

The Efficacy of Lexical Stress Diacritics  
in the English Comprehensibility and Accentedness of Korean Speakers

by

KEUN KIM

A Thesis Submitted in Partial Fulfillment  
of the Requirements for the Degree of

MASTER OF ARTS

in the Department of Linguistics

© Keun Kim, 2019  
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.

We acknowledge with respect the Lekwungen-speaking peoples on whose traditional territory the university stands and the Songhees, Esquimalt, and WSÁNEĆ peoples whose historical relationships with the land continue to this day.

## **Supervisory Committee**

The Efficacy of Lexical Stress Diacritics  
in the English Comprehensibility and Accentedness of Korean Speakers

by

KEUN KIM

### **Supervisory Committee**

Dr. John Archibald, Department of Linguistics, University of Victoria  
**Supervisor**

Dr. Hossein Nassaji, Department of Linguistics, University of Victoria  
**Departmental Member**

## Abstract

### Supervisory Committee

Dr. John Archibald, Department of Linguistics

Supervisor

Dr. Hossein Nassaji, Department of Linguistics

Departmental Member

In the field of second language (L2) pronunciation research and teaching, relatively less research has been conducted on the efficacy of suprasegmental features (e.g., stress and intonation) than that of segmentals (individual sounds) (Thomson & Derwing, 2015). However, previous studies reported that suprasegmental errors are as much or more responsible for accentedness and comprehensibility ratings than are segmentals (Munro & Derwing, 1995a; Kang, Rubin, & Pickering, 2010).

English stress, a suprasegmental feature, can be very challenging for native Korean speakers to acquire due to the different prosodic systems of the two languages. While English has stress at the lexical level, Korean has tone patterns at the phrasal level known as the *Accentual Phrase* (Jun 1996). It is common that Korean speakers of English place stress on the wrong syllable or do not give sufficient auditory prominence to a stressed syllable in a word like *sápphire*. Korean learners of English are often frustrated by not knowing which syllable gets prominence because the English writing system does not provide this information. This point calls for enhanced input that can make stressed syllable salient for L2 learners.

The purpose of the current study, therefore, was to investigate 30 native Korean subjects to examine the efficacy of providing enhanced input (lexical stress diacritics) on the accentedness and comprehensibility of their L2 English. In the pretest, the participants read aloud 15 English sentences without diacritics. Then, the subjects were

given explicit treatment instructions on the production of increased pitch and extended duration as a marker of English stress with musical notation presented. The participants were invited to read aloud novel sentences written with diacritics to mark stress placement. In the treatment task (immediately following the treatment instructions on the same day), the participants read aloud the same sentences from the pretest but with correct stress placement indicated by diacritics. In the posttest, which took place two days after the pretest and the treatment task, participants read 15 sentences without diacritics again to see if the effects of the treatment were retained. Speech samples from three measurement points were rated by three native speakers of English in terms of comprehensibility and accentedness.

Four main findings are reported. First, significant improvement was found between the pretest and treatment task in both comprehensibility and accentedness ratings. This result indicates that the participants gained immediate benefit from utilizing diacritics with stress correctly placed. Second, significant improvement was observed in the pretest vs. the posttest in both accentedness and comprehensibility ratings, demonstrating the effects of the treatment. Third, there was no significant difference between the treatment task and the posttest, suggesting that the participants retained what they learned from the treatment when diacritics were removed. Fourth, the significant improvement in the pretest vs. the posttest was observed across the participants' level of English proficiency and gender, showing promise as a more generally applicable pronunciation teaching technique. Pedagogical and empirical implications along with limitations of the current study are discussed.

## Table of Contents

Supervisory Committee .....	ii
Abstract .....	iii
Table of Contents .....	v
List of Tables .....	vii
List of Figures .....	viii
Acknowledgments.....	ix
Dedication .....	xii
Chapter 1. Introduction .....	1
1.1. Background .....	1
1.2. Outline.....	2
Chapter 2. Theoretical Background and Literature Review .....	4
2.1. Definitions of Key Terms .....	4
2.1.1. Intelligibility, Comprehensibility, and Accentedness .....	4
2.1.2. Segmentals .....	10
2.1.3. Suprasegmentals (Prosody).....	11
2.2. English Stress.....	12
2.2.1. Types of English Stress.....	12
2.2.2. English Stress and its Impact on Intelligibility, Comprehensibility, and Accentedness.....	14
2.2.3. Stress Parameter Model (SPM).....	16
2.2.4. Teaching English Stress .....	19
2.2.5. Acquisition of L2 English Stress .....	20
2.3. Korean Prosody.....	20
2.3.1. The Accentual Phrase in Korean.....	21
2.3.2. Vowel Length Distinction in Korean.....	24
2.4. Input Enhancement and English Orthography .....	25
2.5. Pronunciation Teaching in Public Education System in Korea .....	27
2.6. Pronunciation and Individual Differences .....	31
2.6.1. Level of L2 Proficiency .....	31
2.6.2. Gender.....	32
2.7. The Present Study .....	32
2.8. Research Questions.....	33
Chapter 3. Methodology .....	34
3.1. Participants.....	34
3.2. Instruments.....	35
3.2.1. English Language Learning Background Information .....	35
3.2.2. English Sentence Reading Materials .....	36
3.3. Data Collection .....	37
3.3.1. Participant Recruitment .....	37
3.3.2. Meeting 1 (Pretest, Treatment Instructions, Treatment Task) .....	38
3.3.2.1. Pretest.....	39
3.3.2.2. Treatment Instructions .....	40
3.3.2.3. Treatment Task .....	42

3.3.3. Meeting 2 (Posttest) .....	44
3.4. Data Preparation.....	45
3.4.1. Inclusion and Exclusion of Data .....	45
3.4.2. Using Data Collected from Secondary Recorder.....	45
3.4.3. Randomizing Recording Files.....	45
3.5. Rating Procedures .....	46
3.5.1. Raters .....	47
3.5.2. Meeting 1 (Comprehensibility Rating) .....	47
3.5.2.1. Rater Training .....	47
3.5.2.2. Comprehensibility Rating .....	50
3.5.3. Meeting 2 (Accentedness Rating).....	51
3.5.3.1. Rater Training .....	51
3.5.3.2. Accentedness Rating.....	53
3.6. Data Analysis .....	53
Chapter 4. Results and Discussion.....	55
4.1. Inter-rater Reliability .....	56
4.2. Comprehensibility.....	58
4.2.1. Results from Total Participants.....	58
4.2.2. Results by Level of L2 Proficiency .....	61
4.2.3. Results by Gender .....	62
4.3. Accentedness.....	64
4.3.1. Results from Total Participants.....	64
4.3.2. Results by Level of L2 Proficiency .....	66
4.3.3. Results by Gender .....	68
4.4. Retention of the Effects of the Treatment.....	70
4.5. Summary of Results .....	71
Chapter 5. Implications, Limitations, and Future Directions .....	73
5.1. Implications.....	73
5.1.1. Pedagogical Implications .....	73
5.1.2. Empirical Implications.....	76
5.2. Limitations and Future Directions .....	77
5.2.1. Lack of Control Group and Group Size .....	77
5.2.2. Ecological Validity .....	78
5.2.3. Lack of Treatment on Reduced Vowel .....	79
5.2.4. Future Directions .....	79
Chapter 6. Conclusions .....	81
Bibliography .....	82
Appendix A English Language Background Questionnaire.....	88
Appendix B Recruitment Notice (English Version & Korean Version).....	89
Appendix C Sample Lesson Plan.....	91

## List of Tables

Table 1 Results of Possible Intelligibility and Comprehensibility Combinations (Derwing and Munro, 2015).....	5
Table 2 Results of Possible Intelligibility and Accentedness Combinations (Derwing and Munro, 2015) .....	5
Table 3 Korean Stops in Word-initial Position (adapted from Cho, Jun, and Ladefoged, 2002) .....	22
Table 4 Korean AP Example Sentence (adapted from Jun, 1996) .....	22
Table 5 Participant Characteristics .....	35
Table 6 List of Sentences without Diacritics for Pretest and Posttest .....	36
Table 7 List of Sentences with Diacritics for Treatment task.....	37
Table 8 Practice Sentences.....	42
Table 9 Order of Speech Samples.....	46
Table 10 Sentences from the practice session.....	50
Table 11 Sample Data Arrangement (Comprehensibility) .....	53
Table 12 Total Scores (Comprehensibility) .....	54
Table 13 Inter-rater Reliability .....	57
Table 14 Means Comprehensibility Scores for Tasks .....	58
Table 15 Tests of Within-Subjects Effects (Comprehensibility) .....	59
Table 16 Pairwise Comparison (Bonferroni).....	60
Table 17 Means by Level of Proficiency for Tasks (Comprehensibility) .....	61
Table 18 Tests of Within-Subjects Effects (TIME*Level: Comprehensibility).....	61
Table 19 Means by Gender for Tasks (Comprehensibility).....	63
Table 20 Tests of Within-Subjects Effects (TIME*Gender_Comprehensibility) .....	63
Table 21 Means Accentedness Scores for Tasks .....	65
Table 22 Pairwise Comparison (Bonferroni).....	65
Table 23 Means by Level of Proficiency for Tasks (Accentedness) .....	67
Table 24 Tests of Within-Subjects Effects (TIME*Level_Accentedness).....	67
Table 25 Means by Gender for Tasks (Accentedness) .....	69
Table 26 Tests of Within-Subjects Effects (TIME*Gender_Accentedness) .....	69
Table 27 Summary of Results.....	72

## List of Figures

Figure 1 Grade 9 English textbook (from NE publishing company).....	28
Figure 2 High school English 1 textbook (from NE publishing company) .....	29
Figure 3 Musical Notation for Treatment Instructions .....	40
Figure 4 Practice Rating Sheet for Comprehensibility .....	49
Figure 5 Praat (Boersma & Weenink, 2019) window showing a waveform display for participant #4 .....	51
Figure 6 Practice Rating Sheet for Accentedness .....	52
Figure 7 Summary of Research Design .....	56
Figure 8 Line Graphs for Mean Comprehensibility Scores .....	60
Figure 9 Line Graphs for Mean Comprehensibility Scores by Level of Proficiency .....	62
Figure 10 Line Graphs for Mean Comprehensibility Scores by Gender .....	64
Figure 11 Line Graphs for Mean Accentedness Scores.....	66
Figure 12 Line Graphs for Mean Accentedness Scores by Level of Proficiency .....	68
Figure 13 Line Graphs for Mean Accentedness Scores by Gender .....	70

## Acknowledgments

Working on a thesis felt like a lonely, solo task at first, but in the end, it was the product made in collaboration with numerous people that I have worked with throughout my master's program.

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr. John Archibald for your continued support and sharing your expertise in the field of second language phonology. I felt extremely privileged to have you as my supervisor. I cannot thank you enough for offering me the directed study, where I was fortunate enough to develop my idea that was the seed of the thesis. I would like to thank you for the opportunity to attend the AAAL 2019 conference, where I was able to interact with researchers who shared similar interests with me. I cannot express my appreciation enough for your contributions of time, ideas, and energy to help me grow and progress as a researcher.

I would like to extend my appreciation to my committee member, Dr. Hossein Nassaji for your insightful, critical yet supportive feedback on my research design and data analysis. The way you delivered the research course equipped me with a firm foundation on how to design a study in the field of applied linguistics. Thank you for your immense contribution to my thesis.

Special thanks to Dr. Murray Munro for providing me with excellent suggestions regarding the interpretation of my research findings.

I would like to thank the following faculty and staff members in the department of linguistics: Dr. Li-Shih Hwang for offering me detailed feedback that shed light on my understanding of applied linguistics and hiring me as a research assistant for your project;

Dr. Suzanne Urbanczyk for helping me expand my scope of understanding regarding prosodic morphology that I was able to incorporate into my thesis; Dr. Ewa Czaykowska-Higgins for making me realize the importance of our own language; Dr. Hua Lin, for letting me teach a class on the Monitor Model in your course and helping me adjust to my life in Victoria; and Jenny Jessa, our awesome graduate secretary, for your genuine care and help.

I am grateful to my friends in the department of linguistics. I thank Dr. Hyekyeong Cheong, whom I spent quality time together sharing research interests. I want to thank Amber for being such a good friend and Sajib, Yiran, Pam, Ruby, Vahid, and Niloo for helping me start a great first year at UVic. I also thank Caroline for being the glue that held us together in the department and Christy, Tess, Mitchell, and Emmanuelle for adding a lot of fun and emotional support to my second year.

I would like to thank friends that I made outside my department: David for making my life in Victoria interesting and your genuine hospitality; Jerry for your insightful input for my mock defence; and Songtao for the nice conversations and meals we shared.

I would like to thank my father, Jaebong Kim, my mother, Kyungae Lee, and my sister, Boreum Kim for all their warm love and encouragement for my study. They believed in me when I was not sure about what I was doing during my study and encouraged me to keep going.

I thank my friends back in South Korea for believing in me and being my guinea pigs for my pilot study: JK, Jude, Y. S. Kang, Sunny, Chris, Songja, H. K. Cho, and Devin K. I also want to thank my colleagues in Korea who were willing to participate in my pilot study: Jihoon Park, Jungwon Lee, and S. H. Lee.

I gratefully acknowledge the funding sources that made my MA work possible: the University Fellowship, the University Awards, the Lucie Daigle Memorial Scholarship in Linguistics, the Henry & Michiko Warkentyne Graduate Scholarships in Linguistics, and the Linguistics Research Fund.

I thank Dr. Madeline Walker, Gillian Saunders, and Nancy Ami, tutors at the Centre for Academic Communication, for providing me with personalized workshop sessions to help me become a competent self-editor and working with me for my mock defence.

Last but not least, I would like to thank the 30 participants for their time commitment and enthusiasm toward the treatment that I provided in my study. Working with them was such a joy and enhanced my awareness of specific areas that I need to look at as a language teacher. This thesis would not have been possible without their dedicated participation.

## Dedication

*To Professor Jordan Baron*

*with admiration*

# Chapter 1. Introduction

## 1.1. Background

As Setter and Jenkins (2005) noted, pronunciation is one of the most important factors in order to achieve successful oral communication. However, teaching second language (L2) pronunciation has often received relatively less attention in a teaching context where communicative competence is highly emphasized (Jenkins, 2000). School teachers and students tend to treat pronunciation as the least useful skill in English as a Foreign Language (EFL) contexts where students are predominantly evaluated by written exams (Szypra-Kozłowska, 2015). It was not until the last few decades that L2 pronunciation has been systematically investigated for its impact on successful communication, but even now relatively less research has been conducted on L2 pronunciation than on other dimensions of language teaching and learning such as grammar and vocabulary (Derwing & Munro, 2005). This implies that there is still much left unexplored on how to teach L2 pronunciation that is beneficial to both L2 teachers and L2 learners.

Many linguistic features such as individual sounds, stress, or intonation can contribute to L2 learners being more or less understandable in their speech. Since L2 teachers and learners have limited time and resources for teaching and learning L2 pronunciation, it is essential to prioritize what elements of the L2 sound system impede or facilitate learners' speaking performance the most (Munro, 2008). If a certain aspect of pronunciation instruction has more promise to improve L2 pronunciation than some other aspects, it would be wise for teachers and learners to pay extra attention to the relevant aspect (Munro, 2008). The question is how we know which aspect has considerable potential in contributing to L2 pronunciation. Since it has been reported that features of first language (L1) influence L2

speech (Archibald, 2017), we can start by examining phonological differences between L1 and L2 to estimate what linguistic aspect brings the most difficulty to L2 users.

For Korean speakers of English, English stress can be very difficult to learn because unlike English, Korean does not have stress at the lexical level. Contrary to English, a stress-timed language, Korean is a syllable-timed language with a narrower pitch range, which makes Koreans' L2 English sound monotonous or exhibit misplaced lexical stresses (Lee, 2001 as cited in Nelson & Kang, 2015). Due to these phonological influences of Korean, particularly in stress and intonation, Korean speakers of English usually have easily recognizable accents whether in their rehearsed or extemporaneous speech (Nelson & Kang, 2015). To tackle difficulties arising from the phonological influences of Korean, developing teaching materials that help L1 Korean speakers notice and practice English lexical stress might show possibility for improving their English pronunciation. Thus, this thesis addresses whether providing lexical stress diacritics in pedagogical materials could help Korean learners of English produce more comprehensible and target-like pronunciation.

## **1.2. Outline**

The thesis is organized into six chapters. Chapter two includes a review of the literature on English stress, the features of Korean prosody, enhanced input, and pronunciation and individual differences as well as the definitions of key terms and the roles of English lexical stress on the part of the listeners. The discussions of theoretical background and literature review in chapter two lead to three research questions that guide the current study. Chapter three describes the research methodology, including participants, instruments, data collection methods, speech sample rating procedures, and analysis of the data. Chapter four presents the results and accompanying discussions in relation to the three research

questions along with a summary of major findings. Chapter five outlines the pedagogical and empirical implications, limitations, and future directions stemming from the current study. Chapter six contains the conclusion of the thesis.

## **Chapter 2. Theoretical Background and Literature Review**

This chapter presents definitions of key terms (2.1.) that are used in the thesis to ensure a common understanding. Section 2.2. reviews previous studies on English stress and its impact on intelligibility, comprehensibility, and accentedness. Section 2.3. elucidates prosodic differences between Korean and English to provide theoretical background and rationale for the study, illustrating why it can be challenging for Korean speakers to acquire English stress. The section also includes previous studies on different stress processing between L1 and L2 English speakers. Section 2.4. reviews studies pertaining to input enhancement. The discussions of input enhancement lead to the necessity of utilizing enhanced input in terms of developing pronunciation teaching materials. Section 2.5. presents the current pronunciation teaching practice in the Korean public education system. Section 2.6. introduces the purpose of the current study, and section 2.7. includes three research questions that guide the study.

### **2.1. Definitions of Key Terms**

#### **2.1.1. Intelligibility, Comprehensibility, and Accentedness**

Intelligibility is defined as “the extent to which a speaker’s message is actually understood by a listener” (Munro & Derwing, 1999). Most often, this is assessed by how accurate listeners’ orthographic transcriptions are (Munro & Derwing, 1999). Comprehensibility refers to listeners’ subjective sense of how easily listeners understand L2 speech, which is usually measured by a 9-point scale (Munro & Derwing, 1999). Accent can be defined as how different an L2 speaker’s pronunciation is from a target variety (Munro &

Derwing, 1995a). Munro and Derwing (1995) showed that it is possible for a speaker to be heavily accented, while still being intelligible.

**Table 1 Results of Possible Intelligibility and Comprehensibility Combinations (Derwing and Munro, 2015)**

<b>Intelligibility</b>	<b>Comprehensibility</b>	<b>Result</b>
High	High	Utterance is fully understood; little effort required
High	Low	Utterance is fully understood; great effort required
Low	Low	Utterance is not (fully) understood; great effort is exerted
Low	High	Probably rare. Utterance is not fully understood; however, the listener has the false impression of having easily determined the speaker's intended meaning

**Table 2 Results of Possible Intelligibility and Accentedness Combinations (Derwing and Munro, 2015)**

<b>Intelligibility</b>	<b>Accentedness</b>	<b>Result</b>
High	High	Utterance is fully understood; accent is very strong
High	Low	Utterance is fully understood; accent is barely noticeable
Low	Low	Not relevant to pronunciation
Low	High	Utterance is not (fully) understood; accent is very strong

Table 1 and Table 2 (Derwing & Munro, 2015) show the partial independence between accentedness and intelligibility and the relationships between intelligibility and comprehensibility. From the two tables, fulfilling intelligibility seems like a pre-requisite for a condition where interlocutors can interact appropriately with their utterances being fully understood. Does this mean that L2 teachers and learners should not be keen on making L2 pronunciation more comprehensible to L1 speakers? It seems that the tendency focusing on intelligibility in L2 pronunciation teaching and learning has been predominant over comprehensibility and accentedness ever since Abercrombie (1949) first put forward that L2 learners just need “a comfortably intelligible pronunciation.” But it should be noted that Abercrombie (1949) defined “*comfortably intelligible pronunciation*” as “pronunciation which can be understood with little or no conscious effort on the part of the listener.” Table 1 confirms that an utterance with low comprehensibility can be fully understood on the condition that intelligibility remains high. However, it should be noted that in this case understanding an utterance can only be accomplished when a great deal of effort is exerted by a listener. This implies that if the listener is not willing to make an effort to understand the utterance, in other words when cooperation is not made on the part of the listener, the communication is likely to break down. Lev-Ari and Keysa (2010) added a good reason why comprehensibility should be paid attention as a measure of understanding L2 speech. In their study, native English speakers were invited to judge 45 trivia sentences, such as “*A giraffe can go without water longer than a camel can*”, recorded by three native English speakers and six L2 speakers of English. The native listeners were asked to indicate its truthfulness on a 14cm line, with one end labeled *definitely false* and the other end *definitely true*. The results showed that the native listeners tended to perceive trivia sentences as less truthful

when spoken by L2 speakers. Lev-Ari and Keysa (2010) attributed the results to listeners' decreased processing fluency, caused by the fact that accented L2 speech is harder to process (Munro & Derwing, 1995b). Processing fluency denotes people's subjective experience of ease or difficulty of processing stimuli when dealing with cognitive tasks (Oppenheimer, 2008 as cited in Trofimovich & Isaacs, 2012). Information that is easier to process is perceived as more familiar, truthful, and pleasant than is more difficulty to process (Reber & Schwarz, 1999; Whitlesea, 1993 as cited in Trofimovich & Isaacs, 2012). Therefore, it is important to identify which aspects of L2 phonology are responsible for making L2 speech less comprehensible, thereby undermining processing fluency for the listeners, and which merely gives the perception of accented speech (Trofimovich & Isaacs, 2012).

As for accentedness, Table 2 shows that an accented speech can be fully understood if intelligibility is kept high. However, speakers with a foreign accent often face linguistic discrimination and social stigma, which is corroborated by Gluszek and Dovidio (2010)'s review on social psychology research investigating L1 speakers' attitudes towards L2 accented speech. Psychological studies on how listeners perceive accented speakers have generally involved using a "matched-guise" technique, where a highly proficient compound bilingual speaker read a script with and without an L2 accent (Gluszek & Dovidio, 2010). The results of these studies reported that listeners judged a speaker with L2 accent to be less intelligent, less competent, and linguistically less proficient than the same speaker without an accent. Non-comprehensible accented speech may bias native speaker (NS) listeners, and they may exert less effort in understanding the non-native speaker (Lindemann, 2002), causing communication problems (Munro & Derwing, 1995a). L2 accented employees encounter discrimination or bias in various lines of employment (Lippi-Green, 2012, as cited

in Derwing & Munro, 2015). For example, Wang, Ardnt, Singh, Biernat, and Liu (2013) investigated customer biases regarding employee L2 accent in call service encounters. One hundred twenty-two American undergraduates were randomly assigned to listen to one of four different recorded phone conversations where a customer calls into a bank to update his address. A native English speaker and an Indian English speaker recorded a favorable (able to update his address) and an unfavorable outcome (unable to update his address) each for the listeners. The listeners were told that the researchers were interested in improving employee service over the phone and were asked to rate the call center employees' accents on a 7-point scale. The findings of the study indicated that a negative accent bias occurred when the service outcomes were unfavorable. In contrast, when the service outcomes were favorable, accent had no impact on the employee evaluation. The researchers suggested that assigning tasks that are likely to bring favorable outcomes to employees with nonstandard accent be more appropriate. However, it is inevitable that workers in customer-service industry deliver unpleasant messages to their customers and deal with a wide range of complaints. It is possible that the findings and suggestion of the study might aggravate discrimination and bias stemming from L2 accented speech by encouraging employers to hire L1 applicants over L2 applicants. In a similar vein, Timming (2017) examined the negative effect of foreign accent on job applicants' employability ratings. Simulated telephone interviews were recorded by five different accents (American-, Chinese-, Indian-, Mexican-accented and British English) in this experiment. In total, 108 men and 115 women with managerial experience rated how likely they would be to hire each applicant for a customer-facing job or non-customer job on a scale of 1 to 7. The results demonstrated clear evidence of accent-based discrimination with Chinese, Indian, and Mexican accents all rated

lower than North American and British accents. It was also found that voices with these three accents were rated significantly lower in customer facing jobs than non-customer-facing jobs. In other words, job applicants with accents are likely to be discriminated by hiring managers in the interactive services that require active engagement with customers.

Consensus among teachers and researchers about the centrality of intelligibility is growing in L2 pronunciation teaching and learning (Munro & Derwing, 2015). However, after having discussed the points above, we cannot completely disregard the importance of assisting L2 learners to have better comprehensibility with less accent. However, it should be noted that the research direction of the thesis does not involve the elimination of a foreign accent. Many researchers have put forward that attaining native-like pronunciation among those who acquire an L2 after early childhood is not a realistic goal but rather an exceptional case (Flege, Munro, & Mackay, 1995). Leading L2 learners to believe that they will achieve this unrealistic goal might do more harm than good (Derwing & Munro, 2005). Thus, research on pronunciation should be able to assess which factors have the strongest impact on L2ers' pronunciation on the basis of empirical evidence and help material developers and teacher devise teaching tools that could help them address those factors with realistic, attainable goals.

Although previous studies attempted to figure out the role of pronunciation in terms of comprehensibility and accentedness, there is little understanding of what factors in pronunciation contribute to comprehensibility and accentedness (Trofimovich & Isaacs, 2012). It seems that there have been differing results regarding which factors affect listeners' comprehensibility and accentedness in the relatively few studies on this topic. For example, Fayer and Krasinski (1987, as cited in Munro & Derwing, 1999) reported that listeners were

distracted and annoyed by non-native patterns in pronunciation and hesitation. On the other hand, Trofimovich and Isaacs (2012) reported that grammatical and lexical errors contributed more to comprehensibility problems than errors in pronunciation when extemporaneous speech samples from 40 native French speakers of English were examined by 60 native speakers of English. Trofimovich and Isaacs (2012) also found that segmental accuracy in L2 speech appeared to affect L1 listeners' accent judgement. However, it is worth noting that Munro and Derwing (1999) argued that differences in methodology of studies as well as in target languages are partially responsible for inconsistent results in studies on pronunciation. Munro and Derwing's (1999) argument suggests that replication of pronunciation of studies is welcome and needed with different target languages or participants with different L1 backgrounds to provide supporting or contradictory evidence to the body of L2 pronunciation research.

### **2.1.2. Segmentals**

Segmentals are the individual consonants and vowels in a language's phonological inventory (Celce-Murcia, M. Brinton, & M. Goodwin, 1996). Using one sound rather than another can entirely change the meaning of words as in the case of minimal pairs such as *alive* [əlaɪv] and *arrive* [əraɪv], which could create communication problems between interlocutors. Thus, inaccurate pronunciation in individual vowels or consonants can lead to misunderstanding (Kelly, 2000). The segmentals are not going to be considered for this current study since the research topic focuses on English lexical stress, which is a suprasegmental feature that is elaborated in the following section.

### 2.1.3. Suprasegmentals (Prosody)

Suprasegmental features, also known as prosody, refer to a speech feature such as stress, intonation, and rhythm that extends over more than an individual segment (Celce-Murcia et al., 1996). Many researchers have maintained that suprasegmental features should be given priority in terms of pronunciation learning and teaching because of their stronger impact on intelligibility and comprehensibility over segmentals (Zielinski, 2015). Chela-Flores (2001) claimed that suprasegmental aspects carry more impact on intelligibility and assist L2 learners with their immediate pronunciation needs. Suprasegmental features play a vital role in conveying essential information about language, and appropriate suprasegmental productions lead to effective communication (Harrison, Kondo, Meng, Tseng, & Viselgia, 2009). Although these features are often regarded as marginal areas of pronunciation teaching and learning, they carry far more importance to successful communication than sometimes recognized, having the greatest impact on listeners' comprehension (Avery & Ehrlich, 1992). Cutler (1998) argued that listeners extensively utilize suprasegmental properties to recognize spoken utterances. Past studies on L2 pronunciation have reported that suprasegmental errors are as much or more responsible for accentedness, comprehensibility, and intelligibility ratings than are segmentals (Munro & Derwing, 1999; Kang, Rubin, & Pickering, 2010). Derwing, Munro, and Wiebe (1998)'s study found that suprasegmental instruction led to better improvement in comprehensibility rating over segmental instruction in extemporaneously produced narratives after a 12-week course of each instruction.

Thomson and Derwing (2015) pointed out that studies on segmentals have been conducted more often than research involving suprasegmentals. However, perceived foreign

accent and intelligibility might be more greatly impacted by improvement in prosodic proficiency than correction on segmental errors (Munro & Derwing, 1999; Derwing & Rossiter, 2003; Field, 2005). This implies that further research on prosody is needed to add to the existing body of research on L2 pronunciation teaching and learning.

## **2.2. English Stress**

English is a lexical stress language, where the syllables receive relative salience in any English word with more than one syllable (Cutler, 2015). This section starts with explaining types of English stress (2.2.1.), followed by the role of English stress in intelligibility, comprehensibility, and accentedness (2.2.2.). Studies on Stress Parameter Model (2.2.3.) are discussed to illustrate L2 speakers' varying degree of difficulty in stress processing depending on their L1. Two competing views on teaching English stress (2.2.4.) are outlined. A study on the acquisition of L2 English stress and its implication for the current study are discussed in 2.2.5.

### **2.2.1. Types of English Stress**

English has two levels of stress; one level is lexical stress, where a syllable gets prominence at a word level, and the other is phrasal stress (also known as sentence stress), where a word within a phrase receives prominence (Archibald, 1998). In lexical stress, stressed syllables often have a higher pitch, longer duration, and increased loudness than unstressed ones, and must have a full vowel quality rather than be reduced to a schwa (Setter & Sebina, 2017). As for phrasal stress, a stressed syllable in one word in a sentence receives greater prominence, which is phonetically realized by an increased pitch, duration, and loudness (Archibald, 1998). Extended vowel duration and higher pitch in the stressed

syllable are likely to be the most salient characteristics of English stress on the part of the listener (Celce-Murcia et al., 1996).

Consider the following examples from Archibald (1998) illustrating examples of lexical stress and phrasal stress in English.

(1) Lexical stress: *cínema*, *banána*

(2) Phrasal stress (sentence stress): My brother hates dogs.

In (1) lexical stress, the first syllable of the word *cinema* is the most prominent, while the middle syllable of the word *banana* is the most prominent. In (2) phrasal stress, the final content word *dogs* normally gets the most prominence in a neutral discourse context. In a non-neutral discourse context, words carrying new or contrastive information are stressed while words expressing old or given information are unstressed with lower pitch (Halliday, 1967). For example, the following dialogue from Hahn (2004) shows how the speaker B introduces new information to the discourse by giving prominence to new information (*always*) in his utterance indicated by ´, while the word *ready* that shows old information is unstressed.

A: Are you ready?

B: I'm *á*lways ready.

The following sentence “I prefer *réd* wine to *whíte* wine.” from Hahn (2004) shows an example of contrastive stress where two parallel elements (the types of wine) are received prominence to signal the contrast in the sentence.

Among these two categories of English stress, the current study addresses lexical stress in English to investigate the efficacy of English lexical stress diacritics in L2 speech as measured by comprehensibility and accentedness.

### **2.2.2. English Stress and its Impact on Intelligibility, Comprehensibility, and Accentedness**

Fraser (2001, as cited in Zielinski, 2015) prioritized six pronunciation features, drawing on their relative impact on listeners' comprehension. Fraser listed lexical and sentence stress at the top of the list with segmental features down the list, claiming that ESL speakers with decent segmental distinctions would be still very difficult to understand if they misplaced lexical and sentence stress.

Culter and Norris (1988) claimed that the occurrence of a stressed syllable initiates access to the mental lexicon in speech perception. Thus, the speaker should produce correct (or close to correct) stressing. Otherwise, the listener is highly likely to experience difficulty processing the message.

The study by Trofimovich and Isaacs (2012), discussed in 2.1.1. (Intelligibility, Comprehensibility, and Accentedness), reported an influential contribution of lexical stress errors to both accent and comprehensibility ratings. In their study, 60 novice raters and three experienced teachers rated the oral production of 40 native French speakers of English. The findings suggested that different linguistic factors were linked to the perception of accent and comprehensibility for novice and experienced raters. For novice raters, lexical stress and rhythm were deemed salient factors influencing listeners' perception of both accent and comprehensibility. For the three experienced teachers, segmental and grammatical accuracy were the strongest predictors of accent and comprehensibility, respectively. Since the majority of English-speaking population does not have a language teaching background, I believe the results from 60 novice raters are more meaningful to understand which linguistic factors affect perceived comprehensibility and accentedness of general population.

Gallego (1990) conducted a study, where 31 native English-speaking undergraduate students provided immediate feedback on communication breakdowns while watching videotaped presentations given by three international teaching assistants (ITA) to draw out factors affecting the intelligibility of the ITAs. The findings of the study indicated that 35.8% of all the problems on pronunciation were attributed to nonstandard word stress, thereby undermining comprehensibility.

Hahn (2004) conducted a study of how native English listeners processed an ITA's short lecture. Three versions of the speech with correct phrasal stress placement, incorrect phrasal stress placement, and absence of phrasal stress placement were recorded. 90 native English speakers were randomly assigned to listen to one of the three versions, and they were asked to recall the lecture's content and to measure the speaker's English speaking ability and the delivery of the lecture. The results of the study indicated that the listeners processed the version with correct phrasal stress placement faster and more accurately than they did in incorrect or missing stress conditions. Furthermore, the listeners evaluated the speaker with correct phrasal stress significantly more favorably than when phrasal stress was misplaced or absent.

Field (2005) investigated the role of lexical stress in his study, where native speaker (NS) listeners and non-native speaker (NSS) listeners were asked to transcribe English disyllabic words with their lexical stress and vowel quality manipulated. The findings showed that intelligibility, "as part of the broad comprehensibility construct" (Kang et al., 2010) for both groups of listeners was greatly impaired by stress misplacement, especially when lexical stress was shifted to the right. He provided the example of the possible intelligibility problem from misplaced stress. If the misstressed word *followed* occurs at the

beginning of an utterance, the listeners might mistake the word for *load* or *flowed*, shaping their expectations as to what is likely to follow based on the mistaken interpretation.

A study by Kang (2010) revealed that suprasegmental features such as word stress and rhythm are responsible for up to 50% of the variance in accent rating for 11 ITAs from different L1 backgrounds.

### **2.2.3. Stress Parameter Model (SPM)**

The Stress Parameter Model (SPM) hypothesizes that speakers' difficulty with stress perception would vary depending on the regularity of stress location in their L1 (Peperkamp & Dupoux, 2002). Their conclusions are that speakers whose L1 does not use stress to distinguish between lexical items (non-contrastive stress) have difficulty discriminating words in an L2 that does (contrastive stress). They labeled this phenomenon "stress deafness" and concluded that French speakers whose L1 is non-contrastive stress are deaf to stress contrasts.

Lin, C., Wang, M., Idsardi, W., and Xu, Y. (2014) discussed differences between contrastive-stress languages and non-contrastive-stress languages as follows. Stress is lexically contrastive in languages such as English, Russian, Spanish, and German, which can include minimal pairs of words that differ only in stress placement as in the case of the English words *trusty* [trásti] (trustworthy) and *trustee* [trásti:] (board member of a foundation). Stress placement has relatively more variability in languages with lexically contrastive stress. Unlike French stress placement, which is always word or phrase final, and Finnish with its stress being word-initial, stress placement in English is not fixed on a certain syllable.

Previously conducted empirical studies have continually indicated that speakers whose L1 is non-contrastive have greater difficulty discriminating minimal pairs of non-words that differ only in the location of stress, which is called “stress deafness” (Dupoux, Peperkamp, & Sebastián-Gallés, 2001). “Stress deafness” was observed in L1 French speakers even with advanced proficiency in Spanish (Dupoux, Sebastián-Gallés, Navarrete, & Peperkamp, 2008). Dupoux et al., (2008) examined three groups of L1 French speakers with different Spanish proficiency (beginner, intermediate, and advanced). In a sequence recall test, the subjects were asked to listen to randomized sequences of minimal pairs that differed in stress position (e.g. [númi] [numí] [númi] [numí]) and to recall the correct order of the sequences. The findings from the task showed that regardless of their Spanish proficiency, the late learners of Spanish had difficulty encoding stress contrasts. Also, a lexical decision task involving a list of word and nonword minimal pairs that differed only in the position of stress reported that the participants had difficulty utilizing stress to access the lexicon. Based on the results of the study, Dupoux et al., (2008) concluded that “stress deafness” is better considered “as a lasting processing problem” that is not fixable through extensive exposure to an L2 with contrastive stress.

In line with Dupoux et al.,’s (2001, 2008) studies, Lin et al., (2014) tested stress processing among three language groups: Korean and Mandarin learners of English and English monolinguals. They hypothesized that Mandarin speakers would have an advantage over Korean speakers in English stress processing since Mandarin and English have contrastive stress at the word-level while Korean does not. To support their hypothesis, they conducted a sequence recall task and a lexical decision task, both adapted from Dupoux et al. (2008). The findings exhibited that both Mandarin and English speakers outperformed

Korean speakers in stress processing of real words and nonwords. They concluded that the performance of L2 speakers could be influenced by the characteristics of the stress system in their L1.

In response to “Stress deafness”, Archibald (2018) argued that the concept of “stress deafness” results from researchers equating variability with unpredictability. Stress in English can be placed in any syllable (variability in English stress), but it does not mean that stress placement is unpredictable. He put forward that English stress placement is systematically decided by various factors such as quantity-sensitivity, parts of speech, etc, maintaining that English stress is predictable.

I believe “stress deafness” can be controversial because of its dichotomous nature, classifying L2 speakers into either deaf or non-deaf categories in terms of stress processing while ruling out any possibility of the acquisition of L2 stress by non-contrastive L1 speakers. Instead, it should be understood that speakers might have varying degrees of difficulty in terms of stress processing depending on the characteristics of stress in their L1 (Lin, C et al, 2014), but it does not mean that L2 speakers whose L1 has non-contrastive stress are completely deficient in processing contrastive L2 stress. Lin et al., (2014) reported that the Korean speakers did display a certain degree of accuracy rates in the sequence recall task and the lexical decision task, suggesting that Korean speakers do have stress processing abilities to some extent. The key takeaway from the studies regarding “stress deafness” is that it can be indeed very challenging for L1 Korean speakers of English to process stress in English. Therefore, L1 Korean speakers need special assistance such as enhanced input (2.4.) that might help them notice English stress structure.

#### **2.2.4. Teaching English Stress**

It seems that there has been contradicting views on teaching English stress. Jenkins (2000) argued that comprehensive teaching of word stress is not reasonable, asserting that “word stress rules are so complex as to be unteachable.” She also claimed that word stress seems to be important to native English listeners, but it rarely brings about intelligibility problems in interactions between non-native speakers of English. However, she did acknowledge the importance of the relationship between word stress and aspirated sounds because of the effect of on segments (English aspiration only occurs before a stressed vowel), so misplacement of word stress might result in possible intelligibility problems. Celce-Murcia et al., (1996) put forward that English word stress is not consistent, and learners can be confused by this lack of predictability on stress placement.

In contrast, Dauer (2005) considered word stress teachable with a handful of basic rules applicable to 85% of polysyllabic words in English, saying that “Jenkins’s lack of attention to word stress is hard to understand.” She also believed that students need to be taught word stress because it does not appear in the writing system.

Given the consideration of these two contrasting approaches regarding teaching lexical stress and discussions in 2.2.2. (English Stress and its Impact on Intelligibility, Comprehensibility, and Accentedness), it seems apparent that English lexical stress plays a certain degree of role in communication in English. With the advancement of internet and mobile technologies, English learners can be readily exposed to stress patterns of English through auditory input from movies, TV series, or YouTube videoclips. However, some L2 learners seem to have difficulty turning English stress patterns given in auditory stimuli alone

into working knowledge of their own. They might need something extra that could help them notice English lexical stress given in input.

### **2.2.5. Acquisition of L2 English Stress**

Archibald (1997) investigated the acquisition of English stress by speakers of Chinese and Japanese in his longitudinal study, where participants were tested on both production and perception tasks related to stress assignment. The participants were tested in November and tested again in March of the following year. They were not given any treatment regarding English stress between these two tests. The results indicated that any great change in their stress patterns was not observed and the participants did not seem to process English stress based on grammatical category or syllable weight. He contrasted the results with those of his previous study (Archibald, 1993), where speakers of contrastive-stress languages such as Spanish were able to process metrical structures in English. His studies suggest that speakers of non-contrastive-stress languages such as Korean and Japanese are less likely to naturally acquire English stress like Spanish speakers would do.

### **2.3. Korean Prosody**

This section illustrates phonological characteristics of Korean that might negatively work for Korean speakers in acquiring English stress. The discussion in this section, in conjunction with 2.2.3. (Stress Parameter Model) and 2.2.5. (Acquisition of L2 English Stress), will lead to the necessity of enhanced input (2.4.) for L1 Korean speakers of English.

### 2.3.1. The Accentual Phrase in Korean

English has contrastive stress at the word-level while Korean does not have word-level stress, but, rather, has a prosodic system, which is different from English word stress (Jun 2005). Guion (2005) discussed how Korean prosody differs from English in two main aspects based on Jun (1996, 1998). First, tone patterns are the basic building block of Korean prosody while in English stress accent is the basic building block of prosody. Second, stress accents in English are associated with the lexical word, whereas tone patterns in Korean are associated with Accentual Phrases (AP). Jun (1996) adopted the term AP to denote the intonation contour in Korean delimiting a prosodic element, which is larger than a word level but smaller than intonational phrase. An intonational phrase consists of one or more APs with a final boundary tone. AP might contain several lexical items, and it is not straightforwardly predictable to figure out what constitutes AP from syntactic structure (Jun, 1996). It should be noted that tone patterns do not indicate features found in a tonal language such as Chinese, where saying words that share the same phonemes with different tones can denote different items as in the case of 妈 [mā] (mother) and 马 [mǎ] (horse). Rather, tone patterns in Korean should be understood as High-Low or Low-High patterns observed within an Accentual Phrase.

Jun (1998) reported that standard Seoul Korean has Low-High-Low-High (LHLH) or High-High-Low-High (HHLH) tone patterns. The latter pattern occurs when an AP-initial segment is either aspirated (produced with a little puff of air) or tense (fortis and laryngealized). See Table 3 for minimal contrasts for Korean stops in word-initial position.

**Table 3 Korean Stops in Word-initial Position (adapted from Cho, Jun, and Ladefoged, 2002)**

Lenis (Lax)		Fortis (Tense)		Aspirated	
방[paŋ]	“room”	빵[p*ɑŋ]	“bread”	팡[p <sup>h</sup> ɑŋ]	“bang”
달[tal]	“moon”	딸[t*al]	“daughter”	탈[t <sup>h</sup> al]	“mask”
잘[ʃata]	“to sleep”	짤[ʃ*ata]	“to squeeze”	차다[ʃ <sup>h</sup> ata]	“to kick”
개다[kæta]	“to fold up”	깨다[k*æta]	“to break”	캐다[k <sup>h</sup> æta]	“to dig up”
사다[sata]	“to buy”	싸다[s*ata]	“to wrap”		

\* marks fortis obstruents

These two-tone patterns (LHLH or HHLH) are implemented when an AP has more than four syllables. When AP has fewer than four syllables, two or three surface tonal patterns are realized. When an AP has three syllables, various patterns can occur such as LH (or HH for the HHLH pattern), LLH (or HLH), or LHH (or HHH), and it is not clear what conditions trigger this variation (Jun 1998). In the case of a one or two-syllable accentual phrase, the LHLH pattern is turned into LH while the HHLH pattern is produced as HH. An illustrative example is provided in Table 4.

**Table 4 Korean AP Example Sentence (adapted from Jun, 1996)**

Sentence in English	I hate Younga.		
Syntactic structure	I-TOP / Younga-ACC / to hate-DEC		
Phrasing	{na-nin} {jəŋa-riŋ} {miwəh-e}		
Tone patterns	L H	L H	L H

TOP: Topic marker / ACC: Accusative / DEC: Declarative

In the example, all the tone patterns of AP are realized as LH because each AP has fewer than four syllables and all three AP-initial segments [n, j, m] are neither aspirated nor tense. Since Korean does not have minimal pairs that differ only in the location of stress (Jun, 1995), even if ‘na-nin(LH)’ is spoken in HL tone pattern, this tone change in Korean does not change the meaning. People from Pusan, the second largest city in Korea, say na-nin with

a HL pattern, but it does not confuse non-Pusan people at all. When na-nin is realized in HL pattern by Seoul people, it shows that the speaker has intention to emphasize “who” hates Younga. However, it does not make na-nin mean something else. In other words, there are no minimal pairs in Korean caused by tone change because the Accentual Phrase in Korean is not specific to a lexical item, but it is a property of the phrase.

Guion (2005) tested English stress processing among three groups of participants, a native English group, an early Korean-English bilingual group, and a late Korean-English bilingual group. The early bilinguals had begun learning English at the age of one to six years when they moved to the United States. The late bilinguals’ age of acquisition ranged from 15 to 34 years, and it was determined by their age when they moved to the United States. In experiment 1, the participants were asked to produce disyllabic nonwords in both noun and verb sentence template to investigate the effects of lexical class and four different types of disyllabic structure (CVV.CVCC, CV.CVCC, CV.CVC, and CV.CVVC). In experiment 2, the same participants from experiment 1 were asked to report their preference for initial or final syllable stress on disyllabic nonwords that were aurally presented in noun and verb sentence frame. For example, a pair of sentences “I’d like a [béitist] and “I’d like a [beitíst]” was played. The participants were asked to indicate which one sounded more like a real English sentence to them. The findings of the study exhibited that both early and late Korean-English bilinguals had non-nativelike knowledge in term of the distributional patterns of stress placement. The early bilinguals showed a somewhat reduced knowledge as compared to L1 English speakers, and the late bilinguals exhibited an even greater reduction in knowledge of this distributional property. Guion (2005) concluded that native Korean

speakers' ability to detect lexical stress in English seemed to be negatively affected by the early exposure to this AP system in Korean.

### 2.3.2. Vowel Length Distinction in Korean

As discussed in 2.2. (English Stress), one of the components of English stress is extended duration of vowel sounds. Traditionally, vowel length in Korean was analyzed as contrastive between long and short. (Ladefoged, 1982, as cited in Park, 1994). This vowel length distinction still does exist to present day according to the vowel length distinction rule set by the National Institute of the Korean Language (NIKL). For example, NIKL (2018) indicates that to refer to a *chestnut* in Korean, “밤” [pa:m] should be lengthened while “밤” [pam] should not be prolonged to mean *night*. However, in modern Korean, they are treated as homonyms, and people rely on context to interpret meaning. In other words, modern Seoul Korean does not have a vowel length distinction (Kim & Kim, 1998, as cited by Guion 2005). This tendency can be backed up by Park (1994)'s survey conducted in 1990 on 35 native Korean speakers residing in the United States who received their education in Korea from primary school to post-secondary school. The participants were given a survey with lexical items, each of which consisted of a pair of words identical in orthography but different in meaning. The participants were categorized into four groups by age range and place of birth in Korea. I will introduce the result from the most number of participants, Group A whose 21 subjects were under 35-year-old mainly from Seoul. The findings discovered that participants wrongly marked long vowels as short with 56.8% error rate. Considering this experiment was carried out in 1990 when distinguishing vowel length distinction was still part of school curriculum, which it is not at the moment, it is reasonable to conclude that modern Seoul Korean no longer has a length distinction or tense-lax vowel opposition. The

disappearing vowel length distinction in Korean seems to partially contribute to Korean speakers' lack of high sensitivity to length distinction in English, which is an essential component of the English stress system, causing a certain degree of difficulty in their English language learning.

As discussed throughout this chapter, segmental and suprasegmental factors are intertwined in complex ways to contribute to L1 English speakers' judgements of L2 English speakers' speech in terms of comprehensibility and accentedness. Given the different prosodic features found between Korean and English, it seems apparent that Korean speakers of English have to learn that English has stress accent and stress is a lexical property, which is very different from their L1. Considering the different prosodic systems of the two languages and stress not being displayed in the spelling system, it seems evident that Korean learners of English might need pedagogical materials that help them notice English lexical stress.

#### **2.4. Input Enhancement and English Orthography**

Having discussed 2.2. (English Stress) and 2.3. (Korean Prosody), it seems clear that Korean speakers of English have difficulty processing English stress from their L1 interference. This led me to hypothesize that they might benefit from enhanced input that helps them notice English stress and process input to intake (processed and internalized input).

Input enhancement refers to any type of tailored input to make a certain grammatical or phonological feature of input visibly or acoustically more salient (Smith, 1981). According to Schmidt (2001)'s noticing hypothesis, input must be noticed by L2 learners for input to become intake, which is processed and internalized input. A growing number of

studies have looked into the ways of making input more noticeable to L2 learners by utilizing enhancement methods with typographical cues such as underlining, boldfacing, or color coding (Lee & Huang, 2008). A number of recent studies show that written input helps L2 learners' acquisition of L2 phonology (Hayes-Harb & Hacking, 2015). De Bot (1983) reported that visual feedback employing intonational contours is helpful in teaching L2 intonation, inducing L2 learners to notice the forms in oral L2 input. Escudero, Hayes-Harb, & Mitterer (2008) demonstrated that words' written forms promote L2 learners' better perception of challenging contrasts in the auditory input. In their study, 20 L1 Dutch speakers learned English nonwords containing /ɛ/-/æ/ sounds, which is a highly problematic contrast for Dutch speakers of English. They were put into two groups; one group of subjects learned the words by matching the words' auditory forms to corresponding pictured meanings while the other group were given additional spelled forms of the words. The results indicated that subjects who had access to written forms showed better word recognition while subjects who did not were confused with the contrast in question.

However, there have been some contradictory results utilizing input enhancement for the acquisition of phonological features. Hayes-Harb and Hacking (2015) investigated the influence of lexical stress marks on L1 English speakers' ability to learn Russian lexical stress contrasts. They also investigated whether the efficacy of stress marks differed depending on the presentation of the words in Cyrillic (the Russian Alphabet) or the Latin alphabet. 44 subjects without Russian language learning background and 29 with Russian language learning experience participated in the study. While inexperienced learners were randomly distributed into four learning conditions (Latin script with stress diacritics, Cyrillic script with stress diacritics, Latin script without stress diacritics, and Cyrillic script without

stress diacritics), experienced learners were randomly assigned to only the two Cyrillic conditions, learning twelve Russian nonwords in six lexical stress minimal pairs. After they learned the new nonwords through auditory stimuli while being asked to match each word to its corresponding picture during their learning phase, they took a test where participants were told to match pictures with auditory input. The test results found no beneficial effects of the diacritic stress marks and no difference in performance associated with the Cyrillic vs. Latin letter condition. The Hayes-Harb & Hacking (2015)'s study only investigated participants' lexical stress learning at perception level matching auditory stimuli with pictures in laboratory-based involving non-existing words. Further research investigating the efficacy of stress marks on the production of existing words is needed to add to the results of the previous research.

In orthographies of European languages, the position of the stress is not displayed in the spelling except for Spanish, which indicates where the stress should fall on a syllable in a word with irregular stress placement (van Heuven, 2008). It is mostly because when a context is given, speakers of the languages can recognize the words by identifying the constituent phonemes, which makes indicating where stress should be placed with diacritics unnecessary in many cases (van Heuven, 2008). However, as discussed in 2.2.3. (Stress Parameter Model), Korean speakers of English seem to have varying difficulty acquiring English stress, which led me to believe that they might benefit from enhanced input that helps them notice English stress and process input to intake.

## **2.5. Pronunciation Teaching in Public Education System in Korea**

In Korea, English education officially starts at grade 3, and classroom instruction mainly focuses on improving students' listening and reading skills since students are not

tested on the English-speaking and writing domains in the College Scholastic Ability Test (CSAT). Teaching English pronunciation, which is an essential dimension that helps L2 learners achieve successful oral communication, is not part of pedagogical materials such as textbooks from grade 3 to grade 12 (Figure 1, 2). Only textbooks for high schoolers (Figure 2) show how to pronounce newly introduced words at the bottom of a page in the form of International Phonetic Alphabet (IPA), which most students do not know how to use. The CSAT-dependent lesson plans and textbooks lacking a pronunciation teaching component result in producing high school graduates with relatively low proficiency in the speaking domain compared to the reading and listening domains.

So far, we have seen a few *sansuhwa* paintings and learned how the old painters tried to show the beauty of life in nature. When you see the paintings, try to imagine what life in nature is like. Let the small things tell you their secret stories. Read the paintings with your eyes and with your heart. By doing so, you will truly be able to bring nature into your heart, as the painters did.

.....  
 · instead    · alone    · lonely    · moment    · peaceful  
 · secret    · truly

### Let's Think!

Why does the writer think the man got on the boat?

**Figure 1 Grade 9 English textbook (from NE publishing company)**

**While You Read** Q1  
If you were Ethan, how would you feel?

for the last few seconds. One of the 10 seniors, Ethan, was especially happy. He had never played in any of the games before. Now, Ethan was finally getting the chance to step onto the grass.

**While You Read** Q2  
Why did the player pass the ball to Ethan?

When the rival team dropped the ball, one 15 of our players recovered it and quickly ran down the field with it. Ethan ran right after him to catch up. As our player got closer to the end zone, he saw Ethan behind him on his left. Instead of running straight ahead, the player kindly passed the ball to Ethan so that he could score a touchdown. 20

senior [si:njər]	determined [ditəˈrmind]	victory [viktəri]	sideline [sáidlain]
allow [əlaú]	especially [ispéʃəli]	recover [rikávər]	catch up
instead of	score [skɔ:r]		

20 Lesson 1

**Figure 2 High school English 1 textbook (from NE publishing company)**

However, even though high schoolers do not get tested on their spoken English, they need English speaking ability when they are in college or after graduation from college. According to Educational Testing Service (ETS), roughly 1,700 companies in Korea including the highest-paying companies like Samsung, LG, and Hyundai require job applicants to achieve level 5 to level 7 on the TOEIC (Test of English for International Communication) Speaking Test. The test is designed to assess English-speaking skills in the context of daily life and the global workplace. The test takers answer 11 questions and their responses are measured by segmental (individual sounds) and suprasegmental features (stress and intonation) as well as cohesion and sentence structures in their answers. The test scores are reported on a scale of eight proficiency level (1=lowest, 8=highest). Even for applicants applying for positions that do not require everyday English use for their work,

submitting a result of the score is mandatory as one of their screening procedures. Those companies also ask their current employees to take the TOEIC Speaking Test in order to be considered for promotion. According to the ETS Korea website, 63 universities in Korea use the TOEIC Speaking Test as a language requirement for graduation or for scholarship applications.

As discussed above, having a certain level of English speaking is required for Koreans to land quality jobs or for job advancement. However, the public English education system in Korea does not adequately assist students in improving their English speaking skills due to the negative washback (the effect of testing on teaching and learning) of the CSAT. Negative washback takes place when a test's content or structure is determined by a limited definition of language ability, and thus constrains the teaching and learning context (Talyor, 2005). The negative washback of the CSAT lead school teachers and learners to focus only on what is tested, neglecting the speaking domain of English. This phenomenon seems to be prevalent in many EFL contexts. Sicola and Darcy (2015) claimed that students who have been exposed to an educational system where passing standardized tests is prioritized, and to a classroom context that is highly teacher-centered and conducted in L1 to teach metalinguistic knowledge about English (grammar rules, vocabulary lists, etc.) with little chance to practice pronunciation, later are likely to become English teachers in the same system, repeating what they have learned. As a result, this cycle becomes self-perpetuating. Currently, many job applicants, employees, and college students in Korea rely heavily on English cram schools to get a satisfactory level on the TOEIC Speaking test. In response to the problems posed, the current study was to examine cost-effective and feasible measures that help Korean speakers of English improve their English pronunciation.

## **2.6. Pronunciation and Individual Differences**

In the field of individual variables in L2 pronunciation attainment, researchers have focused on age of acquisition (AOA), length of residence (LOR), extent of L1 and L2 use, language learning context, gender and ethnic identity (Edwards, 2017). I discuss relationships between L2 pronunciation and level of L2 proficiency (2.6.1.) and gender (2.6.2.), which are related to the research questions (2.8.) in the current study.

### **2.6.1. Level of L2 Proficiency**

Darcy, Ewert, and Lidster (2012) point out that there is a paucity of research guiding teachers in determining level-appropriate pronunciation activity. In terms of theorizing instructional differences on the basis of L2 proficiency level, very small number of researchers, such as Gilbert (2001a; b, as cited in Darcy et al., 2012), Jenner (1989, as cited in Darcy et al., 2012), and Murphy (1991, as cited in Darcy et al., 2012) contributed (Darcy et al., 2012). Darcy et al. (2012) note that most pronunciation teaching materials are designed for high-level learners. It appears that L2 learners at each level could take advantage of straightening up specific pronunciation priorities (Gilbert, 2001a; Jenner, 1989; Missaglia, 1999, as cited in Darcy et al., 2012). For example, learning complex sentence stress patterns is not ideal for L2 users who have difficulty generating complex sentences in the first place (Darcy et al., 2012). This means that some pedagogic techniques do not work for beginners and require the techniques to be adapted to accommodate beginners. The discussions also raise a question on whether there are L2 pronunciation teaching tools that can be generally applicable across level of L2 proficiency.

### **2.6.2. Gender**

Past studies reported differing results regarding the influence of gender on L2 pronunciation (some reported a significant effect of gender, while others did not), but females usually got higher ratings than males in the majority of the studies that indicated gender as a significant predictor of perceived foreign accent (Piske, MacKay, & Flege, 2001).

### **2.7. The Present Study**

The main purpose of the current study was to delve into the efficacy of English stress diacritics (enhanced input) for L1 Korean learners of English in terms of comprehensibility and accentedness. The study investigated whether or not the treatment on how to use diacritics displaying stress placement helped the participants in the study produce speech samples with better comprehensibility and accentedness. To determine the effects of treatment, the study used a one group pretest and posttest design. There was a two-day gap between the pretest and the posttest. In addition to the pretest and the posttest, a treatment task was also rated by native listeners to examine whether the presence of diacritics assisted L1 Korean speakers' performance in reading sentences aloud.

To verify if the use of diacritics can be generally applicable to a wide range of L2 English learners, the study explored whether the participants showed different performance depending on their perceived speaking proficiency (beginner, intermediate, and advanced) and gender or not.

## 2.8. Research Questions

In light of the issues and previous studies reviewed above, the following three research questions were formulated.

1. What effects does enhanced input have on L1 Korean speakers' reading English sentences out loud, as measured by comprehensibility and accentedness ratings?
2. Do participants show better performance on the sentence reading task when diacritics are present compared to the reading task without diacritics?
3. Do factors such as English proficiency (beginner, intermediate, and advanced) or gender influence performance on the sentence reading tasks?

## Chapter 3. Methodology

This chapter describes the methods used to gather the quantitative data and is divided into the following six sections. 3.1. outlines the participants' demographic information. 3.2. introduces two instruments that were used to collect data. 3.3. explains in detail how the data for the study were collected. 3.4. describes how the collected speech samples were prepared for native English listeners to judge. 3.5. presents how the collected data were rated by the native English listeners. 3.6. explains what statistical methods were employed to answer the three research questions in the study.

### 3.1. Participants

The study involved 30 Korean speakers of English with less than one year of experience in English speaking countries. The participants ranged in age from 17 to 46 years, with a mean age of 26.43. The participants' length of residence in English speaking countries ranged from 1 to 10 months with a mean of 5.06 months. They had begun learning English in their home country as early as age five and as late as age 12, with a mean starting age of 8.2 years. The majority (73.33%) of the participants had begun learning English at the age of eight when they first received English education at elementary school. Four participants first started learning English upon entering middle school because English education started at Grade 7 at that time.

24 participants were enrolled in English language learning institutes in a western part of Canada and five participants were working holiday visa holders who were not enrolled in language schools at the time of recruitment. The remaining one participant was a local high school student. Each participant reported their own perceived English-speaking proficiency

level from beginner to advanced (11 beginner, 17 intermediate, 2 advanced). Table 5 summarizes participant information.

Participation in the current study was completely voluntary, and all the participants were compensated in cash (CAD \$10) for their involvement.

**Table 5 Participant Characteristics**

Age (Years)	Mean	26.43
	Range	17 – 46
Gender	Male	14
	Female	16
Level of Speaking Proficiency	Beginner	11
	Intermediate	17
	Advanced	2
Length of Residence (Months)	Mean	5.06
	Range	1 – 10
Starting Age for Learning English (Years)	Mean	8.2
	Range	5 – 12
Purpose of Visit to Canada	Study	25
	Work	5

*Note. n=30*

## 3.2. Instruments

### 3.2.1. English Language Learning Background Information

In order to obtain participants' background information including age, gender, and hometown, participants were asked to complete a short questionnaire (Appendix A). To protect the identity of research participants, they were asked to use a pseudonym. The questionnaire also included items regarding the participants' prior English learning experience, such as length of residence in English speaking countries, age of acquisition, self-assessed level of English-speaking proficiency, and years of English instruction.

### 3.2.2. English Sentence Reading Materials

Two sets (Table 6 and Table 7) of 15 English sentences were used to elicit speech samples from participants' reading aloud tasks. The 15 sentences in Table 6 were identical with those in Table 7, but sentences in Table 7 contain diacritics showing where lexical stress is placed. Table 6 was employed for the pretest and the posttest, and Table 7 was used for the treatment task. To minimize the possibility of subjects showing better performance in shorter sentences compared to longer ones, each sentence has 13 syllables with three stress diacritics. I created 15 sentences based on the level of vocabulary set by the Ministry of Education in South Korea (Ministry of Education, 2015).

**Table 6 List of Sentences without Diacritics for Pretest and Posttest**

#	Sentences	# of syllables	# of diacritics
1	I will take the comment as a personal insult.	13	3
2	It will be focusing on the specific protest.	13	3
3	I will take a photograph of it before purchase.	13	3
4	I will make progress in my industrial career.	13	3
5	I have a rehearsal for the national contest.	13	3
6	I have to complain about the horrible machine.	13	3
7	I will watch a beautiful parade this Saturday.	13	3
8	I need to consume fresh produce for my surgery.	13	3
9	I need to announce a recall of the computers.	13	3
10	I have an important project about kangaroos.	13	3
11	I like the biology of marine animals.	13	3
12	All the locals pass the democratic agenda.	13	3
13	I got a refund for the perfume in Korea.	13	3
14	I will get a discount for the Olympic season.	13	3
15	There has been a huge decrease in police arrests here.	13	3

**Table 7 List of Sentences with Diacritics for Treatment task**

#	Sentences	# of syllables	# of diacritics
1	I will take the comment as a personal insult.	13	3
2	It will be focusing on the specific protest.	13	3
3	I will take a photograph of it before purchase.	13	3
4	I will make progress in my industrial career.	13	3
5	I have a rehearsal for the national contest.	13	3
6	I have to complain about the horrible machine.	13	3
7	I will watch a beautiful parade this Saturday.	13	3
8	I need to consume fresh produce for my surgery.	13	3
9	I need to announce a recall of the computers.	13	3
10	I have an important project about kangaroos.	13	3
11	I like the biology of marine animals.	13	3
12	All the locals pass the democratic agenda.	13	3
13	I got a refund for the perfume in Korea.	13	3
14	I will get a discount for the Olympic season.	13	3
15	There has been a huge decrease in police arrests here.	13	3

### 3.3. Data Collection

#### 3.3.1. Participant Recruitment

The Application for Research Ethics Approval for Human Participant Research was approved by the Human Research Ethics Board at the university on October 9, 2018. Upon receiving approval, I first contacted a forum moderator for an online Korean community site to obtain approval for posting a recruitment notice (Appendix B). After obtaining approval from the moderator, the notice was uploaded. 15 participants were recruited from the online community. To recruit the remaining 15 participants, I contacted the director of the English Language Centre (ELC) at the university via email with detailed information on the current study. The director allowed me to talk about my research to prospective participants during

their coffee social in the lobby of the ELC on Fridays. After a couple of visits to the ELC, I was able to recruit the remaining 15 participants.

### **3.3.2. Meeting 1 (Pretest, Treatment Instructions, Treatment Task)**

I met with each participant in person two times, and both meetings were held on the university campus in a quiet room without distraction within the Department of Linguistics. When the departmental room was not available, a group study room in the university library was used. The participants were not recorded in a professionally-setup environment such as a speech research lab with a recording booth because this type of setting might cause anxiety for participants. Instead, the participants were recorded by a high-quality recorder in a quiet setting.

During the first meeting, the participants was given detailed information on the purpose of the study and the time commitment needed for the study. They were also informed that the current study was conducted to fulfill the master's program that the author was taking. To ensure participants' anonymity in the study, each participant was asked to use their pseudonym and was assigned a participant number from 1 to 30. The first meeting was designed for participants to complete five tasks as follows: 1) read and sign the participant consent form, 2) provide English Language Background Questionnaire (Appendix A), 3) take the pretest (reading 15 sentences without diacritics), 4) receive the treatment instructions to learn how to use and apply English stress diacritics, and 5) take the treatment task (reading 15 sentences with diacritics).

Before participants started working on each task, instructions were offered verbally in their native language (Korean) for their better understanding. As for the consent form, a verbal summary of the form was offered to each participant explaining the objective of the

study, anonymity, and withdrawal of participation etc. The participants were given enough time to read and sign two copies of the consent form. One copy of was left with each participant, and the other was taken by the researcher. In the case of the English Language Background Questionnaire (Appendix A), the participants were asked to indicate their self-assessed English-speaking proficiency: beginner, intermediate, or advanced. The average time spent on Meeting 1 for five different tasks was approximately 25 minutes.

### 3.3.2.1. Pretest

The pretest was administered to evaluate the participants' reading performance without the presence of English stress diacritics. The Participants were presented a list of 15 sentences (Table 6) and given enough time to go through the list so as not to make them feel rushed or anxious. To properly control any extraneous variables and elicit speech samples reflecting the participants' current level of English knowledge, they were not allowed to consult any types of materials such as a dictionary or ask questions to the researcher to check how to pronounce the words in the list. After they orally expressed that they were ready, the list was taken away.


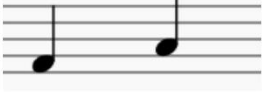

The participants were invited to look at the screen in front of them. In order to maximize the participants' attention to each sentence in Table 6, only one sentence showed up on the screen at a time, and the next sentence came up after the participants finished reading the previous sentence. Their reading aloud tasks were recorded using a high-quality ZOOM H2N recorder in MP3 file format at 44,100Hz sampling frequency and 256 kbps bit-rate. A built-in iPad recorder was used as a secondary recorder just in case of unexpected technical difficulties caused by the primary recorder.

The participants read sentences more than once in one of the following situations.

- 1) The participants expressed their intention of recording again.
- 2) Reading was abruptly interrupted due to the participants' coughing or laughing.
- 3) An unexpected noise occurred outside the recording room.

After successfully being recorded on the pretest, each participant generated 15 separate recording files, with each file containing one sentence shown in Table 6. Then, they moved onto the treatment instructions.

### 3.3.2.2. Treatment Instructions

banána	today	pencil
		
ba      na      na	to      day	pen      cil

**Figure 3 Musical Notation for Treatment Instructions**

In this 10-minute treatment instructions, participants were given a chance to familiarize themselves with the use of English stress diacritics. In the current study, English orthographic input was enhanced (modified) by diacritics to make lexical stress more salient to learners (see 2.4. for more detailed discussion of enhanced input). The participants were informed of two things about the use of enhanced input (diacritics). First, the diacritic means where English lexical stress is placed, and second, the use of the diacritic entails higher pitch and longer duration when diacritic-indicated syllables are pronounced. The following shows the detailed instructions given to the participants.

Each participant was presented with three slides containing musical notation (Figure 3) and instructed to pronounce a syllable with diacritic with slightly higher pitch

and longer duration than the rest. The participants were informed that the use of musical notes in Figure 3 was not meant for them to “sing in exact tune” like fa-la-fa as in banana. The musical notation was employed to visually provide the participants with better understanding on how to give relatively higher pitch on stressed syllables than unstressed ones. I explained the concept of pitch as the relative highness or lowness of the voice. Then, I demonstrated the pitch changes in the slides and asked participants to repeat. The participants were not given speech samples that were recorded by native English speakers.

As discussed in 2.2.1 (Types of English Stress), pitch changes, length (duration), loudness (intensity), and vowel quality are involved to give a syllable prominence. However, due to humans’ limited capacity for information processing, it is extremely difficult for language users to pay equal attention to all linguistic features at the same time. (Segalowitz & Hulstijn, 2005). The Trade-off Hypothesis (Skehan, 2009) posits that because of these capacity limitations, learners need to trade off their attentional resources between all the processes a task requires. They cannot pay equal attention to all aspects of language performance (accuracy, fluency, and complexity). Skehan (2009:511) argued that “committing attention to one area, other things being equal, might cause lower performance in others.” In the same vein, for English learners whose L1 does not have lexical stress system, it could be cognitively demanding to consider all the aspects of English stress. Asking them to consider pitch changes, duration, loudness, and vowel quality might increase task complexity or accuracy at the expense of fluency. As Celce-Murcia et al. (1996) noted, extended vowel duration and higher pitch in the stressed syllable are the most salient characteristics of English stress on the part of the listener, only these two features

were adopted for participants to give prominence to a stressed syllable, while lowering task complexity for them.

After the brief instruction on the pitch change and extended duration given to stressed syllables, the participants were asked to read aloud five sentences (Table 8) that are not relevant in the study. The purpose of reading five novel sentences was not to evaluate the participants' reading performance, but to merely confirm their understanding on the use of English lexical stress diacritics. The participants were asked to read again when they misplaced lexical stress in the practice sentences. The average time spent on the treatment instructions was about 10 minutes.

**Table 8 Practice Sentences**

<b>Practice sentences</b>
1. I don't like going to the festival.
2. You need a balanced meal.
3. I love the University of Victoria.
4. I don't like his series of dramatic actions.
5. Can you describe her appearance?

### 3.3.2.3. Treatment Task

The treatment task was conducted to see if there was any difference in participants' performance in reading aloud the sentences when lexical stress was kept correct with the aid of English stress diacritics.

In the treatment task, the participants were given a list of 15 sentences in Table 7. All the sentences were exactly the same as those used in the pretest except the existence of

diacritics showing stress placement. The participants were presented the list (Table 7) with diacritics and given enough time to go through the list by themselves. They were instructed to pronounce syllables with diacritics with higher pitch and longer duration as covered in the treatment instructions. After they orally expressed that they were ready, the list was taken away.

When the recording began, the participants read each sentence appearing on the screen. The participants read sentences more than once in one of the following situations.

- 1) The participants misplaced lexical stress.

(In the treatment task, the participants were expected to read the sentences with correct stress placement indicated by diacritics.)

Example: Participant #5 put stress on the second syllable for “protest” in the sentence #1 in her first try. She read aloud the sentence again with correct stress placement.

- 2) The participants expressed their intention of recording again.
- 3) Reading was abruptly interrupted due to participants’ coughing or laughing.
- 4) An unexpected noise occurred outside the recording room.

After successfully being recorded on the treatment task, each participant generated 15 separate recording files, with each file containing one sentence shown in Table 7.

At the end of Meeting 1, the participants were reminded of their schedule for Meeting 2. Some of them asked what they would do in the second meeting, but they were not informed on the nature of Meeting 2 so that any possible variables that might have affected the results of the study were excluded.

### 3.3.3. Meeting 2 (Posttest)

Meeting 2 occurred two days after the initial meeting. For example, Meeting 1 for participant #1 took place on October 25, 2018 and Meeting 2 was held on October 27, 2018. In Meeting 2, the participants took the posttest, which was designed to see how participants performed in the sentence reading task when the help of English stress diacritics was removed after two days the treatment had been given. This posttest was administered to determine whether there were the effects of the treatment. All the participants from Meeting 1 returned for Meeting 2.

The participants were given a list of 15 sentences (Table 6) and given enough time to go through the list. Many of them expressed their frustration from not remembering stress patterns in the list. Some of the participants asked me where the stress was placed on certain words, but I did not give them answers. The list was taken away when recording began. The participants read each sentence aloud on the screen as they did in the pretest and the treatment task.

Participants read sentences again in one of the following situations.

- 1) The participants expressed their intention of recording again.
- 2) Reading was abruptly interrupted due to participants' coughing or laughing.
- 3) An unexpected noise occurred outside the recording room.

After successfully being recorded on the posttest, each participant generated 15 separate recording files, with each file containing one sentence shown in Table 6.

### **3.4. Data Preparation**

#### **3.4.1. Inclusion and Exclusion of Data**

After each participant finished the pretest, the treatment task, and the posttest, I listened to sound files to verify if the quality was good enough to be used for data analysis. When the participants generated multiple recordings corresponding to a sentence due to situations described in 3.3.2.1. (Pretest), 3.3.2.3. (Treatment Task), and 3.3.3. (Posttest), the last sound file for a sentence was selected for data analysis with the remaining files discarded. After this reviewing process, each participant yielded 45 recording files (15 sentences \* 3 (pretest, treatment task, posttest)).

#### **3.4.2. Using Data Collected from Secondary Recorder**

As Participant #20's speech samples for the pretest recorded by the primary recorder were completely lost due to unknown technical difficulties, recording files from the secondary recorder (a built-in iPad recorder) were used instead. The volume of speech samples from the secondary recorder was 79.0 dB on average while it was 90.0 dB on average from the primary recorder. MP3Gain, an audio normalization software tool, was used to adjust the different volumes of speech samples collected from two different recorders.

#### **3.4.3. Randomizing Recording Files**

1,350 recording files (15 sentences \* 3 (pretest, treatment task, posttest) \* 30 participants) were left after redundant speech samples were discarded. GoldWave, a digital audio editing software product, was utilized to incorporate 45 sentences (15 sentences \* 3 (pretest, treatment task, posttest)) of each participant into a single file with a randomized

order. Randomization of speech samples was adopted to prevent raters from easily detecting the source of speech samples. 30 separate files (30 participants) were generated through this process, and an average playing time of each file was about four minutes.

**Table 9 Order of Speech Samples**

Sentence #	Order of Presentation to Raters		
	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>
1	Pretest	Treatment task	Posttest
2	Posttest	Treatment task	Pretest
3	Posttest	Treatment task	Pretest
4	Pretest	Treatment task	Posttest
5	Treatment task	Pretest	Posttest
6	Posttest	Pretest	Treatment task
7	Treatment task	Pretest	Posttest
8	Treatment task	Posttest	Pretest
9	Pretest	Treatment task	Posttest
10	Treatment task	Posttest	Pretest
11	Treatment task	Pretest	Posttest
12	Posttest	Treatment task	Pretest
13	Posttest	Pretest	Treatment task
14	Treatment task	Pretest	Posttest
15	Pretest	Treatment task	Posttest

### 3.5. Rating Procedures

I met with three raters two times (one time for comprehensibility rating and the other for accentedness rating) in a speech research lab. The second meeting was held a day after the first rating was completed. Each meeting was comprised of two parts: a rater training and an actual rating. Each meeting lasted approximately four hours including a 7-minute break every five sets of stimuli.

### **3.5.1. Raters**

Three native English speakers, born and raised in North America, were hired to judge the speech samples. All of them held a bachelor's degree in linguistics and were enrolled in a graduate program at a university in Canada at the time of data analysis. None reported being diagnosed with any hearing disorders. All of them were female with a mean age of 28. Each rater was compensated for their work in the form of cash worth 100 dollars (CAD) after the completion of two rating procedures (comprehensibility rating and accentedness rating).

### **3.5.2. Meeting 1 (Comprehensibility Rating)**

#### **3.5.2.1. Rater Training**

The raters were provided with a 15-minute training session on how to grade participants' recording samples. The training session started with presenting the definitions of intelligibility and comprehensibility in the L2 pronunciation context as follows.

##### **1) Intelligibility**

- Definition: How much listeners can understand a speaker's message.
- It is usually measured by assessing how accurate listeners' orthographic transcriptions are.

##### **2) Comprehensibility**

- Definition: How easily listeners understand L2 speech
- High comprehensibility: Listeners can quickly process speech without extra effort, attention, or energy.
- Low comprehensibility: A speech sample requires extra attention, effort, or energy for listeners to understand.

Even though intelligibility was not part of the rating constructs in the current study, its definition was introduced so that the raters could have a clearer understanding of what they were looking at. To make sure of the raters' understanding of the notion of comprehensibility, each rater was asked to state the definition in their own words. As their answers were deemed acceptable, they moved on to a practice rating session.

The practice session as well as the actual rating was conducted on the Google Surveys system, where the raters assigned a perceived comprehensibility rating by clicking a dot below a number from 1 (low comprehensibility) to 9 (high comprehensibility) as shown in Figure 4. This 9-point scale was adopted from Munro and Derwing (1995), but I labeled 1 as low comprehensibility as opposed to high comprehensibility as in Munro and Derwing (1995). The raters received a weblink in their email message connecting to the Google Surveys system for the practice rating.

1-1

	1	2	3	4	5	6	7	8	9	
Impossible to understand	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Extremely easy to understand

1-2

	1	2	3	4	5	6	7	8	9	
Impossible to understand	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Extremely easy to understand

1-3

	1	2	3	4	5	6	7	8	9	
Impossible to understand	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Extremely easy to understand

**Figure 4 Practice Rating Sheet for Comprehensibility**

The raters were told that they were going to be listening to a set of three identical sentences collected from three different times (Table 10). The raters were aware that the order of the sentences was randomized. These sentences were completely unrelated from 15 sentences that the participants read during the data collection procedure. The stimuli were recorded by the author with varied lexical stress placements.

**Table 10 Sentences from the practice session**

- |   |
|---|
| <ol style="list-style-type: none"><li>1. You will describe the picture on your computer.</li><li>2. Soccer is the most popular sport for Canadians.</li><li>3. I have created things to appeal to all the students.</li></ol> |
|---|

The raters were instructed to keep the following in mind when rating the stimuli.

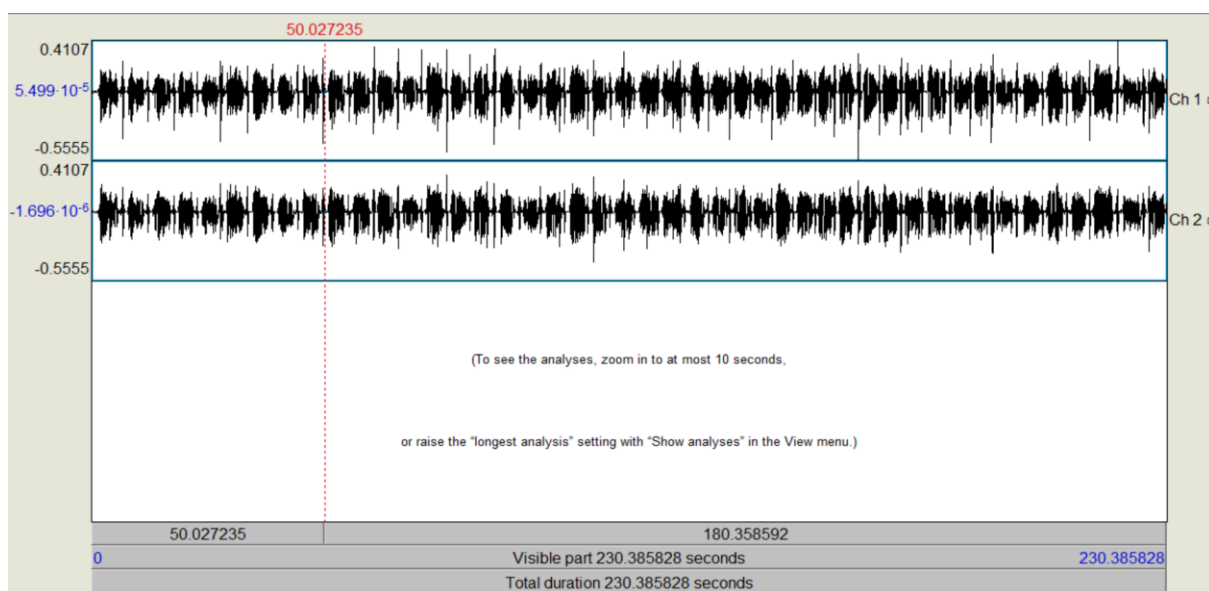
1. Rate each speech sample in relation to one another.
2. Pretend you are hearing speech samples of each speaker for the first time.
3. If possible, use a full scale.
4. Slow rate of speech and naturalness should not affect your judgement. Save them for accentedness rating.
5. Make your judgement after each file is fully heard.

The stimuli were played by Praat (Boersma & Weenink, 2019), free software for the analysis of speech in phonetics, through high-quality stereo speakers in the lab. Speech samples were replayed only when the raters asked for them, or technical difficulties arose. After each set of three sentences was judged, raters shared how they graded each set. This debriefing only happened during the training session, not during the actual rating session. Upon completion of rating three sets of stimuli, the raters moved on to the actual rating session.

### 3.5.2.2. Comprehensibility Rating

The basic procedure for the actual comprehensibility rating was the same as the practice rating. The stimuli were played through Praat (Boersma & Weenink, 2019) starting from participant #1 to #30 in ascending order. Flat areas in Praat window show silences

between speech samples (Figure 5), and they were useful indicators showing where a speech sample started and ended when I needed to replay a speech sample upon raters' requests.



**Figure 5 Praat (Boersma & Weenink, 2019) window showing a waveform display for participant #4**

To maximize raters' rating performance and alleviate their fatigue, the raters were given a 7-minute break every five sets of stimuli, which took approximately 25 minutes for them to rate. The rating sheets were downloaded from the Google Survey system during each break and saved on the author's laptop. It took three hours and thirty minutes for the whole actual comprehensibility rating.

### 3.5.3. Meeting 2 (Accentedness Rating)

#### 3.5.3.1. Rater Training

Meeting 2 was held a day after the comprehensibility rating. The raters were presented with the definition of accentedness as the extent to which a listener judges L2 speech that differs from the native speaker norm. The raters were reminded of the overall procedure of how the rating should be done. They were informed of the differences

between comprehensibility and accentedness and were reminded that they were not going to judge how much they understood the stimuli.

They were provided with the same stimuli from the practice session of the comprehensibility rating for practice (Table 10), but this time they were asked to rate the degree of foreign accent in each sample. The Google Survey system was employed again with a 9-point scale with 1=very strong accent and 9=no foreign accent (Figure 6).

1-1

	1	2	3	4	5	6	7	8	9	
Very strong accent	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	No foreign accent

1-2

	1	2	3	4	5	6	7	8	9	
Very strong accent	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	No foreign accent

1-3

	1	2	3	4	5	6	7	8	9	
Very strong accent	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	No foreign accent

**Figure 6 Practice Rating Sheet for Accentedness**

After each set of three sentences was judged, the raters shared how they graded each set as they did in the comprehensibility rating. This debriefing only happened during the training session. Upon completion of rating three sets of stimuli, the raters moved on to the actual rating session.

### 3.5.3.2. Accentedness Rating

The basic procedure for the actual accentedness rating was the same as the practice rating. Unlike the comprehensibility rating session, where the stimuli were played in ascending order, this time they were played starting from participant #30 to #1 in descending order. The rest of the procedure was the same as the comprehensibility rating session. It took approximately three hours and thirty minutes for the actual accentedness rating.

## 3.6. Data Analysis

The comprehensibility and accentedness judgements were tabulated (e.g. Table 11) on two separate Excel spreadsheets to gather mean ratings for each speech sample on the pretest, the treatment task, and the posttest. Then, 15 mean ratings per reading aloud task were added to find a total score of each participant (Table 12). The same procedure was repeated for accentedness ratings. The maximum score that a participant could get was 135 (15 sentences \* 9-point rating scale), and the minimum was 15 (15 sentences \* 1).

**Table 11 Sample Data Arrangement (Comprehensibility)**

Rater	Participant	Sentence	Pretest	Treatment task	Posttest
1	1	1	6	5	5
2	1	1	5	7	8
3	1	1	7	7	8
<b>Mean ratings (Sentence 1)</b>			<b>6</b>	<b>6.333333</b>	<b>7</b>
1	1	2	8	6	7
2	1	2	5	6	8
3	1	2	8	8	6
<b>Mean ratings (Sentence 2)</b>			<b>7</b>	<b>6.666667</b>	<b>7</b>

			..		
			..		
			..		
			..		
			..		
1	1	15	7	6	8
2	1	15	7	6	8
3	1	15	9	9	9
<b>Mean ratings (Sentence 15)</b>			<b>7.666667</b>	<b>7</b>	<b>8.333333</b>

**Table 12 Total Scores (Comprehensibility)**

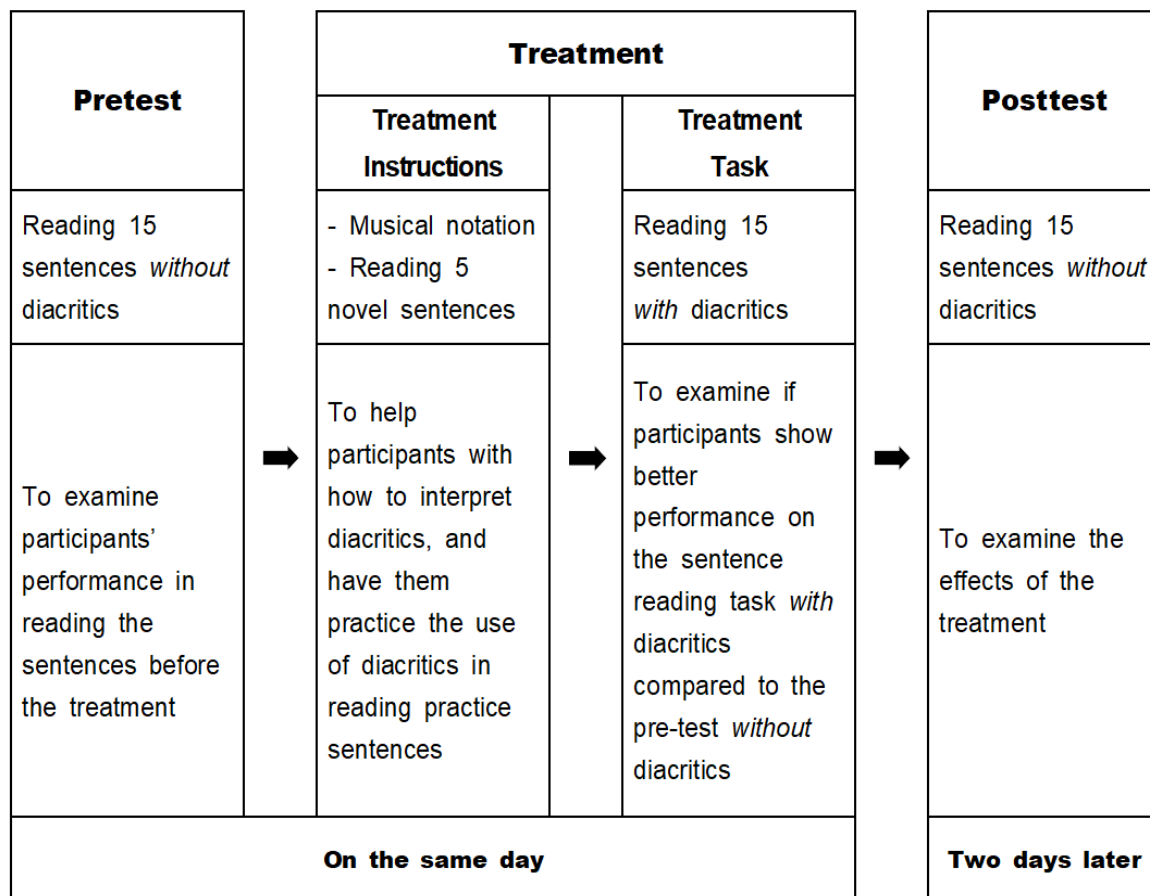
Participant	Pretest	Treatment task	Posttest
1	99.66667	103	106
2	93.66667	101.3333	98.66667
3	91	106	95
4	109.3333	113.3333	119.6667
		..	
		..	
		..	
		..	
30	101.3333	116.6667	108.6667

Since the same participants in the study were measured three times (pretest, treatment task, posttest) to see changes in their sentence reading performance, a one-way repeated-measures analysis of variance (ANOVA) was carried out on Statistical Package for the Social Sciences (SPSS).

## Chapter 4. Results and Discussion

This chapter presents results and accompanying discussion of the present study. Section 4.1. presents the inter-rater reliability among the three native listeners. Section 4.2. includes the results of statistical analysis on comprehensibility scores, and Section 4.3. presents those of statistical analysis on accentedness scores. In both 4.2. and 4.3., the results from the total participants ( $n = 30$ ) are reported first, followed by the results by the participants' level of English proficiency (beginner, intermediate, and advanced) and gender (14 male, 16 female). Section 4.4. discusses the retention of the effects of the treatment. Section 4.5 summarizes the key findings of the study.

The purpose of the present study was to investigate the efficacy of providing lexical stress diacritics on written sentences on the accentedness and comprehensibility of L2 English of L1 Korean speakers. Figure 7 shows the summary of the research design.



**Figure 7 Summary of Research Design**

In order to assess the effects of the treatment on English comprehensibility and accentedness of L1 Korean speakers over three measurement points, a one-way repeated measures ANOVA was performed.

#### **4.1. Inter-rater Reliability**

Intraclass correlation coefficients (ICC) were used to calculate the inter-rater reliability among the three raters on the pretest, the treatment task, and the posttest. A high degree of interrater reliability was found neither in comprehensibility nor in accentedness ratings by using the average measure intraclass coefficient (two-way mixed). The results of the inter-rater reliability are provided in Table 13.

**Table 13 Inter-rater Reliability**

Intraclass correlation coefficients (average measures / two-way mixed)			
	Pretest	Treatment task	Posttest
Accentedness	0.472	0.25	0.362
Comprehensibility	0.356	0.516	0.584

In a high-stakes standardized speaking test such as the TOEIC Speaking Test, having a high degree of agreement between raters is a must because test takers' should get the same results whoever grades their speech samples. To ensure inter-rater reliability, testing organizations train raters in a rigorous way and require them to recalibrate on a regular basis (Derwing & Munro, 2015). Unlike a standardized test, the nature of the current study relied on the raters' subjective judgements in terms of comprehensibility and accentedness ratings. M. Munro (personal communication, March 17, 2019) provided several reasons for seemingly low reliability of the ratings in my study. First, the ICC always increases as the number of raters goes up, so if I had used more raters (15 or more raters), it would have yielded higher ICC. As a by-product, having more raters could also lead to a bigger effect size in the results, since having more raters should result in a more reliable rating of each participant. Second, if the variability in ratings was not large across speakers, in other words scores were clustered instead of being spread on the whole 9-point scale, that might tend to lower the between-rater correlations. Munro suggested including some native speaker voices in the rating set to give the listeners a reference point and asking the listeners to use the whole 9-point scale. Munro does not give raters any training, but provides them with a short no-feedback practice session. In the instructions that I gave

to the three raters (see 3.5. Rating Procedures), I told them to use the whole scale if possible, but there was no native speaker voice to give them a reference point. The points that M. Munro made could be addressed in the future related studies to possibly have a higher degree of inter-rater reliability. As another way to increase inter-rater reliability, having more test items (i.e., sentences) could also be considered.

## 4.2. Comprehensibility

### 4.2.1. Results from Total Participants

Table 14 illustrates that on average the highest score was observed in the treatment task, where the participants performed the sentence reading task with diacritics right after the treatment instructions (mean = 103.966 ; SD = 7.812), followed by the posttest, which took place two days after the training (mean = 103.388; SD = 7.892), and the lowest was gathered from the pretest, where they were assessed before the treatment (mean = 98.888; SD = 6.059).

**Table 14 Means Comprehensibility Scores for Tasks**

TIME	<i>n</i>	Comprehensibility score	
		<i>M</i>	<i>SD</i>
Pretest	30	98.889	6.060
Treatment task	30	103.967	7.812
Posttest	30	103.389	7.893

*Note.* The maximum score is 135.

To find out whether these differences were significant, inferential statistics were examined. First, Mauchly's Test of Sphericity showed that the assumption of sphericity was met,  $\chi^2(2) = 5.253$ ,  $p = .072$ . Since the Mauchley's test was non-significant, I was able

to use the Sphericity Assumed row in Tests of Within-Subjects Effects (see Table 15), which shows that there was a significant effect of the treatment on the comprehensibility scores of the participants ( $F(2, 58) = 12.489, p = .000, \eta_p^2 = .301$ ).

**Table 15 Tests of Within-Subjects Effects (Comprehensibility)**

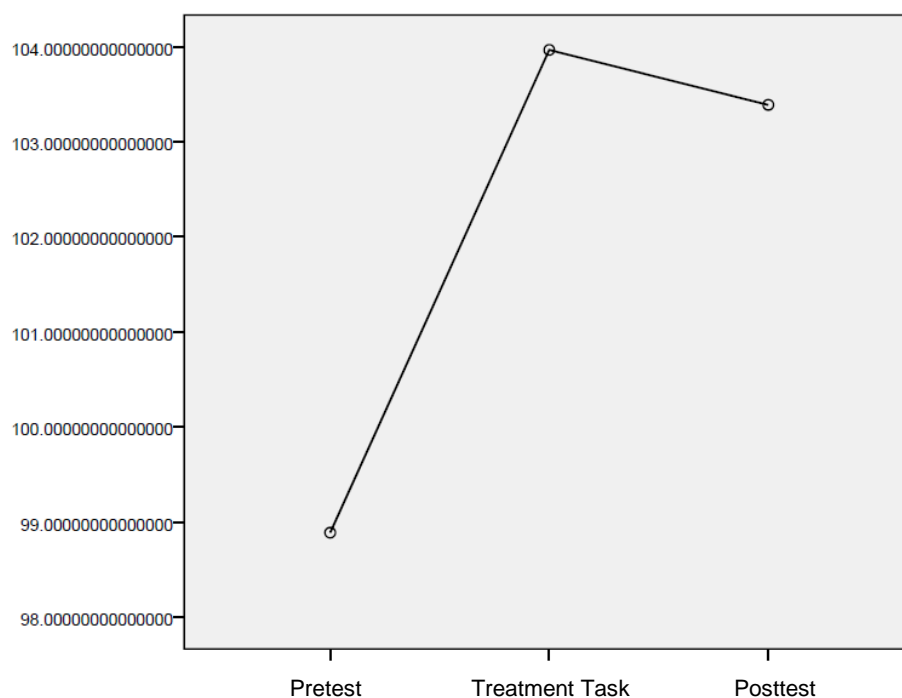
Source		df	F	P	Partial Eta Squared
TIME	Sphericity assumed	2	12.489	.000	.301
Error (TIME)	Sphericity assumed	58			

Table 15 confirmed that an overall significant difference in means was observed. To examine where these differences came from, the Bonferroni post hoc test (Table 16) was carried out. The test revealed that there was a significant difference in comprehensibility between the pretest (no diacritics) and the posttest (no diacritics) ( $p = 0.003$ ), and between the pretest (no diacritics) and the treatment task (with diacritics) ( $p = 0.001$ ), but no significant difference between the treatment task (with diacritics) and the posttest (no diacritics) ( $p = 1.000$ ). The comparison of the pretest and the treatment task was required to answer the second research question “Do participants show better performance on the sentence reading task when diacritics are present compared to the reading task without diacritics?”. Compared to the pretest, the participants’ comprehensibility significantly improved in the treatment task, which suggests that reading sentences with correct stress placement indicated by diacritics led the participants to produce speech with better comprehensibility. Examining a comparison between the pretest and the posttest was critical in this study to verify the effects of the treatment without the aid of diacritics. The comparison of the pretest and the posttest suggests that the

participants successfully applied what they learned from the treatment to the posttest when the diacritics were removed. No significant difference was found between the treatment task and the posttest, meaning that the participants maintained their knowledge gained from the treatment.

**Table 16 Pairwise Comparison (Bonferroni)**

(I) TIME	(J) TIME	Mean Differences (J-I)	P
Pretest	Treatment task	5.078	.001
Pretest	Posttest	4.500	.003
Treatment task	Posttest	-.578	1.000



**Figure 8 Line Graphs for Mean Comprehensibility Scores**

#### 4.2.2. Results by Level of L2 Proficiency

To answer the third research question “Do factors such as English proficiency (beginner, intermediate, and advanced) or gender influence performance on the sentence reading tasks?,” a factorial ANOVA was conducted. The findings demonstrate no statistically significant interaction between level and TIME,  $F(2, 54) = .646, p = .632$  (Table 17), indicating that changes in comprehensibility ratings over time are not significantly different across the three levels (11 beginner, 17 intermediate, 2 advanced) of proficiency.

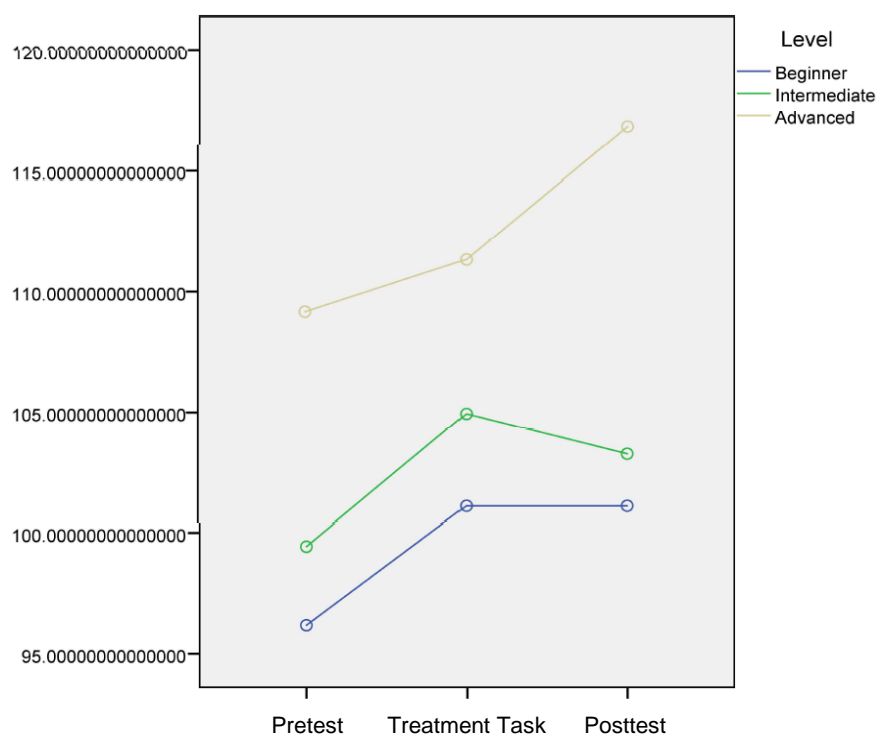
**Table 17 Means by Level of Proficiency for Tasks (Comprehensibility)**

Time	Level	n	M	SD
Pretest	Beginner	11	89.182	5.450
	Intermediate	17	99.431	5.426
	Advanced	2	109.167	.236
	Total	30	98.889	6.060
Treatment task	Beginner	11	101.121	8.040
	Intermediate	17	104.941	7.507
	Advanced	2	111.333	2.828
	Total	30	103.967	7.812
Posttest	Beginner	11	101.121	8.176
	Intermediate	17	103.275	6.631
	Advanced	2	116.833	4.007
	Total	30	103.389	7.782

**Table 18 Tests of Within-Subjects Effects (TIME\*Level: Comprehensibility)**

Source		Df	F	P	Partial Eta Squared
TIME*Level	Sphericity assumed	4	.646	.632	.046
Error (TIME)	Sphericity assumed	54			

The line graphs in Figure 9 show change in comprehensibility ratings over three measurement points. The horizontal x-axis refers to time (pretest, treatment task, posttest), and the vertical y-axis refers to comprehensibility ratings. The three different colors outline the comparison among the participants' English proficiency levels. It is worth noting that advanced learners display somewhat different change compared to beginner and intermediate learners. Advanced learners show better performance in the posttest (without diacritics) than in the treatment task (with diacritics) while it is not the case for beginner and intermediate learners.



**Figure 9 Line Graphs for Mean Comprehensibility Scores by Level of Proficiency**

#### 4.2.3. Results by Gender

The effect of gender on the tasks as in the third research question is reported in the thesis for the following two reasons. First, the results can add to the existing body of research concerning the effect of gender on the degree of L2 accent and comprehensibility.

Second, the findings provide an account of the gender-treatment interaction, showing if the effects of the treatment is generally applicable across gender.

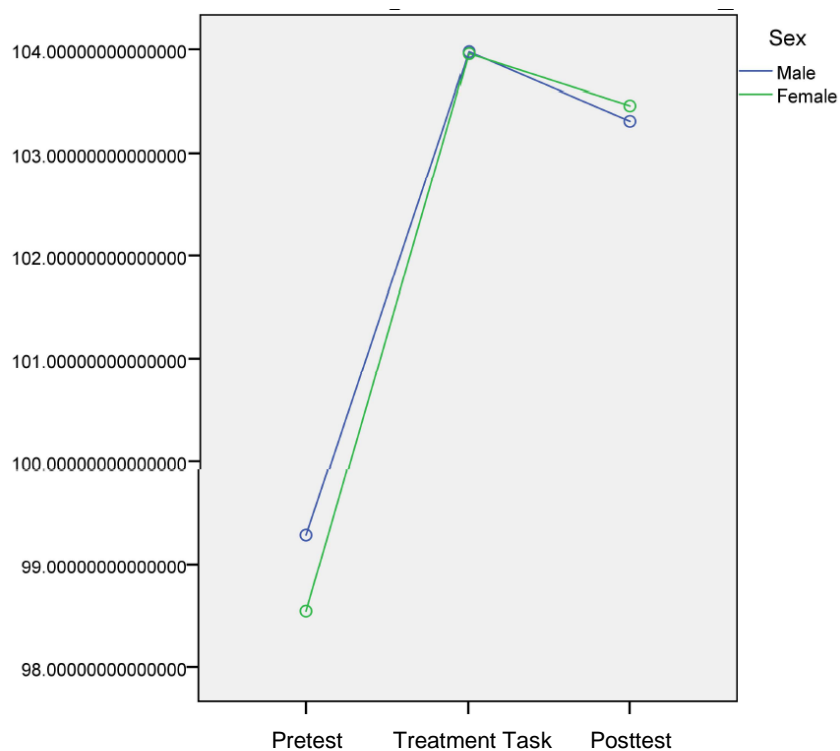
A factorial ANOVA was conducted to discover the interaction effect of the participants' gender on the pretest, the treatment task, and the posttest. There is no statistically significant interaction between gender and TIME,  $F(2, 56) = .088, p = .916$  (Table 19), indicating that changes in comprehensibility ratings over time are statistically equivalent across gender.

**Table 19 Means by Gender for Tasks (Comprehensibility)**

Time	Gender	n	M	SD
Pretest	Male	14	99.286	5.264
	Female	16	98.542	6.834
	Total	30	98.889	6.060
Treatment task	Male	14	103.967	9.492
	Female	16	103.958	6.317
	Total	30	103.967	7.812
Posttest	Male	14	103.310	8.180
	Female	16	103.458	7.901
	Total	30	103.389	7.893

**Table 20 Tests of Within-Subjects Effects (TIME\*Gender\_Comprehensibility)**

Source		df	F	P	Partial Eta Squared
TIME*Gender	Sphericity assumed	2	.088	.916	.003
Error (TIME)	Sphericity assumed	56			



**Figure 10 Line Graphs for Mean Comprehensibility Scores by Gender**

### 4.3. Accentedness

#### 4.3.1. Results from Total Participants

A repeated measures ANOVA (sphericity assumed) determined that mean accentedness ratings significantly differed between the three time points ( $F(2, 58) = 81.516, p = .000, \eta_p^2 = .738$ ). Post hoc tests using the Bonferroni correction (Table 22) revealed that a significant difference ( $p = 0.000$ ) was observed in accentedness between the pretest (no diacritics) and the posttest (no diacritics), and between the pretest (no diacritics) and the treatment task (with diacritics), but no significant difference was found between the treatment task (with diacritics) and the posttest (no diacritics) ( $p = .064$ ).

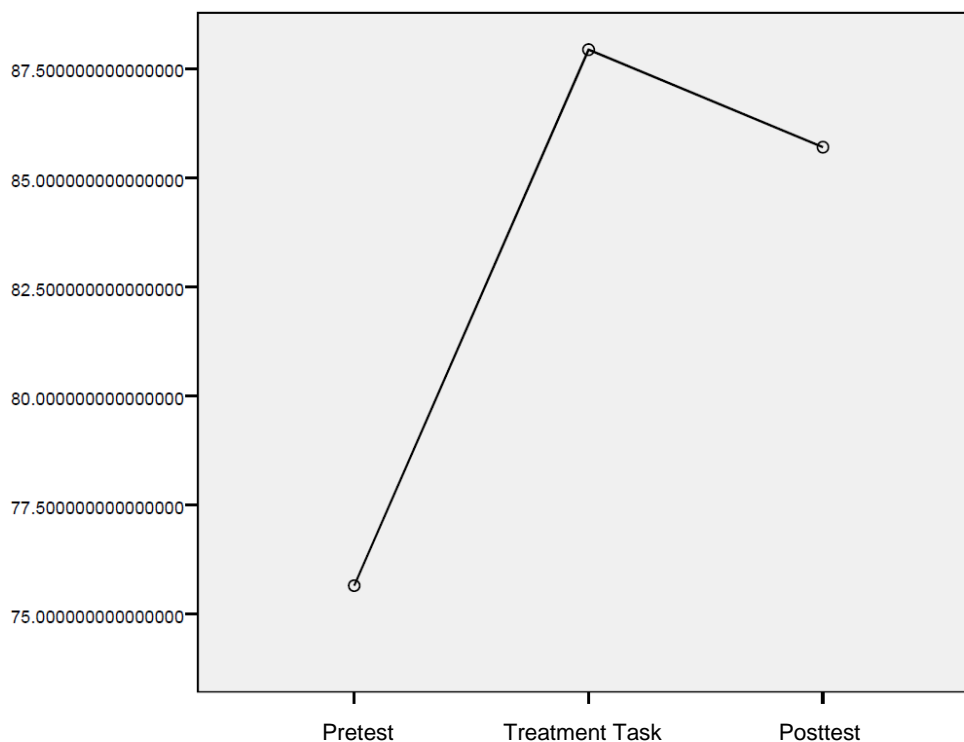
**Table 21 Means Accentedness Scores for Tasks**

TIME	<i>n</i>	Accentedness score	
		<i>M</i>	<i>SD</i>
Pretest	30	75.656	10.310
Treatment task	30	87.933	10.184
Posttest	30	85.700	10.573

*Note.* The maximum score is 135.

**Table 22 Pairwise Comparison (Bonferroni)**

(I) TIME	(J) TIME	Mean Differences (J-I)	p
Pretest	Treatment task	12.278	.000
Pretest	Posttest	10.044	.000
Treatment task	Posttest	-2.233	.064



**Figure 11 Line Graphs for Mean Accentedness Scores**

The participants' accentedness significantly improved in the treatment task, compared to the pretest, which suggests that reading sentences with diacritics present assisted the participants in producing speech samples with better accentedness. The comparison of the pretest and the posttest suggests that the treatment had a significantly positive effect on improving the participants' accentedness in the sentence reading aloud task. No significant difference was seen between the treatment task and the posttest, suggesting that the participants retained the knowledge they gained from the treatment when the assistance was removed.

#### 4.3.2. Results by Level of L2 Proficiency

To answer the third research question "Do factors such as English proficiency (beginner, intermediate, and advanced) or gender influence performance on the sentence

reading tasks?,” a factorial ANOVA was conducted. The findings demonstrate that there is no statistically significant interaction between level and TIME,  $F(4,54) = .579$ ,  $p = .679$  (Table 22), indicating that changes in accentedness ratings over the three measurement points are not significantly different regardless of the participants’ perceived level of L2 proficiency.

**Table 23 Means by Level of Proficiency for Tasks (Accentedness)**

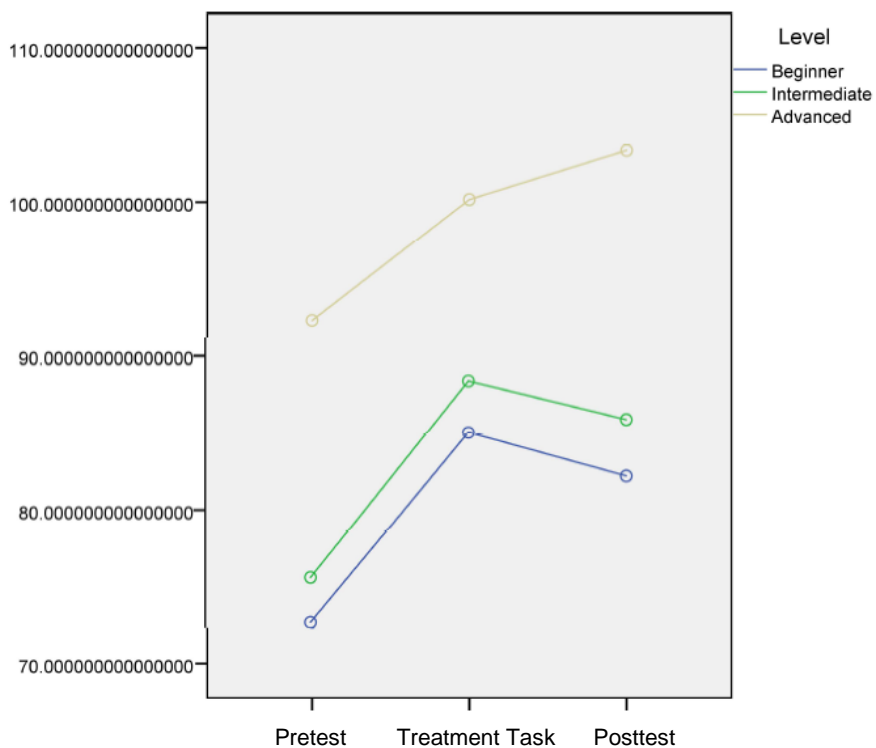
Time	Level	n	M	SD
Pretest	Beginner	11	72.697	8.013
	Intermediate	17	75.609	10.538
	Advanced	2	92.333	3.230
	Total	30	75.656	10.310
Treatment task	Beginner	11	85.030	11.119
	Intermediate	17	88.373	9.243
	Advanced	2	100.167	3.064
	Total	30	87.933	10.184
Posttest	Beginner	11	82.212	9.145
	Intermediate	17	85.882	9.998
	Advanced	2	103.333	7.071
	Total	30	87.700	10.573

**Table 24 Tests of Within-Subjects Effects (TIME\*Level\_Accentedness)**

Source		df	F	P	Partial Eta Squared
TIME*Level	Sphericity assumed	4	.579	.679	.041
Error (TIME)	Sphericity assumed				

The line graphs in Figure 12 display changes in accentedness ratings for three groups of participants over three measurement points. As found in comprehensibility

ratings (Figure 9), the advanced participants' accentedness ratings increased over time while the beginner and intermediate participants' rating slightly went down in the posttest.



**Figure 12 Line Graphs for Mean Accentedness Scores by Level of Proficiency**

#### 4.3.3. Results by Gender

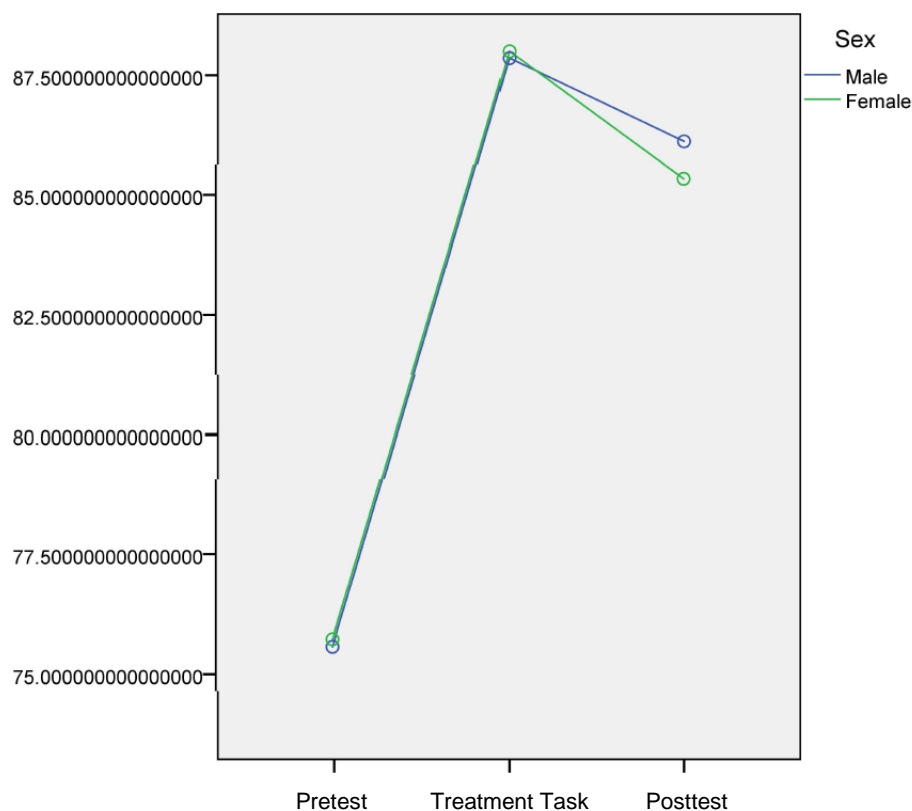
A factorial ANOVA suggests that there is no statistically significant interaction between gender and TIME,  $F(2, 56) = .134$ ,  $p = .875$  (Table 24), indicating that changes in accentedness ratings over the three measurement points are equivalent across gender. These results demonstrate that the effects of the training are not gender-specific.

**Table 25 Means by Gender for Tasks (Accentedness)**

Time	Gender	n	M	SD
Pretest	Male	14	75.571	9.528
	Female	16	75.729	11.260
	Total	30	75.656	10.310
Treatment task	Male	14	87.857	10.340
	Female	16	88.000	10.385
	Total	30	87.933	10.184
Posttest	Male	14	86.119	10.366
	Female	16	85.333	11.075
	Total	30	85.700	10.573

**Table 26 Tests of Within-Subjects Effects (TIME\*Gender\_ Accentedness)**

Source		df	F	P	Partial Eta Squared
TIME*Gender	Sphericity assumed	2	.134	.875	.005
Error (TIME)	Sphericity assumed	56			



**Figure 13 Line Graphs for Mean Accentedness Scores by Gender**

#### **4.4. Retention of the Effects of the Treatment**

Before I started to conduct the study, my hypothesis was that the participants' performance would significantly improve in the treatment task with the help of diacritics. However, significant improvement in the posttest was not expected in my initial hypothesis because the 15-minute treatment (10-minute treatment instructions and 5-minute treatment task) seemed short, and the participants were not given any assignments to reinforce what they learned between the treatment and the posttest, which took two days later. Thomson and Derwing (2015) noted that global improvement in comprehensibility and intelligibility requires longer duration of instruction such as weeks or even months of instruction, not hours or days. Although the results of the current study do not pinpoint specific reasons as

to why the effects of the treatment were maintained during the posttest, I can surmise that the explicit instruction in the treatment might have raised the participants' awareness of English lexical stress, which led to significant improvement in the posttest. This surmise can be backed up by Thomson and Derwing (2015) maintaining that pronunciation will continue to improve after explicit training, particularly where instruction helps raise learners' awareness of pronunciation aspects.

Although I did not conduct a formal interview on how the participants felt about the treatment in this quantitative study, I did have a brief conversation with each participant at the end of the experiment. Some of the participants mentioned that they did not have much awareness about English stress before, and the treatment served as a wake-up call for them to study more about English lexical stress in the future. When presented the compensation, some participants even mentioned that they felt like they should pay for the treatment, rather than receive the money.

The results of the current study suggest that even short-term intervention can lead to significant pronunciation gains. To verify whether there were more permanent effects of the treatment, a delayed post-test (e.g. with a two-week delay) could be adopted in future studies.

#### **4.5. Summary of Results**

Statistical analyses of the quantitative data produced statistically significant differences between the pretest (without diacritics) and the posttest (without diacritics) in accentedness as well as comprehensibility ratings, suggesting the effects of the treatment. Significant improvement was observed between the pretest and the treatment task. The result suggests that the presence of lexical stress diacritics assisted the participants in

producing speech samples with improved comprehensibility and accentedness. The improvement in comprehensibility and accentedness ratings over time was statistically equivalent across the English proficiency level and gender of the participants. A significant difference was observed neither in comprehensibility nor accentedness ratings between the treatment task (with diacritics) and the posttest (without diacritics). These results suggest that the participants successfully applied what they learned from the treatment to the posttest, where the diacritics were not present, with the effects of the treatment maintained.

Table 27 shows the summary of the results. Overall, in terms of the effects of the treatment, the significant improvements were observed in both comprehensibility and accentedness. It is notable that bigger mean differences and p-value (pretest vs. treatment task and treatment task vs. posttest) were found in accentedness than in comprehensibility.

**Table 27 Summary of Results**

(I) TIME	(J) TIME	Category	Mean Differences (J-I)	p
Pretest	Treatment task	Comprehensibility	5.078	0.001
		Accentedness	12.278	.000
Pretest	Posttest	Comprehensibility	4.5	0.003
		Accentedness	10.044	.000
Treatment task	Posttest	Comprehensibility	-0.578	1
		Accentedness	-2.233	0.064

\* pretest: reading-aloud task without diacritics

\* treatment task: reading-aloud task with diacritics

\* posttest: reading-aloud task without diacritics

## **Chapter 5. Implications, Limitations, and Future Directions**

The chapter addresses the implications and limitations of the current study and proposes future research directions drawing on the study's major findings. Empirical implications as well as pedagogical implications are presented, followed by limitations stemming from the research design of the study. Future research directions are suggested based on the limitations of the current study.

### **5.1. Implications**

#### **5.1.1. Pedagogical Implications**

Teachers rely heavily on their intuition rather than research in terms of deciding which features have the greatest effect on pronunciation in class (Levis, 2005). Many language teachers seem to neglect pronunciation due to lack of training, not feeling confident enough in their knowledge in pronunciation (Foote, Holtby, & Derwing, 2011). Furthermore, non-native English teachers find teaching pronunciation a further obstacle to overcome because of their insecurity in being non-native speakers (Levis, Link, & Barriuso, 2016). Their insecurity mostly comes from lacking confidence in their ability to successfully model English pronunciation, and most teachers end up teaching about English (grammar rules etc.) in their L1 (Sicola & Darcy, 2015).

The findings of the present study can provide insight into the way pronunciation instructions can be devised by material developers as well as native and non-native classroom practitioners.

Material developers should consider inserting diacritics to show where stress is placed in a short paragraph or a dialogue in a textbook. However, I do not suggest every

part of a textbook should be full of diacritics as it would be distracting for learners to focus on the content of the materials. Rather, it would be a good idea to have diacritics in a small part per chapter where the focus is on pronunciation. Material developers should make sure that diacritic-embedded parts cover frequently used English words first before less frequently used words. For example, English words *simultaneous* and *status* are both listed in the national curriculum of English in Korea as essential words that students need to learn (Ministry of Education, 2015). However, in terms of prioritizing which word should be first included in the diacritic-mediated pronunciation teaching parts, high frequency word *status* should come first before *simultaneous*. To figure out word frequency, material developers and teachers can refer to the British National Corpus (BNC)<sup>1</sup>, which is composed of real life written and spoken texts. According to the BNC, *status* is seen 9,061 times whereas *simultaneous* is seen 625 times in the corpus. Organizing the sequence of diacritic-embedded parts based on frequency would allow students to learn what they need to know most in real-life communication.

As discussed in 2.2.(English Stress), stress placement in English is not fixed on a certain syllable, meaning that stress can vary across syllable positions within words. In terms of deciding on which stress pattern should be taught first using diacritics, I would recommend teaching words that are stressed on the initial syllable first for the following reasons. Vocabulary studies exhibit that there is a significantly strong tendency that English words get stressed on the initial syllable, and this tendency is even stronger in real speech samples (Culter & Carter, 1987, as cited in Culter, 2015). Although approximately a quarter of the words in English have unstressed initial syllables, most of them have low

---

<sup>1</sup> <https://www.english-corpora.org/bnc/>

frequency of occurrences (e.g. *acquire*) (Culter, 2015). The words most often used in real speech are shorter and more likely to be stressed on the word-initial syllable (e.g. *garbage*, *borrow*, *numbers*), implying that stressed syllable in spoken English is most likely to be the beginning of words (Culter, 2015).

For learners to reinforce stress placement in English, a range of activities can also be included in a textbook. I present two sample activities (paired-reading activity and creative writing) that I designed in Appendix 3.

Teachers can simply make a handout if their current textbooks do not have diacritic-embedded activities. Non-native teachers can refer to a free online dictionary such as The Merriam-Webster Dictionary<sup>2</sup> to check stress that they are unsure of. Teachers can be creative in terms of developing a handout while considering the level of their students and key vocabulary items in a chapter they are teaching. Using diacritics can be incorporated into teaching a grammar point where stress shift brings out a changing part of speech on its own (e.g. *prótest*: noun, *protést*: verb). A sample lesson plan can be found in Appendix 3.

The findings of the study offer an idea of developing a language learning program in the field of computer-assisted language learning (CALL). Developers in the field of CALL might consider devising a computer program or a mobile app that allows users to see English lexical stress when they type or paste paragraphs or sentences. A program should be able to parse the syntactical structure of sentences as English has a number of word pairs that change stress depending on the part of speech (e.g. *prótest*: noun, *protést*:

---

<sup>2</sup> <https://www.merriam-webster.com/>

verb). If the program can be equipped with a text-to-speech technology (TTS), L2 learners will be able to get auditory input as well as visual input from diacritics.

Many English classes in an EFL setting tend to focus on improving learners' reading comprehension with less attention paid to pronunciation (Szpyra-Kozłowska, 2015). Nevertheless, it is important to note that some Korean learners of English are in need of being equipped with a certain level of speaking proficiency after secondary school as discussed in 2.5. (Pronunciation Teaching in Public Education System in Korea). Since the current public education system does not properly accommodate the needs, students are directed to take English speaking classes at private institutes. In this regard, the findings of the current study are significant because it can provide pointers of how reading-focused EFL classes can be organized to support learners' needs of having speaking skills with good comprehensibility. The results of the study suggest that enhanced input can be beneficial for improving comprehensibility as well as accentedness. When the enhanced input is combined with a wide range of reading activities (e.g. reading aloud, creating a story), it is possible that the outcome of classroom instructions could translate into improving learners' pronunciation even in reading-focused EFL situations. Thus, the present study contributes to teachers developing pronunciation teaching tools based on the empirical findings.

### **5.1.2. Empirical Implications**

As discussed in 2.4. (Input Enhancement), various types of enhanced input have proven effective in L2 learners' acquisition of L2 phonology. To my knowledge, however, there has been no empirical study of whether providing lexical stress diacritics affects L2 learners' pronunciation comprehensibility and accentedness. Although Hayes-Harb and

Hacking's (2015) study (see 2.3. for details) investigated the usefulness of the lexical stress marks, the focus of their study was to see the relationship between lexical stress marks and Russian nonword-learning ability, not pronunciation. The present study explored the role of English lexical stress diacritics in relation to L1 Korean speakers' English comprehensibility and accentedness. The participants demonstrated immediate improvement in both constructs (comprehensibility and accentedness) with the presence of diacritics in the treatment task, and the effects of the treatment were maintained in the posttest. Thus, the findings of the current study provide empirical evidence on the role of lexical stress diacritics in perceived comprehensibility and accentedness. The findings of this thesis support positive results from pre-existing studies regarding the usefulness of enhanced input. To my knowledge, the current study was the first to examine a linkage between lexical stress diacritics and L2er's pronunciation. Therefore, the current study can serve as a starting point for further related empirical studies to be pursued.

## **5.2. Limitations and Future Directions**

Although significant effects of the treatment were found in the current study, the study reported here has several limitations in terms of its research design. Discussions of the lack of control group and ecological validity will lead to future research directions.

### **5.2.1. Lack of Control Group and Group Size**

The current study adopted a one-group pretest–posttest design due to limited time and resources allowed for my study timeline. To assess a more accurate measurement of the effects of the treatment, having a control group that did not have any type of interventions would be ideal.

However, prior studies indicated that external factors are less likely to contribute to significant improvement in research with short-term intervention (Thomson & Derwing, 2015). Given the brief duration of instructional input in the current study (approximately 15 minutes, see 3.3.2.2. Treatment Instructions and 3.3.2.3. Treatment Task), I can assume that the significant improvement is attributed to the treatment.

In terms of group size, I had 30 participants in my study. However, the actual group size per level of proficiency was neither high nor equal, particularly in the advanced group (11 beginner, 17 intermediate, 2 advanced). Future studies could have 30 participants in each group to further verify the effects of speaking proficiency on the use of lexical stress diacritics.

### **5.2.2. Ecological Validity**

As for the methods, the sentence reading-aloud task that I used for assessment might act as a limitation to the study's ecological validity. Although the treatment led to the significant improvement in terms of both comprehensibility and accentedness, the current study does not answer whether the improvement can translate into meaningful pronunciation skills in real life communication. However, Thomson and Derwing (2015) maintain that significant improvement in a controlled context such as reading-aloud tasks could lead to long-term enhancements in pronunciation. Their argument suggests that the treatment of the current study might help the participants have pronunciation skills that could be transferred to meaningful communication in real-life situations.

To examine whether the treatment can lead to pronunciation gains in the real world, future studies may adopt a variety of extemporaneous speaking tasks, such as picture-

description tasks, impromptu presentations or topic-based questions (e.g. what is your most memorable trip and why?).

### **5.2.3. Lack of Treatment on Reduced Vowel**

Suprasegmental distinctions (higher pitch and longer duration) indicate stressed syllables in all lexical stress languages (e.g., Spanish), but English also distinguishes stressed and unstressed syllables segmentally; a stressed syllable must contain a full vowel while an unstressed one has a schwa (Cutler, 2015). However, due to brief duration of the treatment in the current study and the cognitive load on the participants (see 3.3.2.1. Pretest), the use of reduced vowel in unstressed vowels was not considered in the study. Since this feature is important to English stress in terms of production and perception of words (Cutler, 2015), it is worth including the use of schwa in treatment in future studies. To alleviate the cognitive load on learners, it is recommended that introducing features contributing to English stress one by one with enough time to practice in a range of longer duration of treatment (e.g. 8 weeks).

### **5.2.4. Future Directions**

Having discussed the implications and limitations of the current study, future research will ideally have a control group, a delayed post-test (see 4.4. Retention of the Effects of the Treatment), and a pretest and posttest that assesses participants' extemporaneous speaking skills. It would be ideal to include a segmental feature (the use of schwa) as well as suprasegmental features (extended duration and higher pitch) in a treatment session. The current study only examined L2ers with non-contrastive-stress L1.

Therefore, it could be worthwhile investigating English speakers of contrastive-stress languages such as German or Spanish as well.

## Chapter 6. Conclusions

The current study investigated the efficacy of lexical stress diacritics on L2 English pronunciation of L1 Korean speakers. Repeated measurement of sentence-reading tasks was administered to collect and analyze speech samples from 30 L1 Korean speakers. The findings revealed statistically significant improvement in both accentedness and comprehensibility ratings when the diacritics were present in the treatment task, and the effects of the treatment were retained in the posttest that took place two days after the treatment. The efficacy of the treatment was not statistically different across English proficiency levels or gender of the participants, showing promise as a more generally applicable pedagogic technique.

Utilizing diacritics to mark lexical stress does not require a cutting-edge technology nor special training. The method can be easily adopted by material developers, language teachers, and even learners for their independent language learning.

The present study was the first to investigate the linkages between the use of lexical stress diacritics and the participants' reading aloud performance measured by comprehensibility and accentedness ratings. In terms of its research design, the study was limited to testing the participants' reading aloud tasks. The results of the study indicate that further related studies on different L2 groups and different tasks eliciting spontaneous speech in an experimental research design are worth pursuing.

## Bibliography

- Abercrombie, D. (1949). Teaching Pronunciation, *ELT Journal*, 3(5), 113–122.
- Archibald, J. (1993). *Language learnability and L2 phonology: The acquisition of metrical parameters*. Boston, MA: Kluwer Academic Publishers.
- Archibald, J. (1997). The acquisition of English stress by speakers of nonaccentual languages: Lexical storage versus computation of stress. *Linguistics*, 35(1), 167-181.
- Archibald, J. (1998). *Second language phonology*. Philadelphia, PA: John Benjamins Publishing Company.
- Archibald, J. (2017). Transfer, contrastive analysis and interlanguage phonology. In: Kang, O., Thomsson, R. I. and Murphy, J. M. (eds.) *The Routledge Handbook of Contemporary English Pronunciation* (pp. 9-24). New York, NY: Routledge.
- Archibald, J. (2018). Advanced level phonology. In P. Malovrh & A. Benati. (eds.) *The Blackwell Handbook of Advanced Proficiency in Second Language Acquisition*. (pp. 241 - 263). Hoboken, NJ: Wiley Blackwell.
- Avery, P., & Ehrlich, S. (1992). *Teaching American English Pronunciation*. Oxford: Oxford University Press.
- Boersma, P. & Weenink, D. (2019). *Praat: Doing phonetics by computer* [Computer program]. Version 6.0.45. Retrieved from <http://www.fon.hum.uva.nl/praat/>
- Celce-Murcia, M., Brinton, D., & Goodwin, J. M., 1953. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. New York, NY: Cambridge University Press.
- Chela-Flores, B. (2001). Pronunciation and language learning: An integrative approach. *International Review of Applied Linguistics in Language Teaching*, 39(2), 85-101.
- Cho, T., Jun, S., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics*, 30(2), 193-228.
- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 14(1), 113-121.
- Cutler, A. (1998). Prosodic structure and word recognition. In A.D. Friederici (Ed.), *Language Comprehension: A Biological Perspective* (2nd ed., pp. 41-70). Berlin: Springer.

- Cutler, A. (2015). Lexical stress in English pronunciation. In: Reed, M., & Levis, J. (eds.) *The handbook of English pronunciation* (pp. 106-124). Malden, MA: Wiley Blackwell.
- Darcy, I.; Ewert, D; & Lidster, R. (2012). Bringing pronunciation instruction back into the classroom: An ESL teachers' pronunciation "toolbox". In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching Conference*, Sept. 2011. (pp. 93-108). Ames, IA: Iowa State University.
- Dauer, R. M. (2005). The lingua franca core: A new model for pronunciation instruction? *TESOL Quarterly*, 39(3), 543-550.
- De Bot, K. (1983). Visual feedback of intonation I: Effectiveness and induced practice behavior. *Language and Speech*, 26(4), 331-350.
- Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48(3), 393-410.
- Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39(3), 379-398.
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Philadelphia, PA: John Benjamins Publishing Company.
- Derwing, T. M., & Rossiter, M. J. (2003). The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2 accented speech. *Applied Language Learning*, 13(1), 1-17.
- Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress "deafness". *Journal of Acoustical Society of America*, 110, 1606-1618.
- Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S. (2008). Persistent stress 'deafness': The case of French learners of Spanish. *Cognition*, 106(2), 682-706.
- Edwards, J. G. H. (2017) Pronunciation and individual differences. In: Kang, O., Thomsson, R. I. and Murphy, J. M. (eds.) *The Routledge Handbook of Contemporary English Pronunciation* (pp. 385-398). New York, NY: Routledge.
- Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, 36(2), 345-360.
- ETS TOEIC [Web page]. (n.d.). Retrieved from <http://exam.ybmnet.co.kr/toeicswt/company/company.asp>

- Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly*, 39(3), 399-424.
- Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, 97, 3125–3134.
- Foote, J., Holtby, A., & Derwing, T. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada Journal*, 29(1), 1–22.
- Gallego, J. C. (1990). The intelligibility of three nonnative English-speaking teaching assistants: An analysis of student-reported communication breakdowns. *Issues in Applied Linguistics*, 1(2), 219-237.
- Gluszek, A., & Dovidio, J. F. (2010). The way they speak: A social psychological perspective on the stigma of nonnative accents in communication. *Personality and Social Psychology Review*, 14(2), 214-237.
- Guion, S. (2005) Knowledge of English word stress patterns in early and late Korean-English bilinguals. *Studies in Second Language Acquisition* 27(4), 503-533.
- Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, 38(2), 201-223.
- Halliday, M. A. K. 1925. (1967). *Intonation and grammar in British English*. Paris: Mouton.
- Harrison, A., Kondo, M., Meng, H., Tseng, C., & Viselgia, T. (2009). Studying L2 suprasegmental features in Asian Englishes: A position paper: *Interspeech Brighton 2009*, 6(10), 1715-1718
- Hayes-Harb, R., & Hacking, J. (2015). The influence of written stress marks on native English speakers' acquisition of Russian lexical stress contrasts. *Slavic and East European Journal*, 59(1), 91-109.
- van Heuven, V. J. (2008). Making sense of strange sounds: (mutual) intelligibility of related language varieties. A review. *International Journal of Humanities and Arts Computing*, 2(1-2), 39-62.
- Jenkins, J. (2000). *The Phonology of English as International Language*. Oxford: Oxford University Press.
- Jun, S. (1995). A phonetic study of stress in Korean. *The Journal of the Acoustical Society of America*, 98(5), 2893-2893.
- Jun, S. (1996). *The phonetics and phonology of Korean prosody: Intonational phonology*

*and prosodic structure*. New York, NY: Garland Pub.

- Jun, S. (1998). The accentual phrase in the Korean prosodic hierarchy. *Phonology*, 15(2), 189-226.
- Jun, S. (2005). Korean intonational phonology and prosodic transcription. In S.-A. Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing*, pp. 201–229. Oxford: Oxford University Press.
- Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal*, 94(4), 554-566.
- Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, 38, 301–315.
- Kelly, G. (2000). *How to teach pronunciation*. Harlow: Longman.
- Kim, I.D. & Lee, B.C (2015). *G9 English*. Seoul: NE Publishing Company.
- Kim, S. K. (2015). *High School English 1*. Seoul: NE Publishing Company.
- Park, J.W. (1994). Variation of Vowel Length in Korean. In Kim-Renaud, Y. (ed.), *Theoretical issues in Korean linguistics* (pp. 175-187). Stanford, CA: CSLI Publications for the Stanford Linguistics Society.
- Peperkamp, S., & Dupoux, E. (2002). A typological study of stress ‘deafness’. *Laboratory Phonology 7* (pp. 203–240). New York, NY: Mouton de Gruyter.
- Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, 29(2), 191-215.
- Lee, S., & Huang, H. (2008). Visual input enhancement and grammar learning: a meta-analytic review. *Studies in Second Language Acquisition*, 30(3), 307-331
- Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? the influence of accent on credibility. *Journal of Experimental Social Psychology*, 46(6), 1093-1096.
- Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, 39(3), 369-378.
- Levis, J. M., Sonsaat, S., Link, S., & Barriuso, T. A. (2016). Native and nonnative teachers of L2 pronunciation: Effects on learner performance. *TESOL Quarterly*, 50(4), 894-931.

- Lin, C., Wang, M., Idsardi, W., & Xu, Y. (2014). Stress processing in mandarin and Korean second language learners of English. *Bilingualism-Language and Cognition*, 17(2), 316-346.
- Lindemann, S. (2002). Listening with an attitude: A model of native-speaker comprehension of non-native speakers in the united states. *Language in Society*, 31(3), 419-441.
- Ministry of Education (2015). *The national curriculum of English*, Seoul: Ministry of Education.
- Munro, M. J., & Derwing, T. M. (1995a). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45(1), 73-97.
- Munro, M. J., & Derwing, T. M. (1995b). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, 38(3), 289-306.
- Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 49(supplement 1), 285-310.
- Munro, M. J. (2008). Foreign accent and speech intelligibility. In: Edwards, J. G. H., & Zampini, M. L. (eds.) *Phonology and second language acquisition* (pp. 193-218), Amsterdam: John Benjamins Publishing Company.
- Munro, M. J., & Derwing, T. M. (2015). Intelligibility in research and practice: teaching priorities. In: Reed, M., & Levis, J. (eds.) *The handbook of English pronunciation* (377-396). Malden, MA: Wiley Blackwell.
- NIKL (2018). *The rules of Korean orthography and the language use*. National Institute of the Korean Language. Retrieved from [http://www.korean.go.kr/front/reportData/reportDataView.do?mn\\_id=45&report\\_seq=944&pageIndex=1](http://www.korean.go.kr/front/reportData/reportDataView.do?mn_id=45&report_seq=944&pageIndex=1)
- Nelson, C.L., & Kang, S.Y. (2015). Pronunciation and world Englishes. In: Reed, M., & Levis, J. (eds.) *The handbook of English pronunciation* (320-329). Malden, MA; Wiley Blackwell.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32), New York: Cambridge University Press.
- Setter, J., & Jenkins, J. (2005). Pronunciation: State-of-the-art review article. *Language Teaching Vol*, 38(1), 1-17.

- Setter, J., & Sebina, B. (2017) English lexical stress, prominence and rhythm. In: Kang, O., Thomsson, R. I. and Murphy, J. M. (eds.) *The Routledge Handbook of Contemporary English Pronunciation* (pp. 137-153). New York, NY: Routledge.
- Segalowitz, N., & Hulstijn, J. (2005). Automaticity in Bilingualism and Second Language Learning. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371-388). New York, NY: Oxford University Press.
- Sicola, L. and Darcy, I. (2015). Integrating pronunciation into the language classroom. In: Reed, M., & Levis, J. (eds.) *The handbook of English pronunciation* (pp. 471-487). Malden, MA; Wiley Blackwell.
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30(4), 510–532.
- Smith, M. S. (1981). Consciousness-raising and the second language learner. *Applied Linguistics*, 2(2), 159
- Timming, A. R. (2017). The effect of foreign accent on employability: A study of the aural dimensions of aesthetic labour in customer-facing and non-customer-facing jobs. *Work, Employment & Society*, 31(3), 409-428.
- Szpyra-Kozłowska, J. (2015). *Pronunciation in EFL instruction: A research-based approach*. Buffalo, NY: Multilingual Matters.
- Taylor, L. (2005). Washback and impact. *ELT Journal*, 59(2), 154-155.
- Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics*, 36, 326–344.
- Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism-Language and Cognition*, 15(4), 905-916.
- Wang, Z., Arndt, A., Singh, S., Biernat, M., & Liu, F. (2013). "you lost me at hello": How and when accent-based biases are expressed and suppressed. *International Journal of Research in Marketing*, 30(2), 185-196.
- Zielinski, B. (2015). The segmental/Suprasegmental debate. In: Reed, M., & Levis, J. (eds.) *The handbook of English pronunciation* (pp. 397-412). Malden, MA: Wiley Blackwell.

## Appendix A

### English Language Background Questionnaire

**Participant Number:** \_\_\_

1. Name (pseudonym):
  
2. Current age (international age counting system):
  
3. Hometown (name of city):
  
4. Gender:
  
5. Years of residence in a country where English is spoken.  
(include all countries and length of residence)
  
6. At what age did you first learn English?
  
7. Your level of class in the current English institute (if applicable)
  
8. Self-assessment in speaking ability in English

Beginner	Intermediate	Advanced
----------	--------------	----------

9. Years of English instruction received:

**Appendix B**  
Recruitment Notice (English Version & Korean Version)

**Korean Participants Needed (English Version)**

I am conducting research for my MA thesis on the topic of how Korean speakers improve their English pronunciation.

**Task**

You will read English sentences. (High school level of English in South Korea)

**Requirement**

-Native Korean speakers with less than 1 year of experience in English speaking countries and who are willing to commit their time to the recording for two days. The second recording will take place two days after the first recording.

**Time**

First day: 30 minutes / Second day: 5 minutes

(The actual time for participants to read the sentences is 4 minutes. The rest of the time will be spent on filling out a questionnaire, signing a consent form, explanation on the nature of the research, and a training session to improve pronunciation)

**Venue**

Clearihue Building, University of Victoria  
(Participants will be notified of exact recording location afterwards)

**Compensation**

Participants will receive \$10 in cash after the experiment

**Contacts**

Text me at XXX-XXX-XXXX or email me at [xxx@xxx.ca](mailto:xxx@xxx.ca) if you are interested.

It is completely anonymous and confidential.

## 한국인 실험 참가자 모집 (Korean Version)

안녕하세요. 저는 빅토리아 대학교 응용언어학 석사 과정 대학원생입니다. 한국인의 영어 발음 향상을 위한 실험에 참여하실 분을 모집합니다.

### 실험개요

한국 고등학교 교육과정 수준의 단어로 구성된 영어 문장을 읽어주시면 됩니다.

### 참가자격

- 영어권 국가 거주 경험이 1년 미만인 한국인 누구나 가능
- 이틀에 걸친 녹음에 **모두 참여 가능한** 분. (첫 번째 녹음 후 이틀 후에 두 번째 녹음이 이루어 집니다. 예. 월요일 첫 번째 녹음, 수요일 두 번째 녹음)

### 소요시간

실험 첫째 날: 약 30분 / 실험 둘째 날: 약 5분  
(실제 영어 문장 읽기에 소요되는 시간은 약 4분이며 그 외 시간은 설문지 작성, 실험 참여 동의서 작성, 연구에 대한 설명, 영어 발음 향상을 위한 트레이닝에 사용됩니다)

### 실험장소

빅토리아 대학교 Clearihue Building (영어 문장 녹음이 이루어지는 장소는 실험참가자에게 개별적으로 알려드립니다.)

### 사례금

본 실험에 참여하실 경우 사례금으로 **10 달러를 현금으로 지급**하여 드립니다.

### 실험참여 희망 및 문의사항

실험에 참여하길 희망하시는 분은 아래 연락처로 문자 혹은 이메일을 보내주세요.

XXX-XXX-XXX

[xxx@xxx.ca](mailto:xxx@xxx.ca)

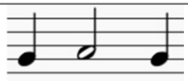
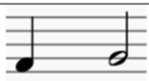

본 실험 참가자의 개인정보는 연구 이외의 목적으로는 사용되지 않으며 철저히 비밀이 보장되어 관리됩니다.

## Appendix C

### Sample Lesson Plan

#### Objectives

1. Ss will be able to learn how to give prominence to a stressed syllable.
2. Ss will learn 5 sets of noun/verbs pairs (*produce, project, increase, object, and conflict*) that only differ in the location of stress and be able to work out the rules on their own and apply the rules to new sets of pairs (*transport, record, and progress*).

Activity	Brief description of activity	Time (Min)	Materials
Intro (WC)	<div style="display: flex; justify-content: space-around; align-items: flex-start;"> <div style="text-align: center;"> <p>banana</p>  <p>ba na na</p> </div> <div style="text-align: center;"> <p>today</p>  <p>to day</p> </div> <div style="text-align: center;"> <p>pencil</p>  <p>pen cil</p> </div> </div> <p>Ss familiarize themselves with the use of English stress diacritics. Ss are presented with three slides above and instructed to pronounce a syllable with diacritic with slightly higher pitch and longer duration than the rest. The musical notation is employed to provide Ss with better understanding on how to give relatively higher pitch and longer duration on stressed syllables than unstressed ones.</p> <p>T models the pronunciation and ask Ss to repeat. T asks a couple of students to read the words on the slide.</p>	5	Musical notation charts
Paired-reading activity (P)	<p>In pairs, one student is given a story with diacritics showing stress placement, and the other student given a story without diacritics. The student with diacritics reads</p>	10	Per pair: Three sheets of paper (one with diacritics)

	<p>the following story with special attention to the stressed syllables to their partner. Their partner puts diacritics on the stressed syllables when listening to their counterpart. After the activity, they check out the answers and switch roles.</p> <p>-----</p> <p>Story: My parents on the farm produce fresh produce every year. They project a decrease in the harvest of apples by 20 percent because of a fire. They have a project to decrease their work force in order to rebound from the fire. Sometimes they have me work for them without paying me. I think I need to object to their demand, which might create conflict with my parents. They say they love me all the time, but their actions conflict with their words. It seems that their sole object in their life is to be rich.</p>		and two without diacritics) Appendix 2.
Working out the rule (G)	In groups, Ss figure out why orthographically identical words have different stress and induce rules out of the given story. Ss comes up to the board to write their own answers. Ss check the answers with T.	5	
Creative Writing (G)	In groups of 4, Ss create a story using the same 5 sets of noun / verbs pairs from Paired-reading activity ( <i>produce, project, increase, object, and conflict</i> ) and 3	20	

	new pairs ( <i>transport, record, and progress</i> ) that differ only in stress placement. Ss must use all the words (16 in total) to create a coherent story and put diacritics on the stressed syllables.		
Presentation (WC)	A presenter for each group reads their story to the whole class with special attention on the stressed syllables. Other than the group presenting, the rest of the groups evaluate a story based on the evaluation criteria.	10	Evaluation sheet