive ecology of teacher-student interaction in classroom conversations. In D. Hicks (Ed.), *Discourse, learning, and schooling* (pp. 29-62). New York: Cambridge University Press.

- Field, J. (2004). An insight into listeners' problems: too much bottom-up or too much top-down? *System*, 32(3), 363-377.
- Flege, J. E. (1981). The phonological basis of foreign accent: A hypothesis. *TESOL Quarterly*, 15(4), 443-455.
- Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance*, 32(5), 1276.
- Gardner, R. C., & MacIntyre, P. D. (1991). An instrumental motivation in language study: Who says it isn't effective? *Studies in Second Language Acquisition*, 13(1), 57-72.
- Garnes, S. (1977). *American students' perception and production of Romanian plosives*.
- Garson, D. (2006). *Reliability analysis*. Retrieved July 11, 2006, from <http://www2.chass.ncsu.edu/garson/pa765/reliab.htm>
- Gilbert, J. (1993). *Clear speech: Pronunciation and listening comprehension in North American English* (2 ed.). New York: Cambridge University Press.
- Goh, C. C. M. (2000). A cognitive perspective on language learners' listening comprehension problems. *System*, 28(1), 55-75.

- Goldstein, L. S. (1999). The relational zone: The role of caring relationships in the co-construction of mind. *American Educational Research Journal*, 36(3), 647-673.
- Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "l" and "r". *Neuropsychologia*, 9, 317-323.
- Guion, S. G., Flege, J. E., Liu, S. H., & Yeni-Komshian, G. H. (2000). Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics*, 21(2), 205.
- Herron, C., Dubreil, S., Cole, S. P., & Corrie, C. (2000). Using instructional video to teach Culture to beginning foreign language students. *CALICO Journal*, 17(3), 395.
- Hilles, S., & Sutton, A. (2001). Teaching adults. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (pp. 385-399). Boston, MA: Heinle & Heinle.
- Horani, L. A. (1995). *The effect of a physician's pronunciation on nurses' perceptions of the physician's medical competency*.
- Huh, J. (2003). *Scaffolding effects of explicit pronunciation instruction in Korean on adult ESL learners' learning of some problematic English sounds*. University of Texas at Austin, Austin, TX.



- Ingram, J. C. L., & Park, S.-G. (1998). Language, context, and speaker effects in the identification and discrimination of English /ɪ/ and /I/ by Japanese and Korean listeners. *Journal of the Acoustical Society of America*, 103(2), 1161-1174.
- Izumi, S. (2002). Output, Input Enhancement, and the Noticing Hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition*, 24(4), 541-577.
- Izumi, S., Bigelow, M., Fujiwara, M., & Fearnow, S. (1999). Testing the Output Hypothesis: Effects of Output on Noticing and second language acquisition. *Studies in Second Language Acquisition*, 21(3), 421-452.
- Johnson, J. S., & Newport, E. L. (1989). Critical Period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60-99.
- Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145-166). San Diego, CA: Academic Press.
- Jones, R. H. (1997). Beyond "listen and repeat": Pronunciation teaching materials and theories of second language acquisition. *System*, 25(1), 103-112.
- Katchen, J. E. (1997). *Off-Air video for high school English classes: Some considerations*.

- Kendrick, H. (1997). Keep them talking! A project for improving students' L2 pronunciation. *System*, 25(4), 545-560.
- Kim, H.-K. (1999, June 25, 1999). *Interlanguage phonology of Korean learners of English*. Retrieved March 6, 2003, from [http://odin.prohosting.com/hkkim/cgi-bin/kaeps/il\\_phon.htm#substitution](http://odin.prohosting.com/hkkim/cgi-bin/kaeps/il_phon.htm#substitution)
- Krashen, S. D. (1981). The "Fundamental Pedagogical Principle" in second language teaching. *Studia Linguistica*, 35(1-2), 50-70.
- Krashen, S. D., & Pon, P. (1975). An error analysis of an advanced learner of ESL: The importance of the Monitor. Working papers on bilingualism, No. 7.
- Kreidler, C. W. (1971). *Teaching English spelling and pronunciation*.
- Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor, Mich: Ann Arbor-University of Michigan Press.
- Laurillard, D. (1995). Multimedia and the changing experience of the learner. *British Journal of Educational Technology*, 26(3), 179.
- Levey, S. (2004). The discrimination and the production of English vowels by bilingual Spanish/English speakers. *Journal of the Acoustical Society of America*, 115(5), 2604.
- Li, D. (1998). "It's always more difficult than you plan and imagine": Teachers' perceived difficulties in introducing the communicative approach in South Korea. *TESOL Quarterly*, 32(4), 677-703.

- MacKain, K. S., Best, C. T., & Strange, W. (1981). Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics*, 2, 369-390.
- Major, R. C., Fitzmaurice, S. F., Bunta, F., & Balasubramanian, C. (2002). The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly*, 36(2), 173-190.
- Maken, M. A., & Ferguson, C. A. (1987). Phonological universals in language acquisition. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology*. Cambridge: Newbury House.
- Maun, I. (1986). Cognates and the Monitor Theory. *British Journal of Language Teaching*, 24(1), 29-33.
- McNerney, M., & Mendelsohn, D. (1987). Putting suprasegmentals in their place. *TESL Talk*, 17(1), 132-140.
- Medgyes, P. (2001). When the teacher is a non-native speaker. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (3rd ed., pp. 429-442). Boston, MA: Heinle & Heinle.
- Meyers, C., & Holt, S. (1998). *Pronunciation for success: Student workbook*. Burnsville, MN: ESL Videos.

- Moll, L. C., Tapia, J., & Whitmore, K. F. (1993). Living knowledge: The social distribution of cultural resources for thinking. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations* (pp. 139-163). Cambridge: Cambridge University Press.
- Morosan, D. E., & Jamieson, D. G. (1989). Evaluation of a technique for training new speech contrasts: Generalization across voices, but not word-position or task. *Journal of Speech and Hearing Research*, 32(3), 501.
- Munro, M. J. (1998). The Effects of noise on the intelligibility of foreign-accented speech. *Studies in Second Language Acquisition*, 20(2), 139.
- Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 49(s1), 285-310.
- Munro, M. J., & Derwing, T. M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate. *Studies in Second Language Acquisition*, 23(4), 451.
- Nemser, W. (1971). Approximative systems of foreign language learners. *International Review of Applied Linguistics in Language Teaching*, 9(2), 115-123.
- Noddings, N. (1984). *Caring*. Berkeley: University of California Press.
- Oller, J. W., Jr., & Ziahosseiny, S. M. (1970). The Contrastive Analysis Hypothesis and spelling errors. *Language Learning*.

- Pica, T., & Others, A. (1996). Language learners' interaction: How does it address the input, output, and feedback needs of L2 learners? *TESOL Quarterly*, 30(1), 59-84.
- Pisoni, D. B., Aslin, R. N., Perey, A. J., & Hennessy, B. L. (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 297-314.
- Ponterotto, D. (1990). Age and efficacy of Monitor use in an EFL classroom setting.
- Prator, C. H. (1971). Phonetics vs. phonemics in the ESL classroom: When is allophonic accuracy important? *TESOL Quarterly*.
- Prawat, R. S., & Floden, R. E. (1994). Philosophical perspectives on constructivist views of learning. *Educational Psychology*, 29(1), 37-48.
- Reves, T., & Medgyes, P. (1994). The non-native English speaking EFL/ ESL teacher's self-image: An international survey. *System*, 23(3), 353-367.
- Richter, U. M. (2003). *Integrating web-based learning into the curriculum*.
- Salomon, G., & Perkins, D. N. (1998). Individual and social aspects of learning. *Review of Research in Education*, 23, 1-24.

- Samimy, K. (1999). To be a native or non-native speaker: Perceptions of "non-native" students in a graduate TESOL program. In G. Braine (Ed.), *Non-native educators in English language teaching* (1st ed., pp. 127-144). Cambridge, MASS.: Newbury House.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 206-226.
- Share, D. L. (2004). Orthographic learning at a glance: On the time course and developmental onset of self-teaching. *Journal of Experimental Child Psychology*, 87(4), 267.
- Sharf, D. J., & Others, A. (1988). Relationship between the discrimination of /w-r/ and /t-d/ continua and the identification of distorted /r/. *Journal of Speech and Hearing Research*, 31(2), 193.
- Scherer, K. R. (1984). Emotion as a multicontent process: A model and some cross-cultural data. In P. Shaver (Ed.), *Review of personality and social psychology: Emotions, relationships, and health* (pp. 37-63). Beverly Hills: Sage Publications.
- Shohamy, E., & Inbar, O. (1988). Construct validation of listening comprehension tests: The effect of text and question type.
- Stockwell, R. P., & Bowen, J. D. (1965). *The sound of English and Spanish*. Chicago: University of Chicago Press.

- Suarez, A., & Meara, P. (1989). The effects of irregular orthography on the processing of words in a foreign language. *Reading in a Foreign Language*, 6(1), 349.
- Swain, M. (1988). Manipulating and complementing content teaching to maximize second language learning. *TESL Canada Journal*, 6(1), 68-83.
- Swain, M. (1993). The Output Hypothesis: Just speaking and writing aren't enough. *Canadian Modern Language Review*, 50(1), 158-164.
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16, 371-391.
- Takagi, N. (2002). The limits of training Japanese listeners to identify English /r/ and /l/: Eight case studies. *Journal of the Acoustical Society of America*, 111(6), 2887-2896.
- Takagi, N., & Mann, V. (1995). The limits of extended naturalistic exposure on the perceptual mastery of English //r/ and /l/ by adult Japanese learners of English. *Applied Psycholinguistics*, 16, 379-405.
- Tatsuki, D. (1999). *Alleviating comprehension problems in movies: Working paper* (No. 0286-0295).
- Truitt, S. N. (1995). Beliefs about language learning: A study of Korean university students learning English. *Texas Papers in Foreign Language Education*, 2(1), 1-14.

- Tsui, A. B. M., & Fullilove, J. (1998). Bottom-up or top-down processing as a discriminator of L2 listening performance. *Applied Linguistics*, 19(4), 432-451.
- VanDijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- VanPatten, B. (1984). Communicative values and information processing in L2 acquisition. Michigan.
- VanPatten, B. (1988). How juries get hung: Problems with the evidence for a Focus on Form in teaching, *Language Learning* (Vol. 38, pp. 243-260).
- VanPatten, B. (1990). Attending to form and content in the input: An experiment in consciousness, *Studies in Second Language Acquisition* (Vol. 12, pp. 287-301).
- VanPatten, B. (2002). Processing the content of Input-Processing and Processing Instruction research: A response to DeKeyser, Salaberry, Robinson, and Harrington, *Language Learning* (Vol. 52, pp. 825-831).
- VanPatten, B. (2002). Processing Instruction, prior awareness, and the nature of second language acquisition: A (partial) response to Batstone, *Language Awareness* (Vol. 11, pp. 240-258).
- VanPatten, B., & Cadierno, T. (1993). Input Processing and second language acquisition: A role for instruction, *Modern Language Journal* (Vol. 77, pp. 45-57).



- VanPatten, B., & Oikkenon, S. (1996). Explanation versus Structured Input in Processing Instruction, *Studies in Second Language Acquisition* (Vol. 18, pp. 495-510).
- Vigil, N. A., & Oller, J. W. (1976). *Rule fossilization: A tentative model*: Language Learning.
- Wardhaugh, R. (1970). The contrastive analysis hypothesis. *TESOL Quarterly*, 4(2), 123-130.
- Weir, R. H. (1962). *Language in the crib*. The Hague: Mouton Publishers.
- Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology*, 24(5), 672.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, 49-63.
- Werker, J. F., & Tees, R. C. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, 75, 1866-1878.
- Wertsch, J. V. (1991). A sociocultural approach to socially shared cognition. In J. M. Levine & S. D. Teaslo (Eds.), *Perspectives on socially shared cognition*. Washington, DC: American Psychological Association.

- Witten, C. (2000). Using video to teach for sociolinguistic competence in the foreign language classroom. *Texas Papers in Foreign Language Education*, 5(1), 143.
- Wu, Y. a. (2001). English language teaching in China: Trends and challenges. *TESOL Quarterly*, 35(1), 191-194.
- Yamada, R. A. (1993). Effects of extended training on /r/ and /l/ identification by native speakers of Japanese. *Journal of the Acoustical Society of America*, 93, 2391(A).
- Yeni-Komshian, G., Johns Hopkins Univ., B., MD. Neurocommunications Lab., & Others, A. (1968). *A pilot study on the ability of young children and adults to identify and reproduce novel speech sounds.*
- Zampini, M. L. (1998). The relationship between the production and perception of L2 Spanish stops. *Texas Papers in Foreign Language Education*, 3(3), 85.
- Zhao, Y. (1997). The effects of listeners' control of speech rate on second language comprehension. *Applied Linguistics*, 18(1), 49-68.

## VITA

Jin Huh was born in 1968 as the eldest son of Sun Huh and Byungsoon Jung in Seoul, The Republic of Korea. He received a B.S. in Computer Science from the Korea Military Academy in 1992, when he was commissioned as a Signal Corps officer. He served in the ROK Army from 1992 to 1996 as a platoon leader, staff officer, and company commander. In 1999, he was granted a B.A. in English Language and Literature by the Seoul National University. After working two and a half years in the Department of English at the Korea Military Academy, he entered the Foreign Language Education Program at the University of Texas at Austin in 2001. He was awarded a Master's degree at the same university in 2003.

Permanent Address: Department of English, Academic Board

Korea Military Academy

Gongneung-dong, Nowon-goo

Seoul 139-799, Republic of Korea

E-mail Address: [jinhuh48@yahoo.com](mailto:jinhuh48@yahoo.com)

This dissertation was typed by the author.