교육학박사학위논문

Exploring the Role of Prosody in Korean EFL High School Students' Oral Reading Fluency

한국인 고등학교 영어 학습자의 영어 읽기 유창성에서 운율적 자질의 역할에 관한 탐구

2023년 8월

서울대학교 대학원

외국어교육과 영어전공

유 지 선

Exploring the Role of Prosody in Korean EFL High School Students' Oral Reading Fluency

by

Jiseon Ryu

A Dissertation Submitted to
the Department of Foreign Language Education
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in English Language Education

At the
Graduate School of Seoul National University

August 2023

Exploring the Role of Prosody in Korean EFL High School Students' Oral Reading Fluency

한국인 고등학교 영어 학습자의 영어 읽기 유창성에서 운율적 자질의 역할에 관한 탐구

지도교수 이 병 민

이 논문을 교육학 박사 학위논문으로 제출함

2023년 7월

서울대학교 대학원
외국어교육과 영어전공
유 지 선

유지선의 박사학위논문을 인준함

2023년 7월

위 원 장 ______________________

부위원장 ______________________

위 원 ______________________

위 원 ______________________

위 원 ______________________

Exploring the Role of Prosody in Korean EFL High School Students' Oral Reading Fluency

by

Jiseon Ryu

A Dissertation Submitted to the Department of Foreign Language Education

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in English Language Education

at the Graduate School of Seoul National University

July 2023

APPROVED BY DISSERTATION COMMITTEE:

____
YOUNGSOON SO, COMMITTEE CHAIR

____
IN CHULL JANG

____
BYUNGMIN LEE

____
EUNJOU OH

____
JUNGEUN CHOI

Abstract

Exploring the Role of Prosody in Korean EFL High School Students' Oral Reading Fluency

Jiseon Ryu

English Major, Dept. of Foreign Language Education

The Graduate School of Seoul National University

Oral reading fluency (ORF), defined as the ability to read text accurately, quickly, and expressively, is considered a crucial skill for reading comprehension success. Founded on the theoretical underpinnings of information processing models and numerous empirical studies, ORF has drawn much attention from reading researchers and educators. Given its significance, various interventions have been developed and incorporated into the classroom, alongside numerous assessment tools for measuring ORF in both first language (L1) and second language (L2) contexts. However, there has been ongoing criticism of ORF assessment tools for their exclusive focus on measuring reading accuracy and rate, while neglecting the critical component of reading prosody.

Reading prosody refers to the melodic quality of oral reading such as tempo, pauses, pitch, and rhythm. Research has shown that as readers become more proficient and fluent in reading, their oral reading becomes smoother and

more speech-like, progressing from a word-by-word, hesitant manner to a more fluid style. In addition, the importance of reading prosody has been underscored as it plays a critical role in reading comprehension by holding information being processed in working memory or acting as a cognitive scaffold for semantic processing (Frazier et al., 2006). Even though little is known about reading prosody so far, a growing body of research has shown that reading prosody is an essential component of ORF and reading comprehension in L1 reading.

In the L2 context, ORF also has been increasingly establishing itself as a legitimate component of reading skill, and evidence has been accumulated suggesting its significant relationship with reading comprehension. However, the research of reading prosody is at the beginning stage in the L2 context.

Therefore, the present study attempted to explore various aspects of the roles reading prosody plays in L2 reading. First, it examined distinctive characteristics of reading prosody features differing according to fluency skill. Second, this research sought to evaluate the supposition regarding the characterization of ORF, specifically assessing whether the combined factors of reading accuracy and rate (i.e., text reading efficiency) and reading prosody can be considered a single construct, known as ORF. Additionally, the investigation aimed to examine the role of reading prosody in accounting for reading comprehension.

A total of 90 Korean high school students learning English as a foreign

language (EFL) participated in the study. The students were subjected to a battery of reading assessments to evaluate their decoding skills, text reading efficiency (TRE), and reading comprehension. To examine and measure reading prosody features, spectrographic analysis was employed as the methodology. By employing a speech sound wave analysis tool such as Praat (Boersma & Weenink, 2023), the visuo-graphic representation of sound waveform was generated. The observation of this spectrogram made it possible to identify and extract various prosodic features.

As a result of spectrographic analyses, six features of reading prosody were extracted: intrasentential pause ratio (IntraP_ratio), intrasentential pause duration (IntraP_duration), ungrammatical pause ratio (UGP_ratio), intersentential pause ratio (InterP_ratio), overall intonation contour (Pitch_SD), and pitch changes at the sentence-final position (Pitch_SF). These features were subsequently analyzed using a series of one-way analysis of variance (ANOVAs), confirmatory factor analysis (CFA), or a series of hierarchical linear regressions.

The study found that pause-related variables such as IntraP_ratio, IntraP_duration, and UGP_ratio were distinctively different among students with different fluency skills, indicating that these variables may be indicative of a student's level of fluency. In contrast, it was found that Pitch_SD did not exhibit significant differences based on fluency level, suggesting that the overall pitch

pattern may remain similar across Korean L2 readers regardless of their level of fluency. Meanwhile, significant variations in pitch changes at the sentence-final position and pause duration across sentences were only evident among higher fluency groups, indicating that these changes primarily manifest among L2 readers who have attained a sufficient level of decoding and fluency.

In addition, it was found that reading prosody and TRE align and contribute to the unified construct of ORF. However, among reading prosody features, only pause-related variables were indicative of ORF.

Lastly, the present study demonstrated that reading prosody plays a crucial role in predicting reading comprehension by acting as a mediator between decoding skills and reading comprehension. However, the results indicate that reading prosody did not account for additional variance in reading comprehension beyond TRE, especially when TRE was given priority in the order of entry in the hierarchical linear regression. When another regression analysis was conducted to inspect the relationship between reading prosody and TRE, a substantial overlap of these variables was detected, suggesting that the effect of reading prosody on reading comprehension would be masked if TRE is entered before reading prosody factors. In fact, when reading prosody was given priority in the order of the entry in the regression analysis, it was found to remain significant alongside TRE.

The current dissertation has important pedagogical implications for L2

reading instruction. The findings suggest that focusing on the development of reading prosody is critical for enhancing L2 reading proficiency. L2 teachers should consider incorporating activities that explicitly address reading prosody, such as choral reading or repeated reading, into their instruction to help L2 learners improve their ORF. Given the mediating role of reading prosody in the relationship between decoding skills and reading comprehension, instructors are encouraged to integrate ORF monitoring activities using the ORF scale into their assessment protocols. The current study's outcomes can be leveraged to develop a valid and reliable ORF scale as in previous studies.

Overall, the findings of this study highlight the importance of a comprehensive approach to L2 reading instruction that targets ORF, especially reading prosody to promote efficient and effective reading comprehension in L2 learners.

Keywords: L2 reading prosody, oral reading fluency, text reading efficiency, spectrographic analysis, reading comprehension, reading assessment, second language reading,

Student number: 2018-35246

Acknowledgements

"What does it mean to read in a foreign language?" Or, "what does it mean to be able to read in English?" My doctoral course was a journey that began with a resolute determination to conquer the "ultimate boss" called "READING," seeking answers to these very questions. While I haven't yet unearthed the answers I sought, nor have I completely defeated the boss, this research project has illuminated the paths I should tread and the avenues I aspire to explore. I would like to express my heartfelt gratitude to all those who supported me on my doctoral journey.

Foremost, I extend my sincere thanks to my dear students who willingly participated in the experiments, dedicating their efforts and time to make their teacher "a doctor." Their attentive listening during experiment explanations, enthusiastic engagement in reading and responding despite the weight of responsibility and tension in the presence of the observing teacher with big eyes, and the invaluable moments spent with students who approached the experiments with curiosity and posed questions—all collectively transcended mere experimentation; it was a mutual learning for one another.

I am deeply grateful to the teachers who guided me through the academic journey. First and foremost, my appreciation extends to Professor Lee Byungmin, who guided me as my advisor during both my master's and doctoral pursuits. Without his guidance, I couldn't have even embarked on, let alone completed, my doctoral voyage. His teachings have been the bedrock upon which I've built my comprehension of the essence of reading and the intricacies of research. Particularly during times when these concepts remained elusive to me, his occasionally stern yet always warmly supportive mentorship has been an invaluable compass steering me through my academic years.

I am also grateful to my committee members, Professor So Youngsoon, Jang In Chull, Oh Eunjou, and Choi Jeongeun, who lent their expertise as examiners for my dissertation. Their generous guidance and insightful feedback during the evaluation process have propelled me to delve deeper into the paper and undertake vital revisions. In particular, I reserve special gratitude for Professor Jang In Chull, who not only peppered the paper with "provocative" questions but also provided invaluable guidance about my identity as a researcher and career path.

I must also express my sincere appreciation to Professor Yang Byunggon from Pusan National University's English Education Department. During a juncture when the data analysis reached an impasse, his timely and unwavering response to my email and teaching guided me towards the right course. Lastly, heartfelt thanks go to Professor Oh Sunyoung, who consistently showed warm encouragement, both in the doctoral proposal and through the sporadic emails I sent.

Furthermore, I wish to express my sincere appreciation to my friends and colleagues. I would like to thank Kim Sayeoung, who shared moments with me in the lab throughout my doctoral years. Our shared concerns, shared meals, book discussions, and unforgettable moon-gazing moments on the rooftop while humming *Seoul's Moon* will forever be etched in my memory. To my friends who supported and cheered me on in the data collection process as if it were their own projects, Lee Jinyoung and Yoon Jihyun, I extend heartfelt thanks. I also would like to thank my precious friends, Choi Jieun, Noh Soyoung, Kim Minjin, and Jung Yoonkyung, who silently stood by during my challenging moments, offering warm comfort and encouragement.

Lastly, I want to convey my profound love and gratitude to my parents — who steadfastly supported their stubborn daughter from the shadows — my precious grandmother, and my still beloved mom in the heavens.

Table of Contents

Abstract i
Acknowledgements vi
Table of Contents ix
List of Tables xii
List of Figures xiii

Chapter 1. Introduction 1
1.1 Motivation for the Study and Statement of the Problem 1
1.2 Definition of the Terms 12
1.3 Research Questions 15
1.4 Significance of the Study 20
1.5 The Outline of the Thesis 22

Chapter 2. Literature Review 24
2.1 Oral Reading Fluency 24
2.1.1 Theoretical Frameworks of Oral Reading Fluency 24
2.1.2 The Constructs of Oral Reading Fluency 28
2.2 The Nature of Reading Prosody 31
2.2.1 Connection of Speech Prosody to Reading Prosody 32
2.2.2 Assessment of Reading Prosody 36
2.2.3 Characteristics of Prosodic Features 45
2.2.3.1 Pause Features 49
2.2.3.2 Pitch Features 52
2.3 Reading Prosody and Reading Comprehension 56
2.3.1 Reading Prosody as a Predictor of Reading Comprehension 56
2.3.2 The Directionality of the Relationship between Reading Prosody and Reading Comprehension 60

Chapter 3. Methodology 64
3.1 Participants 64
3.2 Reading Assessment and Procedures 66
3.2.1 General Assessment Procedure 67
3.2.2 Decoding Skill Measurement 68
3.2.3 Text Reading Efficiency Measurement 68

3.2.4 Reading Comprehension Assessment 70
3.3 Reading Prosody Measurement 73
3.3.1 Intrasentential Pausal Ratio/Duration 74
3.3.2 Ungrammatical Pause Ratio 76
3.3.3 Intersentential Pause Duration 78
3.3.4 Pitch Changes at the Sentence-final Position 79
3.3.5 Overall Intonation Contour 82
3.4 Statistical Analysis 85
3.4.1 Statistical Analysis for RQ 1 85
3.4.2 Statistical Analysis for RQ 2 88
3.4.3 Statistical Analysis for RQ 3 92

Chapter 4. Results 99
4.1 Descriptive Statistics and Correlation 99
4.2 Differences in Reading Prosody Features Depending on Reading Skill 108
4.3 Dimensionality of Oral Reading Fluency 119
4.4 Reading Prosody as a Predictor of Reading Comprehension 122
4.4.1 Unique Contributions of Reading Prosody to Reading Comprehension Beyond the Decoding Skills 123
4.4.2 Reading Prosody as a Mediator to Reading Comprehension ... 127
4.4.3 Predictability of Reading Prosody Beyond Text Reading Efficiency 131
4.4.4 Summary of the Results on Predictability of Reading Prosody for Reading Comprehension 137

Chapter 5. Discussion 139
5.1 Different Characteristics of Reading Prosody Features as a Function of Fluency Skills 139
5.2 Oral Reading Fluency as a Unitary Construct 147
5.3 Roles of Reading Prosody in Reading Comprehension 149

Chapter 6. Conclusion 156
6.1 Major Findings 156
6.2 Pedagogical Implications 158
6.3 Limitations and Suggestions for Future Research 162

References **166**
Appendices **182**
Appendix A: Consent Form 182
Appendix B 183
Reading Material A Modified from May (2014) 183
Reading Material B Modified from May (2014) 184
Appendix C 185
Parameter Estimate Coefficients of Two-Dissociable-Constructs Model of ORF 185
Graphic Illustration of CFA Results of Two-Dissociable-Constructs Model of ORF 186
국 문 초 록 **187**

List of Tables

Table 2.1 NAEP Oral Reading Fluency Scale (Pinnell, 1995) 38
Table 2.2 Multidimensional Fluency Scale (Rasinki et al., 2009) 40
Table 2.3 Comprehensive Oral Reading Fluency Scale (Benjamin et al., 2013) 42
Table 2.4 Features of Reading Prosody Measured in the Previous Studies 47
Table 3.1 A Flesch-Kincaid Grade Levels of MCSAT RC Passages 72
Table 3.2 Comparisons of Three Groups on Reading Comprehension and Fluency Skill 87
Table 3.3 CFA Goodness - of - fit Indices 91
Table 3.4 Factor Loading for the Two Components 96
Table 4.1 Descriptive Statistics of the Variables (N=90) 100
Table 4.2 Bivariate Correlation among the Variables 107
Table 4.3 Descriptive Statistics of Reading Prosody Features by Groups 109
Table 4.4 Difference of Reading Prosody Features Depending on L2 Fluency Skills 110
Table 4.5 Post - hoc Comparisons of L2 Reading Prosody Features across the Three Fluency Skill Groups 115
Table 4.6 The Model Fit Indices of Two Alternative Models 119
Table 4.7 Parameter Estimate Coefficients of the Unitary-construct Model 121
Table 4.8 Comparison of the Models with Different Pause Variables 124
Table 4.9 Contributions of Reading Prosody to RC Beyond Decoding Skills 125
Table 4.10 Model Summary of the Mediation Analysis 130
Table 4.11 Contributions of Reading Prosody to RC Beyond TRE 132
Table 4.12 Overlap between Reading Prosody and TRE 134
Table 4.13 Alternative Model for the Contribution of Reading Prosody to RC 135

List of Figures

Figure 3.1 Spectrographic Representation of Measurement of Pause Features .. 77
Figure 3.2 Spectrographic Representation of Measurement of Sentence-final Pitch Declination 81
Figure 3.3 Spectrographic Representation of Measurement of Pitch_SD 84
Figure 3.4 Unitary-Construct Model of ORF 89
Figure 3.5 Two-Dissociable-Constructs Model of ORF 90
Figure 3.6 The Hypothesized Mediation Model of Reading Prosody 97
Figure 4.1 Frequency of Pitch_SD 103
Figure 4.2 Frequency of Pitch_SF 103
Figure 4.3 Frequency of IntraP_ratio 104
Figure 4.4 Frequency of UGP_ratio 104
Figure 4.5 Frequency of IntraP_duration 105
Figure 4.6 Frequency of InterP_duration 105
Figure 4.7 Differences of Pitch_SD across Groups 113
Figure 4.8 Differences of Pitch_SF across Groups 113
Figure 4.9 Differences of IntraP_ratio across Groups 113
Figure 4.10 Differences of UGP_ratio across Groups 114
Figure 4.11 Differences of IntraP_duration across Groups 114
Figure 4.12 Differences of InterP_duration across Groups 114
Figure 4.13 Comparison of Pause Patterns between the Two Participants of Different Fluency Skills 117
Figure 4.14 The Unitary-construct Model of ORF 122
Figure 4.15 RC by Standardized Residual Scatter Plot 125
Figure 4.16 The Mediation of PAUSE between Decoding Skills and Reading Comprehension 130

Chapter 1. Introduction

The current study aims to explore the roles of prosodic rendering that Korean high school students learning English as a foreign language (EFL)[1] make while they read aloud a text. To this end, the current study investigates the reading prosody as one of the three components - accuracy, fluency, and reading prosody - of oral reading fluency. This chapter provides an introduction to the thesis by first highlighting the motivations that led to the study's development. Subsequently, the definition of terms and the research question that this study aimed to address are presented. Finally, the chapter concludes by outlining the structure of the thesis and providing a brief description of each of the following chapters.

1.1 Motivation for the Study and Statement of the Problem

Oral reading fluency (ORF), defined as the ability to read text accurately, quickly, and expressively, is considered a crucial skill for reading comprehension success. The National Reading Panel report (NICHD, 2000) brought ORF to the forefront of first (L1) and second language (L2) reading research and education. In the past, ORF was overlooked as a reading skill

---

[1] In this thesis, the terms *second language* and *foreign language* are not differentiated and used interchangeably since the two terms refer to any language acquired by a speaker in addition to their native language.

(Allington, 1983; Dowhower, 1991), but it has since become one of the most studied, assessed, and debated components of reading (Benjamin, 2012; Kim et al., 2021b).

The importance of reading fluency as a fundamental aspect of reading comprehension is widely accepted in current research (Kim 2020a, 2020b; Kim et al., 2021b; Wolters et al., 2020). This understanding is based on information-processing models that highlight the interplay between lower-level and higher-level processing skills during reading (Fuchs et al., 2001). One of the classic theories in cognitive psychology is the automaticity theory, which was posited by LaBerge and Samuels (1974). The automaticity theory assumes that because working memory capacity is limited, not all the information involved in reading can be processed simultaneously. Therefore, lower-level processing skills such as word recognition and syntactic parsing should be processed automatically, allowing higher-level processing skills to be executed with minimal conscious effort. If a reader devotes too much attention to lower-level processing, higher-level skills such as comprehension may be hindered.

Unlike this bottom-up serial processing model where lower-level processing skills constrain higher-level processing, Stanovich (1980) assumes that human cognition could modulate this processing more flexibly. In the interactive-compensatory model, he claims that higher-level skills such as context facilitation can compensate for the lower-level processing when the efficient execution of the lower-level skills is restrained. In a similar vein, the verbal efficiency theory (Perfetti, 1988) claims that to the extent that word recognition becomes efficient, reading comprehension can operate smoothly.

Although the above three theories slightly differ in explaining how lower-level operates and how much it affects reading comprehension, it is undeniable that they all highlight the critical role of low-level skills in reading comprehension. In this regard, ORF has drawn the attention of numerous reading scholars and educators since the automatic lower-level process can be represented in readers' accurate, efficient, and expressive reading (Fuchs et al., 2001; Rasinski et al., 2011).

First and foremost, central to the research of ORF was the relation to reading comprehension. A plethora of research has investigated whether ORF could be a proxy for reading comprehension (Fuchs et al., 2001; Jenkins et al., 2000, 2003; Kim et al., 2010; Marston et al., 1986; Ryu & Lee, 2021; Shinn et al., 1992; Valencia et al., 2010; Wise et al., 2010). Based on the evidence on the inherent relationship between ORF and reading comprehension, subsequent studies have been conducted to investigate the relative importance of ORF in comparison to other reading-related skills (e.g., listening, decoding, and phonological awareness) (Baker et al., 2012; Baker et al., 2008; Danne et al., 2005; Jenkins et al., 2000, 2003; Kim et al., 2014; Kim et al., 2010; Kim et al., 2011; Kim et al., 2012; Kim & Wagner, 2015; Kim, 2015; Lee, 2020; Riedel, 2007; Roehrig et al., 2008; Silverman et al., 2013; Tilstra et al., 2009; Woo & Jeong, 2018; Woo & Jung, 2012, 2017). In addition, thanks to the high practicality and validity of its assessment, numerous longitudinal studies on students' reading development using ORF measurement outcomes have been conducted, and thus have provided insight into the nature of ORF and reading development. (Al Otaiba et al., 2009; Hintze & Silberglitt, 2005; Jenkins et al., 2007; Kim et al., 2010; Kim et al., 2021a, 2021b; Kim & Wagner, 2015; Kim, 2015; Quirk et al., 2009; Solari et al., 2014; Speece & Ritchey, 2005; Veenendaal et al., 2015).

As such, the importance of ORF in L1 reading development has been extensively researched, thanks to the efforts of numerous scholars. The Direct and Indirect Effects of the Reading model has recently identified ORF as a key "proximal component (p.470)" that directly impacts reading comprehension (Kim, 2020a, 2020b). Figure 1.1 shows that ORF is founded on word reading skills and listening comprehension, and acts as a mediator between these skills and reading comprehension. This research has provided compelling evidence for the critical role of ORF in reading comprehension, highlighting its importance as a key component of reading development.

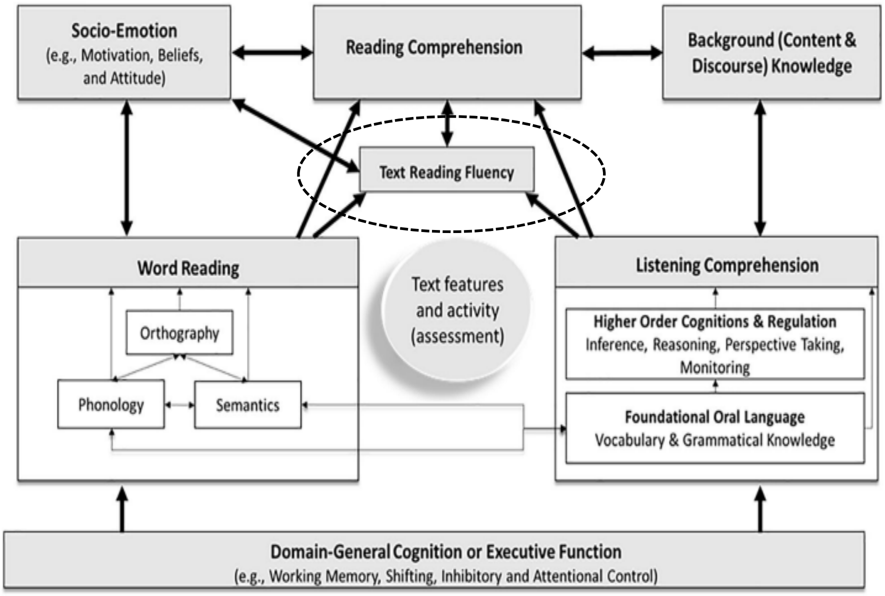


Figure 1.1 Direct and Indirect Model of Reading (DIER; Kim, 2020b)[2]

The growing recognition of oral reading fluency's crucial role in reading has prompted efforts to integrate it into L1 reading instruction and assessment. In order to enhance children's ORF, numerous intervention and instruction programs have been proposed and actively investigated for their impact on ORF development. Of particular interest is the instructional efficacy of repeated reading instruction (Chang, 2012, 2019; Dowhower, 1994; Gorsuch & Taguchi, 2008, 2010; Lo et al., 2011; Taguchi et al., 2012; Taguchi et al., 2004; Therrien

---

[2] Kim has replaced the commonly used phrase "*oral reading fluency*" with "*text reading fluency*" to underscore that this ability extends beyond solely the oral mode, despite the fact that it is commonly assessed through oral reading evaluations. (Kim 2020a, 2020b; Kim et al. 2021a, 2021b). Kim mentioned that this term is distinguished from *text reading efficiency* in that *text reading fluency* encompasses the accuracy, speed, and prosody of reading text while text reading efficiency refers to the construct which only includes the accuracy and speed of text reading.

& Kubina Jr, 2006; Yeganeh, 2013). Additionally, fluency-oriented reading instruction, such as readers' theater, has been shown to be effective in enhancing ORF mostly in L1 contexts (Clark et al., 2009; Keehn, 2003; Lekwilai, 2014; Ng & Boucher-Yip, 2010; Thienkalaya & Chusanachoti, 2020; Young & Rasinski, 2009).

In terms of assessment, ORF has also garnered significant attention among education practitioners as a measure of reading ability due to its quick and straightforward administration, its well-established correlation with reading comprehension, and its capacity to provide useful diagnostic information regarding students' reading development. To this end, norms for ORF have been developed and consistently updated to facilitate the diagnosis and monitoring of students' progress in reading (Hasbrouck & Tindal, 1992, 2006, 2017). A plethora of assessment kits and programs, such as AIMSWEB (Shinn & Shinn, 2002), informal reading inventories (e.g., Roe & Burns, 2010; Leslie & Caldwell, 2021), curriculum-based measures (CBM; Deno & Fuchs, 1987), Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good III & Kaminski, 2002), Gray Oral Reading Test (GORT; Wiederhold & Bryant, 2012), and Woodcock Reading Mastery Tests III (WRMT; Woodcock, 2011), have emerged in the field and are commonly employed to assess children's reading development.

Notwithstanding the advantages of ORF as an assessment tool and its popularity among educators, the method used to measure it has been a topic of debate from the outset. The prevailing approach to assessing ORF involves counting the number of words students read accurately in a minute (Words Correct Per Minute; WCPM). However, this method has been heavily criticized by Samuels (2007), who argues that WCPM merely measures "the speed of barking at print (p.563)," and Morris et al. (2013) who contend that reading rate should be evaluated in "the reading-for-meaning-context (2013, p. 62)."

Another critique of the assessment of ORF stems from empirical evidence on the existence of word callers and gap fillers (Hamilton & Shinn, 2003; Knight-Teague et al., 2014; Meisinger et al., 2009; Meisinger et al., 2010; Quirk & Beem, 2012). Word callers are individuals who exhibit good reading fluency but minimal comprehension, while gap fillers are defined as those who display the opposite pattern of reading (Duke & Carlisle, 2011). Although the extent to which word callers and gap fillers exist among L1 readers is yet to be established, the percentage of these reader groups is known to increase over time (Meisinger et al., 2010; Schwanenflugel & Knapp, 2015). Moreover, Quirk and Beem (2012) reported that a wider gap may exist between reading fluency and reading comprehension among L2 readers, with 55.5% of total participants exhibiting this gap in their study. Since the significance of ORF as a reading skill is founded on its association with reading comprehension, the validity of ORF assessment may only be supported by the concurrent relationship with reading comprehension. In other words, the mere existence of word callers and gap fillers in the reader profile has made it difficult to justify the assessment of ORF.

Lastly, it should be noted that a critical issue of validity arises from the inconsistency between the concept of ORF and the methods used to evaluate it. While the definitions of ORF may differ across studies, they typically include three key aspects: reading rate, reading accuracy, and reading prosody. To address this issue, a widely accepted definition of ORF was proposed by Kuhn, Schwanenflugel, and Meisinger (2010) as follows:

> Fluency combines accuracy, automaticity, and oral reading prosody, which, taken together, facilitate the reader's construction of meaning. It is demonstrated during oral reading through ease of word recognition, appropriate pacing, phrasing, and intonation (p. 240).

Considering this perspective, reading prosody is widely recognized as a fundamental component of ORF (Kuhn et al., 2010; Schwanenflugel & Benjamin, 2016; Wolf & Katzir-Cohen, 2001). However, many widely-used ORF assessment tools have traditionally excluded the measurement of reading prosody, leading to ongoing criticisms of the validity of measuring only WCPM. Hence, it has been argued that the inclusion of reading prosody in ORF assessments is essential for a comprehensive understanding of a student's overall reading abilities (Godde, 2020; Kim et al., 2021a; Kuhn et al., 2010; Wolters et al., 2020)

To date, insufficient attention has been devoted to reading prosody in reading research in comparison to other components. The exclusion of reading prosody from ORF research can be attributed to two primary reasons. Firstly, while measuring ORF traditionally involves a straightforward and apparent approach (i.e., WCPM), a distinct definitional and methodological consensus for evaluating reading prosody has yet to be established. Upon closer examination of existing ORF rating scales, it becomes evident that they are built on divergent assumptions regarding reading prosody. For instance, the creators of the National Assessment of Educational Progress (NAEP) Oral Fluency Scale primarily focused on phrasing, arguing that it could encapsulate other aspects of prosody such as intonation, stress, and pauses (Pinnell, 1995). Meanwhile, Zutell and Rasinski's (1991) Multidimensional Fluency Scale included four reading prosody components: expression and volume, phrasing, smoothness, and pacing. More recently, Benjamin et al. (2013) employed spectrographic analysis to identify two distinct components of reading prosody: intonation and pausing. These different subcomponents suggest that the scales were developed based on varying hypotheses about the nature and characteristics of reading prosody in oral reading.

Another reason for the exclusion of reading prosody from ORF measurement may be the longstanding view that WCPM is a strong predictor of reading comprehension. Despite ongoing debates on the validity and reliability of WCPM measurement, its predictive power has consistently been demonstrated in numerous studies (Fuchs et al., 2001; Jenkins et al., 2000, 2003; Kim et al., 2010; Marston et al., 1986; Ryu & Lee, 2021; Shin & McMaster, 2019; Shinn et al., 1992; Valencia et al., 2010; Wise et al., 2010; Yeo, 2010). For instance, Yeo's (2010) meta-analysis reported a medium to large effect size (r = .69) for the relationship between reading comprehension and ORF, and a more recent meta-analysis by Shin and McMaster (2019) found a similar mean effect size (r = .63). Given the robustness of WCPM as a predictor of reading comprehension, it may have been perceived as unnecessary to include reading prosody as an additional measure in ORF assessment tools. However, this view may overlook the potential contribution of reading prosody to the reading process, as well as the importance of developing more comprehensive measurement tools that can capture multiple aspects of reading fluency.

Empirical studies have provided evidence indicating that incorporating reading prosody can enhance the predictive power of ORF for reading comprehension (Benjamin & Schwanenflugel, 2010; Miller & Schwanenflugel, 2006, 2008; Valencia et al., 2010). For instance, Valencia et al. (2010) reported that adding prosody to a WCPM measure improved the predictability of ORF for reading comprehension, particularly in upper elementary grades where WCPM becomes less effective as a predictor of overall reading ability. Moreover, researchers have explored the possible mechanisms underlying the relationship between reading prosody and reading comprehension, although this area of inquiry remains relatively understudied (Wade-Woolley et al., 2022; Wolters et al., 2020). Also, it is increasingly accepted that a causal relationship between reading prosody and reading comprehension exists (Groen et al., 2019; Klauda & Guthrie, 2008; Miller & Schwanenflugel, 2006, 2008; Veenendaal et al., 2016).

Similar to L1 research, recent studies in L2 reading have indicated that ORF is a critical predictor of L2 reading comprehension and an essential skill to develop (Ahn & Kang, 2016; Ching Pey et al., 2014; Crosson & Lesaux, 2010; Jeon, 2012; Kang, 2011, 2020, 2021; Kim, 2012; Lee, 2014; Lee & Chen, 2019; Lee, 2018, 2020; Newell et al., 2020; Ryu & Lee, 2021; Saiegh-Haddad, 2003; Shin, 2021; Shin, 2018). However, L2 studies on ORF have mainly focused on reading rate and accuracy, without considering the role of reading prosody. Despite the crucial role of reading prosody in L1 reading, its contribution to L2 reading remains largely unexplored. Thus, a critical need exists for investigating the role of reading prosody in L2 reading comprehension, as this can provide insights into the unique challenges that L2 readers face and inform the development of effective L2 reading instruction.

Accordingly, this research endeavors to investigate the role of reading prosody in L2 reading, and more specifically, to scrutinize the characteristics of L2 reading prosody features that may exhibit variance contingent upon reading proficiency levels. Moreover, the predictive validity of reading prosody in L2 reading is evaluated through an examination of its relationships with other ORF components and reading comprehension. To achieve these goals, the study employs a spectrographic analysis approach, which facilitates the objective observation and measurement of prosodic features by generating a visual representation of waveforms that exhibit variation over time (Benjamin & Schwanenflugel, 2010; Benjamin et al., 2013; Miller & Schwanenflugel, 2006, 2008; Schwanenflugel et al., 2004; Schwanenflugel et al., 2015). This methodology could provide a more detailed and nuanced understanding of prosodic features compared to subjective or manual rating methods. By investigating the role of reading prosody in L2 reading, this study could contribute to the development of effective L2 ORF assessment and instruction.

1.2 Definition of the Terms

In this section, a comprehensive elucidation of the technical terminology and measurements employed in the data collection process is provided.

i. Oral reading fluency: oral reading fluency (ORF) is an umbrella term that encompasses its three components - reading rate, accuracy, and reading prosody.

ii. Text reading efficiency: in the present study, text reading efficiency (TRE) was used to refer to the construct which is comprised of reading rate and accuracy only, and does not include reading prosody. Its index is Words Correct Per Minute (WCPM) which indicates the number of words a reader can accurately read within one minute. This term has been introduced and adopted in the literature (Kim et al., 2021a, 2021b) to distinguish it from ORF and highlight that ORF is a broader term that includes reading prosody as well as reading rate and accuracy.

iii. Decoding skill: it refers to the ability of a reader to convert written symbols (i.e., letters and letter combinations) into their corresponding sounds, and then blend those sounds together to form recognizable words. It is a fundamental component of reading proficiency and is often associated with early reading development (Gough & Tunmer, 1986; Wolf & Katzir-Cohen, 2001). As in the previous studies (Kim et al., 2021b; Miller & Schwanenflugel, 2006; Schwanenflugel et al., 2004), the decoding skill was measured by using a reading battery (i.e., Woodcock Reading Mastery Test III - Word Identification subtest) that specifically target word-level decoding abilities.

iv. Fluency skill: in this study, *fluency skill* was used as an index for reading proficiency. As an operational definition, it refers to the ability to read words and sentences fast and accurately. As in the previous studies on L1 reading prosody, this skill was utilized as a criterion to categorize the participants into different skill groups (Benjamin & Schwanenflugel, 2010; May, 2014; Miller & Schwanenflugel, 2006). Following the previous studies, the composite score for *fluency skill* was derived by combining the decoding scores and TRE of the participants, via principal component analysis. The resultant score was used to establish three groups, based on which the investigation into the variation of reading prosody features was conducted.

v. IntraP_ratio: any temporal gap that lasts longer than 100ms is considered a pause (Benjamin, 2012; Dowhower, 1991; Miller & Schwanenflugel, 2006, 2008). IntraP_ratio refers to the percentage of pauses readers made within the sentence. It is calculated by dividing the number of pauses by the total number of spaces within a sentence.

vi. UGP_ratio: it refers to the percentage of ungrammatical pauses readers made within the sentence. The pauses which deviated from phrase and clause boundaries were counted, divided by the total number of spaces within a sentence.

vii. IntraP_duration: it refers to the total duration of pauses made in a sentence.

viii. InterP_duration: it refers to the duration of pauses from the end of a sentence to the following sentence.

ix. Pitch_SF: it refers to the pitch changes at the end of the sentence. To calculate its index, the pitches at the final peak and the endpoint of a declarative sentence were measured. The difference between these two values was determined by subtracting the endpoint pitch from the last peak pitch.

x. Pitch_SD: it refers to the overall intonation contour of a sentence. To determine its index, the pitch was measured at the peak of each word in a sentence, and the standard deviation of these values was calculated. This resulting value was then divided by the total number of sentences that participants read.

1.3 Research Questions

With the aforementioned problems and the necessity of the study, this thesis aims to examine whether reading prosodic features have different characteristics depending on fluency skill, whether reading prosody and text reading efficiency align and contribute to the unified construct of ORF, and whether reading prosody can be valid predictors of reading comprehension. Thus, the current study is guided by the following research question:

1. Does Korean EFL high students' reading prosody (e.g., intrasentential pause ratio, ungrammatical pause ratio, intrasentential pause duration,

intersentential pause duration, pitch changes at the end of the sentence, and overall intonation contour) differ depending on their fluency skill levels?

2. Is ORF a unitary construct that encompasses both reading prosody and text reading efficiency or do these two components represent dissociable skills?

3. To what extent do reading prosody features account for reading comprehension?

    3-1. To what extent does reading prosody account for reading comprehension after controlling for decoding skills?

    3-2. Does reading prosody play a mediating role between decoding skills and reading comprehension?

    3-3. To what extent does reading prosody account for reading comprehension after controlling for TRE?

Based on the theoretical and empirical research, the hypotheses for the study are formulated as follows:

Research Question One:

Hypothesis 1: Prior research conducted in the L1 context has demonstrated notable variations in reading prosody features, such as pause frequencies (e.g., Miller & Schwanenflugel, 2008), pause duration (e.g., Binder et al., 2013), overall intonation contour (e.g., Benjamin & Schwanenflugel, 2010), pitch changes at the end of sentences (e.g., Miller & Schwanenflugel, 2006), with respect to readers' proficiency levels. These studies have found that as readers' proficiency, encompassing decoding skill and text reading efficiency, increases, significant differences in these prosodic aspects emerge. Based on the findings of the previous studies, it was hypothesized that in L2 readers, each cohort with different fluency skills would differ in every aspect of their prosodic passage reading.

Research Question Two:

Hypothesis 2: Even though there has been consensus on the definition of ORF as encompassing reading speed, accuracy, and reading prosody (Kuhn et al., 2010), research specifically focused on the ORF construct itself is limited. What is known from a few previous studies utilizing factor analyses on reading prosody (e.g., Benjamin et al., 2013; Kim et al., 2021a) is that pitch- and pause-related factors are related but distinct components. Conversely, a more advanced factor analysis study examining the ORF construct itself found that ORF, consisting of reading speed and accuracy, pitch factors, and pause factors, is a robust construct (Kim et al., 2021b). Therefore, a hypothesis was formulated as follows: the amalgamation of text reading efficiency and reading prosody would form an indivisible construct of oral reading fluency in L2 reading, in accordance with the widely accepted definition of this term.

Research Question Three:

Hypothesis 3-1: Previous studies have provided evidence that reading prosody contributes to the explanation of reading comprehension variance, even when controlling for decoding skill (Binder et al., 2013; Miller & Schwanenflugel, 2008; Sabatini et al., 2019). However, the extent to which reading prosody explains variance in reading comprehension can vary depending on the developmental stage of readers and languages under investigation. Nevertheless, given that the primary function of reading prosody is to facilitate the comprehension of meaning in written text (Frazier et al., 2006; Kuhn et al., 2010; Kim et al., 2021a), the subsequent hypothesis was postulated: in L2 readers, individual differences in prosodic reading of passages would account for unique variance in reading comprehension, even when accounting for differences in decoding skills.

Hypothesis 3-2: Extant research and theories on the role of reading prosody in reading comprehension have speculated that built upon the decoding skill, reading prosody aids reading comprehension by helping to hold information being processed in working memory (Frazier et al., 2006; Kuhn et al., 2010; Schreiber, 1991, Schwanenfluegel et al., 2004). Although the directionality of the relationship between reading comprehension and reading prosody in the L1 reading context remains uncertain (i.e., whether reading prosody is an outcome of reading comprehension or a contributor to reading comprehension), the following hypothesis is proposed: in L2 readers, reading prosody would mediate the relationship between decoding skills and reading comprehension.

Hypothesis 3-3: In line with Hypothesis 3-1, in general, previous studies have shown that reading prosody could predict reading comprehension abilities of L1 learners even beyond what is accounted for by reading speed and accuracy at the text level (Miller & Schwanenflugel, 2006; Benjamin & Schwanenflugel, 2010). Even though several L1 previous studies revealed inconsistent results (e.g., Schwanenflugel et al., 2004), the following hypothesis is proposed based on the general findings in this field: in L2 readers, individual differences in prosodic passage reading will account for unique variance in reading comprehension, even beyond differences in accurate and speedy passage reading.

1.4 Significance of the Study

The current study aims to expand the knowledge base on L2 reading prosody by examining the extent to which reading prosody features vary in relation to decoding and text-reading skills. Moreover, by investigating the association between reading prosody and reading outcomes (text reading efficiency and reading comprehension) in L2, this study could provide valuable insights into the underlying mechanisms of L2 reading. Despite the centrality of reading prosody in the definition of ORF, prior research has mainly focused on reading speed and accuracy, while reading prosody has been relatively understudied. In particular, research on reading prosody in L2 is scarce, and its associations with and contributions to ORF and reading comprehension have yet to be explored. Therefore, this study aims to address these gaps and shed light on the role of reading prosody in various aspects of L2 reading.

Furthermore, the study's use of spectrographic analysis allows for a more nuanced understanding of the role of reading prosody in L2 reading. Using spectrographic analysis instead of L1-based reading fluency scales is significant because it allows for more precise and objective measurements of L2 reading prosody features. This approach enables the identification of unique characteristics and patterns of L2 reading prosody, providing valuable insights into L2 reading prosody development. Furthermore, the use of spectrographic analysis can provide more specific and accurate information to L2 readers, allowing for targeted interventions and support. From this viewpoint, another pedagogical implication would be drawn. According to Mislevy et al. (2017), the validity of an assessment tool can be claimed if it is developed based on the evidence drawn from the observation of the performance of the construct. Hence, an L2 ORF scale based on direct observation of reading prosody features would justify its validity if the scale is developed by examining the interrelationships among the observable variables and their relevance to the ORF construct. In this sense, the current study has the potential to serve as a foundation for the development of an effective L2 ORF assessment tool that can inform instructional practices and ultimately contribute to the enhancement of L2 reading development.

Assessment of ORF has been widely employed in the L1 classroom due to its practicality and usefulness. When it comes to the assessment of ORF in L2, Grabe and Jiang (2013) suggest that measuring various aspects of ORF can provide teachers with a "fine-grained (p.192)" understanding of students' reading development. In this regard, reading prosody, which Kuhn et al. (2010) argue is "at the heart of the development of reading skill (p.239)," should be integrated into the reading assessment in general. In the EFL context where reading skill is mainly evaluated through silent reading, incorporating ORF, particularly reading prosody, into the reading assessment can act as a developmental window for professional educators and teachers to provide timely and suitable reading interventions. The current study can shed light on how reading prosody features vary depending on reading proficiency and how they are linked to reading comprehension, thus providing L2 reading teachers with a framework for effective instruction and valid assessment.

1.5 The Outline of the Thesis

The current dissertation comprises six chapters. The first chapter provides an overview of the study's background and outlines the research problem and questions. Furthermore, it highlights the significance of the study and defines the terms used throughout the research. Chapter 2 presents a comprehensive review of the literature related to theoretical frameworks and empirical studies on ORF, with a specific emphasis on reading prosody. In Chapter 3, the research methodology is explained, including details on the participants, the overall assessment procedures for reading abilities, as well as the specific methods for measuring reading prosody, and the data analysis method. Chapter 4 presents the results of the study, while Chapter 5 delves into a detailed discussion of the findings. The dissertation concludes with Chapter 6, which summarizes the key findings and provides pedagogical implications. Limitations of the study are also discussed, and recommendations for future research are suggested.

Chapter 2. Literature Review

This chapter presents the theoretical underpinnings that informed the present study and provides a comprehensive review of the literature on ORF and prosody. As reading prosody is considered one of the key elements of ORF, the first section outlines the theoretical framework of reading fluency, followed by a review of the literature on reading prosody and the relationship between reading prosody and comprehension.

2.1 Oral Reading Fluency

This section presents a systematic review of the theoretical frameworks and empirical literature on ORF. The section critically evaluates the role of reading prosody in ORF and examines its relationship with other key fluency constructs, such as accuracy and rate.

2.1.1 Theoretical Frameworks of Oral Reading Fluency

The concept of ORF has been elucidated through the lens of information-processing theories. This section provides an overview of three interconnected theoretical frameworks that contribute to the understanding of fluent reading: the automaticity theory, the verbal efficiency theory, and the interactive and compensatory model.

The automaticity model, first proposed by Laberge and Samuel in 1974, is a seminal theory in the field of reading that has significantly influenced the current understanding of reading processes. This model asserts that reading is primarily a process of automatic word recognition, in which fluent reading is attained through the establishment of automaticity in word recognition. The fundamental concept driving the model is that reading is a hierarchical process in which the automaticity of lower-level processes, such as word recognition, enables the allocation of cognitive resources to higher-level processes. The automaticity model posits that once a skill is described as automatic at a macro level, the underlying subskills and their relationships must also be automatic. The initial phase of this process is focused on the visual code and the "unitization (p. 298)" of visual stimuli, which can range from individual letters to whole words or even common phrases. Over time and with practice, the visual features of these stimuli become unitized and are perceived as a single entity. As this process continues and letter perception becomes more automatic, attention to early visual coding processes decreases. This decrease in attentional demand allows cognitive resources to be redirected to other areas, such as the semantic or meaning code, resulting in a greater capacity to comprehend written language with greater speed and accuracy.

While the automaticity model has made significant contributions to the field of reading, it has been subjected to criticism for its limited account of the reading process and its insufficient attention to the complex interplay between word recognition and comprehension. Some scholars contend that the model oversimplifies the intricacies of reading by disregarding the influence of higher-order cognitive processes, such as syntax and semantics, on reading comprehension (Perfetti & Hart, 2002; Seidenberg, 1997). Others have suggested that the model neglects the crucial role of contextual and background knowledge in facilitating word recognition and comprehension (Adams, 1990; Stanovich, 1980).

The verbal efficiency model, proposed by Perfetti in 1985, posits that skilled readers have greater "verbal efficiency" in processing words and language. This efficiency is achieved through the development of larger and more robust lexical entries for frequently encountered words. In other words, skilled readers can quickly and accurately recognize familiar words because they have developed more efficient connections between the visual, phonological, and semantic processing systems in the brain. The model suggests that the ability to recognize words is dependent on two factors: the quality of the representations of the words in the reader's mental lexicon, and the strength of the connections between these representations and other relevant information such as context, background knowledge, and syntactic structures. This allows skilled readers to comprehend the meaning of the text quickly and accurately.

Lastly, Stanovich's interactive-compensatory model, which emerged in 1980, is a conceptual framework that offers insights into the intricate process of reading. The theory posits that reading is a multifaceted process that entails bottom-up and top-down processing. The former involves the decoding of individual words and their meanings, while the latter involves drawing upon prior knowledge and experiences to comprehend and interpret the text. Interactive models posit that information is synthesized from these multiple knowledge sources simultaneously, and the compensatory assumption states that a deficit in one knowledge source leads to increased reliance on other sources, regardless of their position in the processing hierarchy. Therefore, struggling readers with weak context-free word recognition skills benefit from an additional contextual expectancy process to aid word identification. However, this additional facilitation comes at a cost, as the conscious expectancy process consumes attentional capacity, leaving fewer cognitive resources for comprehension operations that integrate larger text units.

Conversely, proficient readers excel at context-free word recognition. Skilled readers have effectively automatized the recognition of words and subword units to a greater extent compared to individuals with reading difficulties. Moreover, they possess better phonetic segmentation and recoding abilities, enabling rapid word decoding even when visual recognition fails. The rapid context-free word recognition proficiency mitigates their dependence on conscious expectancies derived from preceding sentence context, thereby liberating attentional resources for the execution of higher-level integrative comprehension processes.

While the three theories differ in their respective approaches to explaining or emphasizing key aspects of reading, they all converge on the notion that proficient reading is characterized by the ability to process language with both speed and accuracy and to flexibly adapt processing strategies to suit the demands of the text.

Building upon this premise, it has been highlighted that ORF serves as an indicator of overall reading proficiency (Adams, 1990; Fuchs et al., 2001). While reading a text aloud, the reader undergoes a transformation where written letters are converted into phonological representations. This phonological information is then used to access the oral vocabulary stored within the lexical memory. Furthermore, the reader seamlessly integrates both lexical and syntactic information at various levels, encompassing intra- and inter-sentence structures (Rasinski et al., 2012). In essence, when readers engage in fluent oral reading, these cognitive components are rapidly orchestrated in a seemingly effortless and unconscious manner. Therefore, the shared tenets of the information-processing theories highlight the critical role that ORF plays in supporting and advancing proficient reading performance.

2.1.2 The Constructs of Oral Reading Fluency

ORF is typically operationalized as having three components, namely reading rate, reading accuracy, and reading prosody (Kuhn et al., 2010). As previously discussed, there is a general consensus among reading researchers that proficient reading depends heavily on accurate and automatic word identification (Ehri, 1995). Indeed, scholars in the field of fluency widely acknowledge that accurate and automatic word identification is a pivotal aspect of fluent reading and that supporting factors such as phonemic awareness and letter naming play a critical role in cultivating accuracy and fluency over time (Ehri, 1995; Kuhn et al., 2010; Kuhn & Stahl, 2003; Wolf & Katzir-Cohen, 2001).

Reading prosody, on the other hand, is responsible for conveying the refined meanings that readers derive from texts, utilizing components such as intonation, stress, tempo, and appropriate phrasing to bring the text to life (Hudson et al., 2008; Kuhn et al., 2010). Rather than reading a text in a flat and unvarying tone, readers who possess good prosodic skills apply elements of natural language to the material they are reading. In doing so, they move beyond mere word recognition and deepen their comprehension.

While prosody is often treated as a separate construct from accuracy and rate, it is actually closely linked to these dimensions of fluency. For example, research has shown that readers who demonstrate good prosody also tend to have higher accuracy and rate scores (Benjamin & Schwanenflugel, 2010; Miller & Schwanenflugel, 2006; Álvarez-Cañizo et al., 2018, 2020). For instance, in the study conducted by Miller and Schwaneneflugel (2006), it was shown that readers who read inaccurately are likely to read slowly because they pause and stumble over words, thereby spending more time reading.

To date, research on reading prosody has been limited in exploring the dimensionality or internal structure of ORF and its components (Godde et al., 2020; Wolters et al., 2020). Whether text reading efficiency and reading prosody are unitary, or two related yet distinct skills remains largely unknown (Kim et al., 2021b). Kim et al. (2021b) who investigated the dimensionality of ORF by bifactor and trifactor analyses found that it is best described as a multidimensional structure, specifically a trifactor structure composed of a general ORF factor, two local factors (i.e., Ratings and Pause), and three specific factors (i.e., Prosody: Ratings and Pause, Prosody: Pitch, and Text Reading Efficiency). Also, it indicated that the general ORF factor that captures common ability across text reading efficiency and various aspects of reading prosody was found to be the most reliable, and predicted reading comprehension across time points from Grade 1 to Grade 3 of L1 readers (Kim et al., 2021b).

Furthermore, although previous studies have established a relationship between text reading efficiency and reading prosody (Benjamin et al., 2013; Paige et al., 2014), the nature of this relationship may differ depending on the specific aspects of reading prosody examined. For example, pause structure indicators of reading prosody such as ungrammatical pauses are moderately to fairly strongly related to text reading efficiency, while pitch-related prosody indicators have little or weak relations with text reading efficiency (Binder et al., 2013; Kim et al., 2021a, 2021b; Wolters et al., 2020).

Even though previous studies in L1 reading have provided insight into ORF itself, there is a need for further research to deepen our understanding of the multidimensional nature of ORF and its constituent factors, particularly with respect to the interplay between text reading efficiency and reading prosody in the L2 reading context. While some research has explored this phenomenon in the context of L1 reading, comparatively little attention has been paid to the construct of ORF in the L2 reading domain. Hence, a comprehensive investigation of ORF constructs and their interrelatedness is warranted, as it has the potential to shed light on the mechanisms underlying L2 reading development.

2.2 The Nature of Reading Prosody

Prosody, often referred to as the "music of language," is a critical component of phonology that deals with suprasegmental aspects of speech, including stress, pauses, pacing, and pitch (Dowhower, 1991; Schreiber, 1991). The role of prosody in speech differs from its role in reading (Godde, 2020; Guaïtella, 1999). However, it is essential to examine the literature on speech prosody because the understanding of how reading prosody operates is primarily grounded in theories and studies in the field of speaking. Therefore, this section begins by exploring how prosody, originally seen as a speech characteristic, is connected to reading. Next, it examines how reading prosody has been assessed and utilized in research and education. Finally, the section delves into each reading prosody feature studied in reading research.

2.2.1 Connection of Speech Prosody to Reading Prosody

As for how prosody works in speech, theories and empirical studies have posited the simultaneous interplay between prosody and syntactic structures. (Cooper & Paccia-Cooper, 1980; Ferreira, 1993; Goldman-Eisler, 1972; Koriat et al., 2002; Levelt, 1989; Selkirk, 1986). Selkirk (1986) has presented a prominent speech production model which suggests that prosodic structure is a hierarchical system, in which distinct components are arranged in multi-level phrases. The underlying assumption of this prosodic model is that syntactic structure governs prosodic structure. For example, the clause boundary compels the intonational boundary, while the phrase boundary influences the phonological phrase boundary. This mechanism triggers phonological phenomena at the phrase or clause boundary, such as pauses, pitch changes, and phrase lengthening.

While Selkirk's (1986) model assumes the constraining role of syntax on prosody, there have been divergent views on how syntax affects prosodic systems across scholars. For example, some scholars have been in line with Selkirk, suggesting a direct impact of syntactic structure (Cooper & Paccia-Cooper, 1980; Levelt, 1989), whereas others have suggested that prosodic structures are only loosely mapped onto syntactic structure (Ferreira, 1993; Frazier et al., 2006) or syntax alone cannot account for prosodic phenomena (Allbritton et al., 1996). Taking into account the body of literature on speaking prosody, it becomes apparent that the grammatical structure of sentences exerts an influence on the prosodic structure. However, the exact nature and extent of this influence remain a topic of ongoing scholarly discussion.

In the act of reading, individuals tend to apply similar patterns of prosody that are used in everyday speech, as noted by Schreiber (1980, 1987, 1991). Such patterns are internally represented through the process of inner speech, in which readers apply the rhythm and melody that is inherent to prosody, not only in oral language but also in silent reading (Ashby, 2006; Fodor, 2002). Hirotani, Frazier, and Rayner (2006) who investigated phrasing and intonation in reading underscore the critical importance of prosody in the following statement:

> That reading should be parasitic on the mechanism underlying the comprehension of the spoken language should not surprise us. What would be shocking is if the rich structuring provided by the intonational system could simply be set aside during reading (p.439-440).

Furthermore, Goldman-Eisler (1972) argued that reading prosody is near to an ultimate reflection of grammatical structures, and Koriat et al. (2002) claim that reading prosody is “a window to structural processing (p.270).” Koriat et al. (2002), in particular, examined the pause structure in order to provide evidence for the structural precedence hypothesis. In the structural precedence hypothesis, it is assumed that readers extract syntactic information from sentences and use the information to guide semantic analysis for text comprehension. By examining the pause pattern of young adults’ reading, they concluded that reading prosody reflects the outcome of structural analysis, and helps to keep the outcome in working memory while semantic information is processed.

Not only at the syntactic level, but prosody also serves in the semantic process of reading. At the word level, for example, prosody facilitates the semantic processing of words by disambiguating compound nouns from adjective-noun pairs (e.g., BLACK-bird vs. black-BIRD) (Kitzen, 2001; Whalley & Hansen, 2006). Even the meaning of a heteronym such as CON*duct* and *con*DUCT can be distinguished in the sentence by putting stress on the different syllables (Chafe, 1988; Wade-Woolley & Wood, 2006). Above the word level, prosody carries readers’ intention by putting emphasis on important words in the sentence, and it is used to express readers’ emotions by a rise-and-fall pattern (Laver, 1994). Also, depending on the sentence type, readers tend to lower or raise their voices at the end of the sentence. Therefore, it seems theoretically evident that reading prosody is interwoven with different aspects of reading comprehension.

Despite the critical role that prosody plays in reading, it has been a neglected component in reading research until relatively recently. Rather, only the segmental components of reading, namely phonological awareness - the ability to comprehend the sound units of words - have been heavily scrutinized in studies on reading development, particularly within the context of intervention studies for dyslexia over the past few decades (Al Otaiba et al., 2009; Barker et al., 1992; Carlisle & Katz, 2006; Kim et al., 2013; Nagy et al., 2003; Schatschneider et al., 2004; Wagner et al., 1994; Wolf & Bowers, 1999).

So far, the investigation of prosody has yielded two distinct lines of research: prosodic sensitivity and reading prosody. The former area has studied the relationship between prosodic sensitivity and reading. Prosodic sensitivity, also referred to as prosodic awareness (Chan & Wade-Woolley, 2018) refers to the skills and capacities involved in recognizing, analyzing, and comprehending the characteristics of speech prosody. Although there are relatively few studies examining how prosodic sensitivity affects the reading process, converging findings have shown that prosodic awareness contributes to reading skills independently of phonological awareness (Goodman et al., 2010; Goswami et al., 2010; Holliman et al., 2008, 2010; Jarmulowicz et al., 2008; Kim & Petscher, 2016; Whalley & Hansen, 2006). Based on the systematic review of prosodic sensitivity in the reading process, Wade-Woolley and Heggie (2016) have argued that readers' knowledge of suprasegmental features, in addition to phonological awareness, should be included as a predictor in explaining the development of reading skills.

In another line of reading research, scholars have investigated the role of prosody generated in oral reading. This strand of research, which initially emphasized ORF in the areas of reading instruction and assessment, has consistently demonstrated that children who possess accurate and efficient oral reading skills tend to read with "good" prosody (Benjamin et al., 2013; Rasinski et al., 2009; Sabatini et al., 2019; Kuhn et al., 2010). In this research area, reading prosody refers to the prosodic rendering that makes an oral reading sound like real-life speech. As children become more proficient and fluent in reading, their oral reading becomes smoother and more speech-like, progressing from a word-by-word, hesitant manner to a more smooth style (Benjamin et al., 2013; Kuhn et al., 2010; Schreiber, 1991).

To better understand what constitutes good quality of reading prosody, the next section reviews the methods used to measure reading prosody and the prosodic features that have been examined in previous studies.

2.2.2 Assessment of Reading Prosody

In the L1 reading context, fluency rating scales have been developed and broadly used to examine and assess children's reading prosody in research and practice (Thomson & Jarmulowicz, 2016). Among widely utilized rating scales is the National Assessment of Educational Progress (NAEP) Oral Fluency Scale (Pinnell et al., 1995). In Table 2.1, the specific points and their description of the scale are demonstrated.

As a holistic scale, it provides a 4-point scale that distinguishes word-by-word oral reading (e.g., *1-point: Reads primarily word–by–word. Occasional two–word or three–word phrases may occur – but these are infrequent and/or they do not preserve meaningful syntax. Lack expressive interpretation. Reads text excessively slow*) to expressive, meaning-based oral reading (e.g., *4-point: Reads primarily in larger, meaningful phrase groups. Although some regressions, repetitions, and deviations from the text may be present, these do not appear to detract from the overall structure of the story. Preservation of the author's syntax is consistent. Some or most of the story is read with expressive interpretation. Reads at an appropriate rate.*).

Table 2.1

NAEP Oral Reading Fluency Scale (Pinnell et al., 1995)

Rating	Description
4	Reads primarily in larger, meaningful phrase groups. Although some regressions, repetitions, and deviations from the text may be present, these do not appear to detract from the overall structure of the story. Preservation of the author's syntax is consistent. Some or most of the story is read with expressive interpretation. Reads at an appropriate rate.
3	Reads primarily in three–and four–word phrase groups. Some smaller groupings may be present. However, the majority of phrasing seems appropriate and preserves the syntax of the author. Little or no expressive interpretation is present. Reader attempts to read expressively and some of the story is read with expression. Generally reads at an appropriate rate.
2	Reads primarily in two–word phrase groups with some three–and four–word groupings. Some word–by–word reading may be present. Word groupings may seem awkward and unrelated to the larger context of the sentence or passage. A small portion of the text is read with expressive interpretation. Reads significant sections of the text excessively slow or fast.
1	Reads primarily word–by–word. Occasional two–word or three–word phrases may occur – but these are infrequent and/or they do not preserve meaningful syntax. Lack expressive interpretation. Reads text excessively slow.

Pinnell et al. (1995) developed this scale on the basis of the theoretical rationale that fluent reading entails automatic word reading and syntactic parsing, which are prerequisites for successful reading comprehension. In light of the author's intention, the scale seems to assume that a fluent reader should be able to phrase a group of words in a meaning unit and understand what the text is about. The NAEP scale was used and validated in reading research (Smith & Paige, 2019), and scale scores were shown to be related to other indicators of reading skills (Danne et al., 2005; Pinnell et al., 1995; Valencia et al., 2010).

Another popular reading prosody rating scale is the Multidimensional Fluency Scale developed by Zutell and Rasinski (1991). Unlike NAEP Oral Reading Fluency Scale, this is an analytic scale that provides more specific information on the four dimensions of reading prosody (i.e., expression and volume, phrasing, smoothness, and pacing). The dimensions and the description for each score are presented in Table 2.2.

For instance, each construct is assessed on a 4-point scale as follows: *Expression and Volume* from 1-point (*Reads in a quiet voice as if to get words out. The reading does not sound natural like talking to a friend.*) to 4-point (*Reads with varied volume and expression. The reader sounds like they are talking to a friend with their voice matching the interpretation of the passage.*), *Phrasing* from 1-point (*Reads word-by-word in a monotone voice.*) to 4-point (*Reads with good phrasing; adhering to punctuation, stress and intonation.*), *Smoothness* ranging from1-point (*Frequently hesitates while reading, sounds out words and repeats words or phrases. The reader makes multiple attempts to read the same passage.*) to 4-point (*Reads smoothly with some breaks, but self-corrects with difficult words and/or sentence structures.*), and *Pace* from 1-point (*Reads slowly and laboriously.*) to 4 (*Reads at a conversational pace throughout the reading.*).

Table 2.2

Multidimensional Fluency Scale (Rasinski et al., 2009)

Dimension	1	2	3	4
Expression and Volume	Reads in a quiet voice as if to get words out. The reading does not sound natural like talking to a friend.	Reads in a quiet voice. The reading sounds natural in part of the text, but the reader does not always sound like they are talking to a friend.	Reads with volume and expression. However, sometimes the reader slips into expressionless reading and does not sound like they are talking to a friend.	Reads with varied volume and expression. The reader sounds like they are talking to a friend with their voice matching the interpretation of the passage.
Phrasing	Reads word-by-word in a monotone voice.	Reads in two-or three-word phrases, not adhering to punctuation, stress and intonation.	Reads with a mixture of run-ons, mid-sentence pauses for breath, and some choppiness. There is reasonable stress and intonation.	Reads with good phrasing; adhering to punctuation, stress and intonation.
Smoothness	Frequently hesitates while reading, sounds out words and repeats words or phrases. The reader makes multiple attempts to read the same passage.	Reads with extended pauses or hesitations. The reader has many "rough spots."	Reads with occasional breaks in rhythm. The reader has difficulty with specific words and/or sentence structures.	Reads smoothly with some breaks, but self-corrects with difficult words and/or sentence structures.
Pace	Read slowly and laboriously.	Reads moderately slowly.	Reads fast and slow throughout reading.	Reads at a conversational pace throughout the reading.

The authors of the scale made it on the basis of the observation of students' reading and teacher feedback, and the scale has been revised several times so far (Paige et al., 2012; Rasinski et al., 2009). Evidence for the validity of the scale has shown that scale scores can be a predictor of reading comprehension (Kim et al., 2021a, 2021b; Rasinski et al., 2009). For example, Rasinki et al. (2009) reported that the scale scores can account for significant variance in the reading comprehension ability of students at three grade levels. Specifically, the proportion of variance that was jointly accounted for by the scale scores was found to be .402, .432, and .326 for the third, fifth, and seventh grades, respectively.

Lastly, Benjamin et al. (2013) developed a reading fluency scale based on their spectrographic observation of reading prosodic features called the Comprehensive Oral Reading Fluency Scale (CORF). A detailed description of the CORF is presented in Table 2.3.

Table 2.3

Comprehensive Oral Reading Fluency Scale (Benjamin et al., 2013)

Automaticity		Expression			
Rating	WCPM[a]	Rating	Appropriate intonation	Rating	Natural pausing
8	137+	4	Makes noticeable pitch variations throughout to communicate meaning. Makes appropriate and consistent end-of-sentence pitch changes. One or two exceptions may exist.	4	Pauses may be used to convey meaning. Between-sentence pauses are short but natural. Unexpected pauses occur less than once per sentence.
6	107-136	3	Varies pitch appropriately and makes appropriate end-of-sentence pitch changes most of the time. Some flatness may exist, but intonation effectively communicates meaning overall.	3	May have brief unexpected pauses once or twice per sentence, but pauses seem to be used mainly to distinguish phrases and sentences. Longer pauses are rare and only momentarily interrupt the flow of the text.
4	78-106	2	Intonation is frequently flat or does not match the punctuation or meaning/phrasing of the text. Shows appropriate pitch variation on a few sentences but is flat or unnatural on many others. Overall impression is that intonation does not effectively communicate meaning.	2	Frequent pausing within sentences May also have some lengthy pausing between sentences. May pause often between phrases or three-or four-word groupings.
2	< 78	1	Reads with flat or other unnatural intonation throughout. Does not mark sentence boundaries with distinct pitch changes, except occasionally.	1	Reading is broken and effortful with numerous pauses throughout. Reads primarily in groups of one-or two words without pausing.

Note. [a]The authors of the scale state that the WCPM values were derived from the benchmark (Hasbrouck & Tindal, 2017) set by the L1 3rd grade as a reference. Hence, when a test taker is in a different grade level, the alternative standard for WCPM will be applied.

This scale consists of three dimensions, each with four levels. One dimension is *Automaticity* based on words correct per minute, and the other is *Expression* based on two prosodic dimensions: *Appropriate Intonation* and *Natural Pausing*. A child who receives a rating of 4 on the *Appropriate Intonation* subscale is able to "*make noticeable pitch variations throughout the reading to communicate meaning and use appropriate and consistent end-of-sentence pitch changes, with only a few exceptions.*" In contrast, a child who receives a rating of 1 "*Reads with flat or other unnatural intonation throughout; does not mark sentence boundaries with distinct pitch changes, except occasionally.*" On the *Natural Pausing* dimension, a rating of 4 indicates that "*Pauses may be used to convey meaning. Between-sentence pauses are short, but natural. Unexpected pauses occur less than once per sentence on average*" A rating of 1 suggests that "*Reading is broken and effortful with numerous pauses throughout. Reads primarily in groups of 1-2 words without pausing.*"

The scale was designed to be consistent with definitions of fluent reading as quick, accurate, and expressive reading. Furthermore, the developers of this reading fluency scale emphasized their objective to enhance the existing fluency rating scales by addressing their perceived limitations (Benjamin, 2012; Benjamin et al., 2013). Notably, these limitations were attributed to the fact that the preexisting scales lacked grounding in the spectrographic characteristics of early reading. As a result, this new scale was meticulously designed to incorporate and consider such crucial spectrographic elements to provide a more comprehensive and accurate assessment of reading fluency.

These three reading fluency scales were developed based on observation and experimental research on the L1 readers' oral reading and designed to capture various characteristics of fluent and prosodic reading of L1 readers. While L1 oral reading fluency scales are useful and dependable for evaluating reading fluency, they cannot be used to assess L2 reading fluency accurately due to possible differences in reading prosody features between L1 and L2 readers.

To obtain a valid evaluation of L2 reading fluency, a more precise and unbiased approach to evaluating reading prosody, such as utilizing speech wave analysis tools should be employed instead of subjective rating scales. These tools generate spectrograms, which graphically represent sound waveforms and enable the identification of various prosodic features. Researchers have used Praat (Boersma & Weenink, 2023) to examine the association between prosodic variables related to reading proficiency and other reading skills (Benjamin, 2012; Benjamin & Schwanenflugel, 2010; Benjamin et al., 2013; Binder et al., 2013; Cowie et al., 2002; Miller & Schwanenflugel, 2006, 2008; Schwanenflugel & Benjamin, 2016; Álvarez-Cañizo et al., 2015). The spectrographic analysis allows for the measurement of pause duration and ratio in milliseconds and pitch changes across various scales. This tool is also user-friendly and can be programmed for automatic analysis using the Praat scripting language. Even though there is sparse research using this method (Benjamin, 2012; Miller & Schwanenflugel, 2006), seminal works conducted by Schwanenflugel and colleagues pioneered this approach to examine reading prosody factors and have provided insight into this field (Benjamin, 2012; Benjamin & Schwanenflugel, 2010; Benjamin et al., 2013; Binder et al., 2013; Cowie et al., 2002; Miller & Schwanenflugel, 2006, 2008; Schwanenflugel & Benjamin, 2016; Álvarez-Cañizo et al., 2015).

Because the current study was designed on the foundation of the studies contributed by Schwanenflugel and colleagues, the findings of their research need to be reviewed in detail. In the next section, the findings of the previous studies that analyzed the prosodic features spectrographically are presented and in particular, the results relating to each suprasegmental component are reviewed.

2.2.3 Characteristics of Prosodic Features

The present section focuses on each of the reading prosody features, namely pause and pitch, which have been examined through the use of spectrographic analysis. It is worth noting that, despite the importance of reading prosody for comprehension and fluency, only a limited number of studies have employed this methodology to explore these features. Table 2.1 provides an overview of the existing literature, including the language under scrutiny, the grade level of the participants, and variables that have been measured through spectrographic analysis, such as pause, pitch, and intensity. Subsequently, the literature on the aforementioned pause and pitch features is reviewed.

Table 2.4

Features of Reading Prosody Measured in the Previous Studies

Studies	Language	Grade	Spectrographic Analysis					
			Pause			Pitch		Intensity
			Duration	Ungrammati-cal pause	Frequency	Intonation Contour	F_0 Difference[a]	
Arcand et al.(2014)	French(CAN)	2		x	x	x		
Álvarez- Cañizo et al. (2018)	Spanish	3, 5, adults			x		x	x
Álvarez- Cañizo et al. (2020)	Spanish	1, 3			x		x	
Benjamin & Schwanenflugel (2010)	English (US)	2		x	x	x	x	
Benjamin et al. (2013)	English (US)	2		x	x	x	x	
Binder et al. (2013)	English (US)	adults	x	x	x		x	
Cowie et al. (2002)	English (US)	4, 5	x		x	x	x	x
Kang & An (2011)[a]			x				x	
Kim et al.(2021a)	English (US)	1, 3	x	x	x	x	x	

(Continued)

Studies	Language	Grade	Spectrographic Analysis					
			Pause			Pitch		Intensity
			Duration	Ungrammatic -al pause	Pause frequency	Intonation Contour	F_0 Difference[a]	
Kim et al. (2021b)	English (US)	1, 3	x	x	x	x	x	
May (2014)	English (US)	5		x	x	x	x	
Miller & Schwanenflugel (2006)	English (US)	3	x		x	x	x	x
Miller & Schwanenflugel (2008)	English (US)	1, 2	x		x	x	x	
Schwanenflugel et al. (2004)	English (US)	2, 3	x		x	x	x	

Note. [a] It pertains to the fundamental frequency (F_0) variation observed at the sentence-final position during the reading of declarative or interrogative sentences.

2.2.3.1 Pause Features

As investigated in Koriat et al. (2002) study, the pause structure has drawn particular attention in the research on reading prosody. In the studies of L1 readers, Binder et al. (2013) examined the reading prosody of adults of low literacy and especially, attempted to observe their pause patterns more closely. The researchers collected pause data from various positions of sentences such as between words, after adjectival commas, after phrase final commas, after sentence final commas, and after a quotative. The results of the ANOVA analysis indicated that less skilled readers tended to pause longer than skilled readers ($p < .001$). Specifically, adults with low literacy skills paused longer between words, after adjectival commas, and after phrase-final commas compared to skilled readers. When inappropriate pausing within words and sentences was examined, the results show that low-literacy readers had longer and more inappropriate pausal intrusions than skilled ones within a word or a sentence.

In the same vein, Álvarez-Cañizo et al. (2018) investigated how the pause duration varies depending on its position and reading proficiency of Spanish L1 children in the longitudinal design. They reported that differences in reading skills made significant differences in the pause duration made in word, phrase, and sentence boundaries, indicating pauses became shorter with higher graders. Also, the findings drawn from another longitudinal study (Miller &

Schwanenflugel, 2008) were in line with those of Álvarez-Cañizo et al. (2018). They found that intersentential pauses of second graders were shorter than first graders by 127 milliseconds and pauses at commas significantly diminished by 253 milliseconds with higher grade levels. From these results, they claimed that the duration of appropriate pauses occurring at the phrase and sentence boundaries could be an indicator of reading development. However, in a more recent study (Álvarez-Cañizo et al., 2020), the authors found inconsistent results, reporting that the duration of appropriate pauses made in phrase or clause boundaries did not significantly differ across different groups of children, while the pause duration made in the absence of punctuation marks tended to be shorter for higher graders.

Other attempts were made in order to examine the effect of sentence type on pause duration (Miller & Schwanenflugel, 2006; Álvarez-Cañizo et al., 2018). The results pointed out that in general pause duration did not vary with the sentence types. For example, in the study conducted by Miller and Schwanenflugel (2006) where pauses were measured in different types of sentences (e.g., declarative sentences, declarative quotatives, yes-no questions, adjective commas, and phrase final commas), shorter pauses were general characteristic of better readers, regardless of sentence types, with duration decreasing as skill level increased. In a more recent study for Spanish monolingual children, Álvarez-Cañizo et al. (2018) investigated how the pause duration varies with sentence type and length. As stimuli, a narrative text which included short and long declarative, exclamatory, and interrogative sentences and two types of pause duration data such as grammatically appropriate pauses and pauses made at punctuation marks were measured. Regardless of sentence type, the overall pause duration was longer as readers' grades were lower. The results of both studies indicate that the duration of pauses during reading is not significantly influenced by the type of sentence being read. Rather, the reader's level of proficiency appears to be the primary determinant of pause duration.

Other than pause duration, the frequency and ratio of pauses within sentences were also measured and collected to examine reading fluency. In a longitudinal study, Miller and Schwanenflugel (2008) found the number of pauses, not duration, was a highly reliable indicator of developmental prosodic change. When the relationship with other prosodic and reading skills was examined, the pause ratio was essentially and contemporaneously associated with children's decoding skills. The pattern of pauses also plays a role in the development of adult-like intonation. In another study (Benjamin & Schwanenflugel, 2010), the same result was drawn, suggesting pausal frequency is deeply related to other reading skills, such as overall intonation contour, decoding skill, and reading comprehension. Álvarez-Cañizo et al. (2020) also found the same role of inappropriate pausal intrusions as a developmental indicator as that of Miller and Schwanenflugel (2008). When the impact of lexical frequency on the number of pauses was examined in the study, the researchers found that more pauses are made in reading low-frequency words than in high-frequency words, which confirms the role of decoding skills in making pauses.

To encapsulate, the pause structure has drawn particular attention in research on reading prosody, with low-skilled readers having longer and more inappropriate pausal intrusions than skilled readers. Pause duration at the phrase and sentence boundaries can be an indicator of reading development, with shorter pauses being characteristic of better readers. The frequency and ratio of pauses within sentences have also been measured to examine reading fluency, suggesting that the number of pauses is a highly reliable indicator of developmental prosodic change in L1 reading. Therefore, it can be concluded that reading prosody is an important component of reading comprehension, and especially pause patterns seem to be an important index of L1 reading proficiency.

2.2.3.2 Pitch Features

Although the study of the perception and production of intonation in language research has been going on for many years, the mechanisms connecting these two processes are still not fully understood (Wade-Woolley et al., 2022). As the role of pitch in reading ability has garnered increasing interest in recent years, several studies have provided insight into the relationship between intonation and reading comprehension.

One of the earliest studies to observe the connection between proficient reading and the appropriate use of intonation was conducted by Clay and Imalach (1971). They found that skilled readers tend to end declarative sentences with a fall in pitch. Furthermore, several studies have highlighted the role of F_0 variability, which refers to the range of fundamental frequency that a speaker uses, as a key indicator of English proficiency (Backman, 1979; Levelt, 1989; Rhee et al., 2003).

Several studies have consistently suggested that pitch-related factors are significant predictors of reading comprehension and fluency in L1 settings (Benjamin & Schwanenflugel, 2008, 2010; Kim et al., 2021a, 2021b; Miller & Schwanenflugel, 2006, 2008; Wolters et al., 2020; Álvarez-Cañizo et al., 2018, 2020). For instance, Miller and Schwanenflugel (2006, 2008) found that skilled readers were more likely to use falling pitch at the end of declarative sentences, and the magnitude of pitch declination increased with reading skill. Additionally, changes in pitch were a significant predictor of reading comprehension, even after controlling for reading fluency skills.

Benjamin and Schwanenflugel (2010) also found significant differences in F_0 changes at the sentence-final positions among low, middle, and high fluency groups. Furthermore, the participants of the study tended to render more adult-like intonation as text difficulty increases, and the middle- and high-fluency groups were significantly more adult-like than the low-fluency group. Additionally, Cowie et al. (2002) studied the effect of intonation on reading comprehension among children with and without reading difficulties. They found that children with reading difficulties had a harder time comprehending the meaning conveyed by intonation changes in speech.

Similar results have been found when investigating the reading prosody of Spanish L1 readers. Álvarez-Cañizo et al. (2020) found significant differences in final pitch declination in declarative sentences across reading fluency groups among Spanish-speaking children in the third and sixth grades. In reading interrogative sentences, children with less reading comprehension had smaller pitch changes. In line with this, Álvarez-Cañizo et al. (2018) investigated how the development of reading prosody varies with sentence type and length among Spanish-speaking children in the third and fifth grades in comparison with adults. They found that the children's general intonation contour was highly related to reading fluency and comprehension.

Overall, the literature suggests that pitch-related factors, particularly intonation contour and F_0 changes at the sentence-final position, are significant predictors of reading comprehension and fluency in the L1 setting. However, it is unclear whether this holds true for L2 learners of English, such as L2 Korean speakers. According to research by Kang and Rhee (2011), it has been observed that L2 Korean English learners tend to use a flattened and monotonous form of intonation compared to native English speakers. This implies that the F_0 range can be a defining characteristic that distinguishes between native speakers and L2 speakers, thus suggesting that pitch changes might not become a significant predictor of reading fluency and comprehension in the L2 reading study. In addition, another evidence of the overall flat intonation of Korean English learners came from Kang's study (2013) which investigated F_0 range differences among Korean English learners with and without immersion education experiences. The findings indicated that even though F_0 range differences were noticeable compared to other prosodic elements such as speech rate and pause length, it was still challenging to distinguish between Korean L2 speakers with and without immersion education experiences only with intonation contour. This suggests that a narrower F_0 range could be attributed to the influence of the native language (Scherer, 2000). Alternatively, if Korean EFL learners achieve a certain level of high English proficiency, it is possible that pitch changes may become more distinctive and serve as a prominent characteristic during their speech (Backman, 1979; Rhee et al., 2003).

Based on the available literature, it is not entirely clear whether the narrower F_0 range of Korean English learners is primarily influenced by the Korean language or overall English proficiency. However, the existing evidence does suggest that pitch-related features may not be reliable distinguishing characteristics.

In conclusion, while the literature consistently shows that pitch-related factors are significant predictors of reading comprehension and fluency in the L1 setting, more research is needed to determine whether this holds true for L2 learners of English, such as Korean EFL readers. Research should continue to explore the relationship between pitch-related factors and reading ability among L2 learners of English.

2.3 Reading Prosody and Reading Comprehension

This section provides a literature review on the relationship between reading prosody and reading comprehension. The review focuses on seminal empirical studies that explore the contribution of reading prosody to reading comprehension. Additionally, the review examines issues regarding the directionality of the relationship between reading comprehension and reading prosody.

2.3.1 Reading Prosody as a Predictor of Reading Comprehension

Reading comprehension is a complex phenomenon in which a number of knowledge is drawn and executed simultaneously. For successful reading comprehension, readers should activate phonological and semantic knowledge to decode words, access the mental lexicon, parse sentences into propositions and integrate overall information suitable to the discourse context (Kintsch, 1998). As mentioned in section 2.2.1, prosody serves an important role in the meaning-making process by interacting with syntactic and semantic features.

To date, several studies have provided evidence of the contribution of reading prosody to reading comprehension. For instance, Schwanenflugel et al. (2004) conducted a study to examine the relationship between reading prosody, decoding abilities, and reading comprehension. However, the findings revealed a relatively weak association between reading comprehension and prosodic features, indicating that factors related to reading prosody did not account for a significant amount of the variance in reading comprehension. Specifically, all of the prosodic features combined accounted for a mere 3% of the total variance in reading comprehension, and the adultlike F_0 contour (with a weight of .19) was found to have an indirect effect on reading comprehension.

Miller and Schwanenflugel (2006) conducted a study to explore the bidirectional associations between reading prosody, reading fluency, and reading comprehension. In the investigation of the predictive role of reading prosody in relation to reading comprehension, they discovered that of all the factors related to reading prosody, only pitch changes in questions and declarative sentences (i.e., the pitch factor) contributed to an additional amount of variance (ΔR = 6.7%) in reading comprehension beyond what could be accounted for by fluency skills alone. This finding implies that prosody serves to facilitate readers' comprehension abilities.

In their investigation of the impact of text difficulty on reading prosody, Benjamin and Schwanenflugel (2010) found that reading prosody significantly predicted reading fluency, with pause ratio and sentence-final F_0 changes emerging as significant variables. Additionally, when examining the relationship between reading prosody and reading comprehension, the study revealed that reading prosody factors continued to account for the significant variance in reading comprehension (R = 5.5%) even after controlling for reading fluency. The reading prosody variable that emerged as significant in predicting reading comprehension was sentence-final F_0 changes.

In a longitudinal study, Miller and Schwanenflugel (2008) investigated developmental characteristics of reading prosody and its effect on subsequent reading fluency and comprehension. Analyzing data from children in grades 1 to 3 over a three-year period, the researchers employed a path analysis and found that intonation contour, rather than pausal intrusion, was a better predictor of later reading fluency and comprehension. Based on these findings, the authors concluded that the overall pitch contour is the primary indicator of fluent, prosodic oral reading, which is consistent with the results of previous studies (Miller & Schwanenflugel, 2006; Ravid & Mashraki, 2007; Schwanenflugel et al., 2004).

In addition to spectrographic analysis, studies utilizing rating scales have also provided evidence of a relationship between reading prosody and comprehension, albeit with varying degrees of strength (Paige et al., 2014; Sabatini et al., 2019). Paige et al. (2014) employed the Multidimensional Fluency Scale, an analytical tool, to evaluate the prosodic features of ninth-grade students' oral reading. The study revealed a substantial correlation between prosody and comprehension, as evidenced by a correlation coefficient of $r = .71$. Furthermore, the results demonstrated that prosody accounted for 40.4% of the variation in silent reading comprehension, with vocabulary contributing an additional 9.7%.

In contrast to this substantial variance accounted for by reading prosody, Sabatini et al. (2019) discovered that reading prosody had a marginal effect on reading comprehension. The study utilized the NAEP fluency scale, a holistic measure of reading prosody, to assess the correlation between prosody and comprehension in fourth-grade students. The findings showed a moderate correlation coefficient ($r = .59$) between prosody and comprehension. However, when multiple regression was employed to assess the predictive value of various reading fluency measures, including reading rate, accuracy, and prosody, the impact of reading prosody on reading comprehension was found to be marginally significant ($p = .062$).

To summarize, the existing literature on L1 reading prosody suggests that there is a connection between reading comprehension and reading prosody.

However, the extent of this relationship appears to differ based on the method used to measure reading prosody (i.e., spectrographic analysis or rating scales), the developmental stages of reading, and levels of reading proficiency. The inconsistency in the findings across studies highlights the need for further research in this area (Wolters et al., 2020).

2.3.2 The Directionality of the Relationship between Reading Prosody and Reading Comprehension

There have been concerns regarding the directionality of the relationship between reading prosody and reading comprehension (Chafe, 1988; Veenendaal et al., 2016). In other words, the relationship between reading prosody and reading comprehension has been a subject of ongoing inquiry, with the question remaining unresolved as to whether reading prosody serves as a contributor to reading comprehension or, conversely, represents an outcome of reading comprehension.

A longstanding view is that reading prosody facilitates syntactic and semantic processing, thereby assisting reading comprehension (Frazier et al., 2006; Hirotani et al., 2006; Kuhn et al., 2010; Schreiber, 1991; Schwanenflugel et al., 2004). Frazier et al. (2006) posited that prosody provides the basic structure for “holding an auditory sequence in memory (p. 248)” during reading or speaking. Aligned with the aforementioned contentions, Fodor’s (1998, 2002) influential research on syntactic parsing during the process of reading presents a compelling argument that challenges the notion that prosody has a restricted role in the interpretation of syntax, even when reading silently. Through the presentation of concrete instances, Fodor demonstrates how “implicit” prosody has a significant impact on the decision-making process related to syntax.

In the empirical study, Schwanenflugel et al. (2004) supported the unidirectional relationship by investigating two models to explain the relationship between reading comprehension and reading prosody. In the first model, reading prosody acted as a mediator between decoding skills and reading comprehension, whereas the second model incorporated the reverse relationship. The study found that the unidirectional influence of reading prosody on reading comprehension was significant, but there was no evidence of a reversed relationship in which reading comprehension predicted the reading prosody. In a similar vein, Fernandes et al. (2018) conducted a longitudinal study on the development of reading prosody among Portuguese-speaking children, and their findings indicate that the extent to which reading prosody contributes to reading comprehension is partially significant, contingent upon decoding abilities. Moreover, the study revealed a lack of a bidirectional relationship between reading prosody and reading comprehension.

On the other hand, Chafe (1988) suggested that prosody is an epiphenomenon of reading comprehension. To read prosodically means that readers can comprehend what they are reading, thereby making it sound smooth and natural (Stahl & Kuhn, 2002). Thus, the ability to read prosodically encompasses reading comprehension skills as well as efficient decoding skills.

With respect to the effect of reading comprehension on prosodic reading, Klauda and Guthrie (2008) posited that a bidirectional relationship exists between prosody and reading comprehension. They suggested that the ability to comprehend the macrostructure of a text facilitates the appropriate use of prosodic features in reading. Similarly, Ravid and Mashraki (2007) demonstrated that in Hebrew-speaking fourth-grade children, reading prosody plays a crucial role in reading comprehension, and vice versa. Veenendaal et al. (2016) conducted a study investigating the development of reading prosody and also found that the bidirectional model fits the data better than the unidirectional model. Specifically, the bidirectional model indicated that the relationship between reading prosody and reading comprehension is different depending on the grade level. The study revealed that, in addition to decoding efficiency, reading comprehension played a significant role in contributing to the development of text reading prosody from the fourth grade to the fifth grade. Remarkably, the findings of this study revealed that a significant association between reading prosody and reading comprehension only manifested in the higher grades, specifically in the fifth and sixth grades. Thus, it can be inferred that a certain level of proficiency in prosody should be achieved to facilitate comprehension effectively.

The present review highlights the ongoing debate surrounding the relationship between reading comprehension and reading prosody, suggesting a complex and multifaceted connection between these two variables. While previous research has attempted to shed light on this relationship, the synthesis of the existing literature reveals several gaps in our understanding, particularly in the context of L1 reading comprehension. Despite recent advances in the field, little is known about how reading prosody operates in reading comprehension, thus underscoring the need for further research into the underlying mechanisms that drive this association, both in L1 and L2 contexts.

Chapter 3. Methodology

The aim of the current study is to investigate characteristics of L2 reading prosody varying according to the L2 fluency skills, and the contribution of L2 reading prosody to L2 ORF and L2 reading comprehension. This chapter provides a detailed description of the participants, reading materials, and assessment procedures utilized in the study. Additionally, the chapter outlines the overall methodology used to measure and process the reading prosody features. Data analysis procedures are also presented towards the end of this chapter.

3.1 Participants

Before recruiting participants, ethical approval for the study was obtained from the Seoul National University Institutional Review Board. Participants were recruited using online participant recruitment and paper flyers placed in each high school[3].

Through the recruitment process, 121 Korean EFL high school third graders were selected for this study. The participants were high-school third

---

[3] In return, all participants received a small gift of about 3000 won. Prior to beginning the study, all participants were provided with a participant information sheet and asked to sign a consent form (Appendix A).

graders in Gyeonggi province, Korea, who had received formal education for about ten years. Unlike prior research primarily focused on L1 young readers' reading prosody characteristics (Benjamin & Schwanenflugel, 2010; Kim et al., 2021a; Miller & Schwanenflugel, 2006, 2008), the relatively higher-grade level was chosen considering the participants' decoding efficiency. Previous studies have claimed that efficient decoding is a prerequisite skill to be attained in order to measure participants' reading prosody (Benjamin & Schwanenflugel, 2008, 2010; Kim et al., 2021a; May, 2014; Miller & Schwanenflugel, 2006, 2008; Schwanenflugel et al., 2004). Otherwise, inefficient word reading could be confounded with prosodic features, making it difficult to extract readers' prosodic characteristics. Consequently, to ensure that participants exhibited a requisite level of word reading accuracy and efficiency, the higher-grader L2 learners were judged as suitable candidates for the study.

In order to select the participants who have enough decoding skills to read the target passages for measuring ORF and reading prosody, decoding skill measurement was implemented in advance and the scores were used as the cutoff standard. The detailed procedure for measuring decoding skills is presented in the following section 3.2.2.

In the prior study conducted by Miller and Schwanenflugel (2008), they used the standard score drawn from the Test of Word Reading Efficiency (TOWRE) sight word test, and participants whose scores are between the 25th and 99th percentiles were included in the study. Unlike the previous study, the current study used students' raw scores as a cutoff score because the Woodcock Reading Mastery Test (WRMT) III does not provide standard scores for L2 readers. Therefore, through the descriptive analysis of the word reading assessment, the outliers were excluded from the study and those who distributed between the 25th and 99th percentiles participated in the following experiments. As a result, the data of 90 students' word reading scores were entered into the analysis, and the selected students proceeded to participate in other reading assessments.

The participants in the final entry were attending a high school in Ansan (n=68), Anyang (n=7), and Seongnam (n=15) cities in Gyeonggi province. All of the participants were aged between 17 and 18, and more than half of the students (n=50) were female, and the others were male (n=40).

Among the participants, only two students had been to English-speaking countries to learn English. One student had resided in an English-speaking country for 7 years and attended Kindergarten and elementary school. The other student received education in a secondary school for three years. These participants were included in the study mainly because they were within the normal distribution.

3.2 Reading Assessment and Procedures

In this section, the assessment procedures used in the current study are outlined. First, the general assessment procedure is presented. Then, the measurement procedures for decoding skill, text reading efficiency (i.e., WCPM), and reading comprehension are described.

3.2.1 General Assessment Procedure

The data collection process was carried out over a period of four months, from April to July 2022. The assessment included measures of decoding skill, text reading efficiency, and prosodic reading, reading comprehension.

For the measurement of decoding skills, text reading efficiency, and prosodic reading, the participants were asked to read a series of words and texts aloud, which took approximately 30 minutes. They were assessed individually in a quiet room at their schools. To ensure accuracy and consistency, the assessments were conducted by the researcher of this study. The participants' oral reading was recorded using a ZOOM recorder model H1N for later analysis.

To measure the participants' reading comprehension, the reading sections of the Mock College Scholastic Ability Test (MCSAT) were utilized. These tests were specifically designed by local education offices and the Korean Institute for Curriculum and Evaluation (KICE) to evaluate the reading comprehension abilities of Korean high school students. The tests were administered in a group setting on March 24, April 13, and June 9, with each school's teachers acting as proctors. After the tests, the participants were given formal test reports, and their reading scores were obtained with their consent.

3.2.2 Decoding Skill Measurement

To measure the decoding skills of the participants, the Woodcock Reading Mastery Test III - Word Identification subtest was utilized. This test is comprised of 46 words that increase in difficulty. Participants were asked to read the words aloud on a one-on-one basis, and a point was given only if a whole word was pronounced correctly. Incorrect syllable pronunciation, disjointed syllables, or no response resulted in no points being granted (Binder et al., 2013). The test continued until the participants could not pronounce four consecutive words correctly. The number of correctly read words was counted to calculate the word reading score. The raw scores from the test were used as an indicator of decoding skills, and 90 participants were selected according to their scores, as mentioned in Section 3.1. To ensure interrater reliability, two raters graded the test scores, with the primary rater being the researcher of this study and the second rater being an English teacher with a master's degree in phonetics. The Cohen's Kappa value for decoding skills was 0.819 ($p < .05$), indicating strong agreement between the raters.

3.2.3 Text Reading Efficiency Measurement

The Woodcock Reading Mastery Test III – Reading Fluency subtest was used to measure the participants' text reading efficiency (TRE). Two passages from Grades 5 and 6 with Flesch-Kincaid Grade Level (FKGL) readability indices of 4.5 and 4.3 respectively were utilized for this purpose. The Lexile indices for these texts ranged between 610 to 800. This text difficulty was chosen based on the results of the previous study (Ryu & Lee, 2021), suggesting that these levels of the text would be not too easy or too difficult for Korean high school students. Prior to reading the passages, participants were informed that comprehension questions would follow their reading and they should answer the questions without seeing the text. The aim of including comprehension questions was to evaluate participants' reading fluency while attempting to comprehend the text. To familiarize participants with the passages, brief introductions were provided in Korean by the researchers before they began reading as follows:

> *You are about to read a short passage about the insect. In the text, the facts and information about the insect are presented, like the number of species and the ways they perceive their surroundings and socialize with others. While you are reading aloud, you should try to understand what the text is about. After you finish the reading, I will ask several questions about the text to check your comprehension. Also, you don't have to read as fast as you can. I would like you to read aloud as you normally would if you were reading it to your little siblings or friends.*

Following the assessment, the participants' reading rate was quantified in terms of words read correctly per minute (WCPM), which was computed by multiplying the total number of words by 60, and then dividing the product by the total reading time in seconds. The average WCPM obtained from the two passages was utilized in the subsequent data analysis. The consistency of the WCPM measurement was established through interrater reliability analysis using Cohen's Kappa coefficient ($k = 0.867$, $p < .05$).

3.2.4 Reading Comprehension Assessment

In order to assess participants' reading comprehension, three separate English reading tests were administered. Specifically, participants were given subsets of the MCSAT on three different occasions: March 24, April 13, and June 9. These subsets were developed by the office of Seoul and Gyeonggi education as well as the Korean Institute for Curriculum and Evaluation (KICE), and contained 28 test items intended to measure a wide range of L2 reading skills, including identifying main ideas, extracting factual information, and making inferences. The items were scored using a two or three-point scale based on item difficulty, with a total possible score of 65 points. The reading section of each test lasted approximately 37-38 minutes and included various text genres, such as correspondence, fiction, articles, and expository texts. The mean score of the three reading comprehension assessments was used as a variable for the participants' reading comprehension ability.

When the readability analyses were conducted for the texts included in the test subsets, it was revealed that the first and second test subsets had similar levels of readability, whereas the texts in the third subset were found to be more readable. Despite this difference in readability, the means of the participant's scores on the three tests were similar (Test 1, *M=42.424, SD= .944*; Test 2, *M=42.408, SD= 1.09*; Test 3, *M=41.12, SD= 1.09*). Therefore, the three test scores obtained from the reading assessments were averaged and then included in the subsequent analyses. Table 3.1 presents the results of the FKGL readability analysis conducted on the texts used in the tests.

Table 3.1

A Flesch-Kincaid Grade Levels of MCSAT RC Passages

Item Number	Test 1 (March, 24th)	Test 2 (April, 13th)	Test 3 (June, 9th)
18	6.5	10	4.5
19	5.6	4.9	3.7
20	11.8	8	3.7
21	18	13.4	5.3
22	9	13.4	6.3
23	13.8	16.3	8.8
24	10	11.3	7.6
25	9.8	9.5	6.4
26	9.1	13	7
27	3.9	6.6	5.2
28	7.4	3.4	2.6
29	14.3	15.6	8.4
30	12.2	12.8	10.5
31	11.1	13.4	9.4
32	11.3	6.9	7.6
33	13.9	10	8.7
34	11.3	12.3	7.6
35	14	16.7	8.7
36	9.6	12.3	8
37	10.2	15.4	5
38	14.3	17.3	9.5
39	13.8	12	7
40	11.5	9.4	6.3
41	13.6	14.3	6.3
42	4.4	5.1	3.8
Mean	10.51	11.05	6.69

Note. MCSAT = Mock College Scholastic Ability Test; RC = Reading Comprehension; These tests were developed by the education office of Seoul and Gyeonggi and KICE in order.

3.3 Reading Prosody Measurement

To assess the participants' L2 reading prosody, two narrative passages which were adapted from a prior study on L1 reading prosody (May, 2014) were employed (Appendix B). Previous studies suggested that texts should be highly decodable but have challenging structures for comprehension (Benjamin & Schwanenflugel, 2010; May, 2014). Decodability was considered an essential factor in measuring prosody since calculating the prosody of a reading error is not meaningful. When the preliminary analysis was conducted, the results revealed that an overall accuracy rate of text reading was about 99%, which is suitable for prosody assessment. Furthermore, the readability levels of the passages were determined using the FKGL readability index, with the first passage having an index of 4.9 and the second passage having an index of 6.7. These levels were deemed appropriately challenging for Korean high school students based on the findings of a previous study (Ryu & Lee, 2021). Therefore, these texts were considered appropriate for extracting and evaluating the prosodic features in the participants' reading in English. Also, before participants read aloud the text, they were told brief explanations about the passages as in the text reading efficiency assessment.

In light of the extensive effort required for sound analyses, this study followed the precedent of prior research on L1 reading prosody by analyzing three sentences from each passage (Benjamin, 2012; Benjamin &

Schwanenflugel, 2010; Benjamin et al., 2013; May, 2014; Miller & Schwanenflugel, 2006, 2008; Schwanenflugel et al., 2004). The chosen sentences are highlighted in Appendix B. Prosodic features were extracted and analyzed using Praat (Boersma & Weenink, 2023), while the basic measurement procedure replicated that of prior studies (Benjamin, 2012; Benjamin & Schwanenflugel, 2010; Benjamin et al., 2013; May, 2014). To ensure the reliability of the measurements and analyses, randomly chosen 100 sentences were evaluated by the second rater. These values are reported in each section of the prosody assessment.

In accordance with previous studies (refer to Table 2.1), the selected variables for this study include the ratio of pauses within a sentence, the ratio of ungrammatical pausing, the duration of intrasentential pauses, the duration of intersentential pauses, fundamental frequency (F_0) change at the ends of sentences, and intonation contour. While previous L1 reading prosody studies have examined three or four variables at most, the current study investigates all the variables that have been suggested to be linked to reading fluency or comprehension in earlier research.

3.3.1 Intrasentential Pausal Ratio/Duration

Most of the previous studies which investigated reading prosody in L1 reading have reported that pause factors (e.g., frequency, ratio, and duration) were closely associated with reading fluency and comprehension. As seen in Table 2.1, pause factors were assessed in most of the studies and their significance was confirmed. Hence, the frequencies and duration of pauses were selected as prosodic variables in this study.

Consistent with the methodology of prior studies (Benjamin, 2012; Benjamin & Schwanenflugel, 2010; Benjamin et al., 2013; Cowie et al., 2002; Dowhower, 1991; Kang & Johnson, 2018; Miller & Schwanenflugel, 2006, 2008; Álvarez-Cañizo et al., 2018), any temporal gaps exceeding 100ms or a flat waveform lasting more than 100ms on the spectrograph were regarded as pauses. When judging pauses, the current study adopted the guidelines from Benjamin's previous research on reading prosody (2012):

1. Pausing is defined as any spectrographic flat waveform lasting longer than 100 milliseconds.
2. Respiratory events such as breaths are included in the definition of pausing.
3. A significant reduction in intensity or the absence of clear formant markers are reliable indicators of pausing.
4. Hesitation and pre-articulation, which refer to delays in the production of speech sounds, are considered to be types of pausing in speech analysis.
5. The omission of words is not considered a type of pausing.
6. If a participant produces a phrase or a word without error and subsequently pauses and repeats it, the interval between the end of the initial phrase or word and the start of the repetition is considered a pause, provided that it lasts longer than 100 milliseconds. Similarly, a pause lasting over 100 milliseconds between the end of the phrase and the start of the next phrase is also considered a pause.

To account for the variation in the number of words across the six sentences of the two passages, a pause proportion per passage was calculated instead of pause frequencies. The total number of pauses within sentences or words for each participant was counted and then divided by the total number of spaces between words. The resulting ratio for each sentence was then averaged across the six sentences to obtain a mean pausal ratio per participant.

The duration of pauses was calculated by measuring the interval of each pausal intrusion, adding up the intervals, and dividing the sum by the total number of sentences. These pause ratio and duration variables are referred to as IntraP_ratio and IntraP_duration in the present study.

The preliminary analysis was conducted with a sample of 25 cases, and discussions between the primary and the secondary raters were held in video conferences to address any pertinent issues. Afterward, the secondary rater assessed 100 sentences, and the results were compared to those of the primary rater. To determine the level of agreement between the ratings, Cohen's Kappa was used, indicating substantial agreement between the raters ($k = 0.808$, $p < .05$).

3.3.2 Ungrammatical Pause Ratio

In addition to pausal intrusions, the ungrammatical pause ratio (i.e., UGP_ratio) was measured. These pauses were defined as pauses made by participants in places that did not correspond to syntactic boundaries or other points in the sentence required for meaning or emphasis (Benjamin, 2012;

Benjamin & Schwanenflugel, 2010; May, 2014). Any pauses occurring within a word or after reading errors were also considered ungrammatical. For example, in the sentence “*She could not find a pineapple*,” any pauses within the sentence were considered ungrammatical as in the previous studies (Benjamin, 2012; Benjamin & Schwanenflugel, 2010; May, 2014). In contrast, in a sentence like “*Kate went to the grocery store, but she could not find a pineapple*,” a pause between “*store*” and “*but*” would be deemed acceptable because the juncture between the clauses, as well as the presence of a comma, indicate a break in the sentence. Figure 3.1 presents a visual representation of the pauses observed on the spectrogram of Praat.

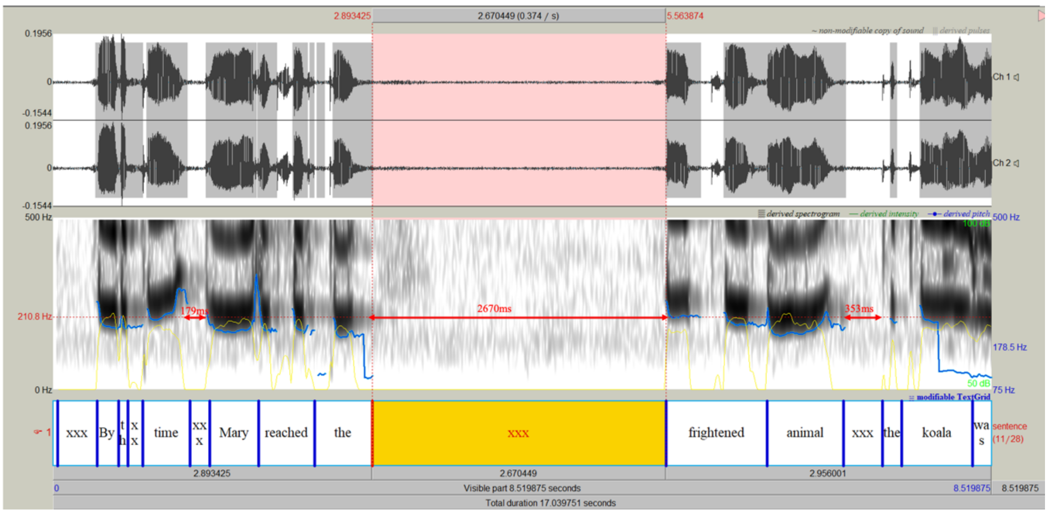


Figure 3.1 Spectrographic Representation of Measurement of Pause Features

In the figure, while reading the sentence "*By the time Mary reached the frightened animal, the koala was…*", the participant seemed to pause at four specific points: between "*the*" and "*time*", "*time*" and "*Mary*", "the" and "*frightened*", and finally, "*animal*" and "*the*" However, the first pause between "*the*" and "*time*" did not qualify as a pause since the time interval between these two words was less than 100ms. The remaining pauses had durations of 179ms, 2670ms, and 353ms, respectively, and were therefore considered as pauses. Among these pauses, the pauses between "*time*" and "*Mary*", as well as between "*the*" and "*frightened*", were deemed ungrammatical pauses since they occurred at positions other than phrasal boundaries.

To calculate the UGP_ratio, the occurrences of ungrammatical pauses were tallied and subsequently divided by the total number of spaces present in the sentence. The resulting mean value, obtained by averaging across the number of sentences, was employed as the designated UGP_ratio index.

Following the assessment of 100 sentences by the second rater, a comparison was made between the results obtained by the primary and secondary raters. Cohen's Kappa was utilized to assess the level of agreement between the two sets of ratings, which revealed a significant level of agreement ($k = 0.858$, $p < .05$).

3.3.3 Intersentential Pause Duration

Punctuations such as question marks, exclamation marks, commas, and periods are known to guide readers to modulate their pitch and pause while reading. In line with previous studies, the duration of temporal gaps in *ms* between sentences was measured as an index of reading prosody. These temporal gaps were identified as flat waveforms on the spectrograph between the end of a sentence and the beginning of the next sentence. The temporal intervals between sentences were aggregated and subsequently divided by the total number of sentences, yielding the computed InterP_duration. Cohen's Kappa was utilized to assess the level of agreement between the two sets of ratings, which revealed a significant level of agreement ($k = 0.820$, $p < .05$).

3.3.4 Pitch Changes at the Sentence-final Position

Previous studies have reported a positive correlation between reading proficiency and the magnitude of the drop in fundamental frequency (F_0) at the end of a declarative sentence (Benjamin & Schwanenflugel, 2010; Benjamin et al., 2013; Miller & Schwanenflugel, 2008; Schwanenflugel et al., 2004). This measurement was obtained by subtracting the last F_0 from the last pitch peak, which was typically measured in Hertz.

In contrast to previous studies, the current study utilizes a different scale for measuring F_0 changes. Previous studies used the Hertz scale to measure pitch changes and reported that pitch changes are a significant feature. However, the literature on phonetics suggests that when measured in Hertz, pitch movements tend to be larger for female voices, as their pitch range is generally higher and wider than that of male voices (Graddol, 1986; Hermes & van Gestel, 1991). Consequently, Hirst and Looze (2021) caution against using the Hertz scale to measure a speaker's pitch range or pitch movements, recommending instead the use of logarithmic scales such as semitones or octaves. In addition, studies using the semitone scale found no significant cross-gender difference in F_0 ranges or modulations (Henton, 1989).

To date, the previous studies on L1 reading prosody have primarily focused on children aged seven to twelve. In these studies, the pitch changes measured in Hertz might have been found to be valid variables as the participants were young children who had not undergone voice changes. However, in the current study, the participants were 12th graders who were more or less young adults. Therefore, to measure F_0 changes in the current study, semitones (re 100 Hertz)[4] were employed instead of measuring in Hertz.

In addition, to identify and extract the vocal nucleus from the last word and the last F_0, a script developed by de Jong and Wempe (2009) was utilized. The measurement of pitch changes is illustrated spectrographically in Figure 3.2, where the word "appearances" is depicted as comprising four syllables with their

---

[4] Semitones can be calculated using the following formulas and the reference pitch is 100Hz in this study. Semitones = $\log_2$(frequency/reference pitch) *12

respective peaks marked in the first tier below the spectrogram. Since the vocalic nucleus of the word was located at the second syllable [pɪ], the mean F_0 was measured at that position. Subsequently, another F_0 was measured at the last point of the intonation line of the syllable [səs].

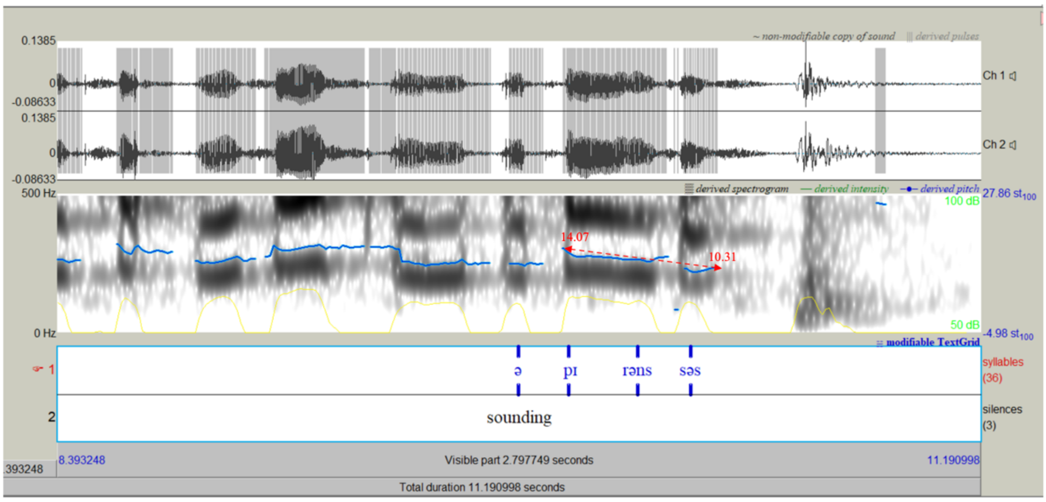


Figure 3.2 Spectrographic Representation of Measurement of Sentence-final Pitch Declination

When measuring F_0 changes, guidelines outlined by Benjamin (2012) were taken into account:

1. The pitch peak of a sentence is defined as the vocalic nucleus of the final word, for which the mean pitch is measured.
2. To determine the final pitch of a sentence, the pitch at the termination of the last word is measured.
3. Extraneous pitch effects arising from the phonetic properties of word-final segments, commonly known as “tails,” are disregarded, and measurements are taken at the natural conclusion of the word.

4. Creaky voice[5] is not included because it can cause the distortion of the data (Yuasa, 2010).
5. Any speech omissions are considered missing data.
6. While incorrect word selection or mispronunciations do not necessarily render the data unusable, in the event of self-correction or repetition by the participant, only the final iteration or correction is considered in the analysis.

To encapsulate, the sentence-final pitch declination for declarative sentences, i.e., Pitch_SF, was calculated by taking the mean F_0 difference between the last peak and the last F_0 in semitones, which was then averaged across all sentences to yield an index.

Prior to measuring the entire dataset, the primary and the secondary raters conducted the preliminary measurement on 25 cases, and subsequent discussions were held in video conferences to address relevant issues. Following this, the secondary rater measured 100 sentences, and the results were compared with those of the primary rater. Cohen's Kappa was used to calculate the level of agreement, which indicated substantial agreement between the ratings (k = 0.726, $p < .05$).

3.3.5 Overall Intonation Contour

Previous research has shown that the general intonation contour observed in children's oral reading can serve as a valid indicator of fluent reading and

---

[5] Creaky voice refers to a vocal effect that occurs when the vocal cords vibrate at a very low frequency at the end of their range, resulting in a sound that is often described as rough or gravelly.

comprehension (Benjamin & Schwanenflugel, 2010; Benjamin et al., 2013; Coots & Snow, 1981; Dowhower, 1987, 1991; Miller & Schwanenflugel, 2006, 2008; Proctor et al., 2005). Early investigations of reading prosody involved comparing children's pitch swing to that of adults' intonation. Subsequently, Benjamin et al. (2013) adopted a novel approach from other research (Bolanos et al., 2013; Cowie et al., 2002), which entailed measuring the variation in pitch changes during oral reading. The current study employed this measurement method. Specifically, the mean and standard deviation of the pitch (in semitones re 100 Hz) for word or sentence stress were determined for each sentence, after which the standard deviations were averaged across sentences. The mean of the standard deviations of the intonation contour was referred to as Pitch_SD. The extraction of intonation contours from sentences adhered to the guidelines established by Benjamin (2012):

1. Sonorant consonants may be incorporated when the syllable boundary is indistinct.
2. Omitted responses are excluded from the analysis.
3. Errors in word choice or pronunciation are not necessarily grounds for exclusion from the data.
4. When assessing speech repetition accuracy, the final iteration should be used.
5. Creaky voice is excluded from the data.

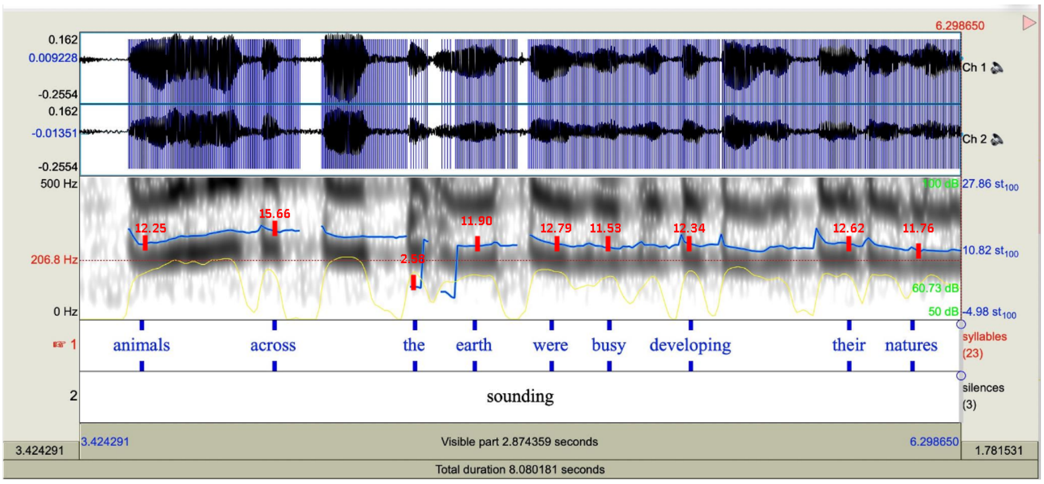


Figure 3.3 Spectrographic Representation of Measurement of Pitch_SD

To identify the peak of each syllable, the researcher utilized a script developed by de Jong and Wempe (2009). Despite the effectiveness of the script in identifying and predicting syllable nuclei, manual adjustments to the position of peak nuclei were sometimes required, particularly because the participants tended to show vowel epenthesis (Shin & Iverson, 2011). Consequently, upon running the script, two raters conducted a meticulous examination of the spectrogram and performed manual adjustments to the peak positions whenever deemed necessary.

Figure 3.3 displays the graphical representation of the pitch contour and the corresponding peak points for each word using Praat software. During the reading task, when a participant reached this section of the sentence "...*animals across the earth were busy developing their natures*..." the pitch values in semitones were extracted from each word. The extracted pitch values ranged from 12.25*st* for the word "*animal*" to 11.76*st* for "*natures*," as shown in the figure. Subsequently, the standard deviation of these pitch values in a sentence was calculated as a measure of intonation variation.

Furthermore, following a series of discussions to deal with measurement procedures and pertinent problems, the two raters measured a total of 100 sentences. Upon calculating interrater reliability using Cohen's Kappa coefficient, the results revealed a moderate level of agreement ($k = 0.752$, $p < .05$) between the two raters.

3.4 Statistical Analysis

The data collected from the study were analyzed using various statistical methods. In order to examine the first research question, principal component analysis (PCA) was used to generate scores for *fluency skill* and a series of one-way ANOVAs were conducted. For the second research question, confirmatory factor analyses (CFA) were conducted. Lastly, a couple of hierarchical linear regressions, PCA, and a mediation analysis were carried out. Detailed descriptions and rationales for modeling are presented in this section.

3.4.1 Statistical Analysis for RQ 1

The primary research question of this study is to explore the variations in prosody features among different groups based on their fluency skills. Prior to conducting ANOVAs, participants were divided into three groups based on their level of reading fluency skills.

To generate a score indicative of L2 readers' fluency skills, a principal component analysis (PCA) varimax rotation was performed. Two variables, namely decoding skills and TRE, were entered into the analysis, in accordance with previous research (Miller & Schwanenflugel, 2006). This factor analysis yielded a good single fit, accounting for 78.2 % of the variance (factor loadings > .884). The resulting factor score[6] was then used as a general indicator of fluency skill.

Based on the tertile of these scores, 30 participants were assigned to each of the three fluency skill groups: Higher, Intermediate, and Lower. To ensure that there was heterogeneity of reading proficiency across the three groups, one-way ANOVAs were conducted for fluency skill and reading comprehension scores. The results indicated that there were significant differences in both fluency skill scores and reading comprehension scores among the groups [fluency skill, $F(2,87) = 167.193, p = .000;$ reading comprehension, $F(2,87) =$

[6] Factor scores are values generated by PCA that represent the degree to which each observation (e.g., TRE and decoding skill) contributes to the underlying factors or components. After PCA identifies the underlying factors, it assigns factor loadings to each variable that indicate how strongly they contribute to each factor. In *SPSS*, factor scores such as regression (default), Barlett, and Anderson-Rubin can be obtained by clicking "Scores" in the Factor Analysis window.

13.282, p = .000]. Table 3.2 displays the results of the post-hoc *Bonferroni* test on fluency skill and reading comprehension.

As seen in Table 3.2, the results of post-hoc tests indicated that the three groups are distinctively different in fluency skills. However, in reading comprehension, the notable differences were found only between the Higher group and the other groups, while the Intermediate and Lower groups were just marginally different (p = .54) [Reading comprehension, M_L= 30.449, *SE*= 10.953; M_I= 37.666, *SE*= 9.502; M_H= 45.600, *SE*= 13.754]. Even though Intermediate and Lower groups were marginally different in reading comprehension abilities, the three groups were deemed heterogeneous in reading proficiency.

Table 3.2

Comparisons of Three Groups on Reading Comprehension and Fluency Skill

Variables	Group (I)	Group (J)	Mean diff.(I-J)	*SE*	*p*
Reading Comprehension[a]	L	I	-7.218	2.993	.054†
		H	-15.152	2.939	.000***
	I	H	-7.933	3.039	.032*
Fluency skill[b]	L	I	-1.105	.121	.000***
		H	-2.120	.119	.000***
	I	H	-1.015	.123	.000***

Note. †p < .10; * p < .05; *** p < .001; L=Lower; I=Intermediate; H=Higher; Adjustment for multiple comparisons: *Bonferroni*
[a]Reading comprehension scores are the mean score of the three MCSAT scores
[b] Fluency skill scores are the factor scores of PCA based on decoding skill and TRE

After forming fluency skill groups, a series of one-way ANOVAs were utilized to evaluate whether significant differences existed in prosodic features (DV) based on participants' reading fluency skills (IV). A series of one-way ANOVAs instead of a MANOVA approach was chosen because this study aimed to identify the characteristics of each prosodic feature affected by reading fluency (Benjamin & Schwanenflugel, 2010; Jaccard & Guilamo-Ramos, 2002). To reduce the overall risk of Type I error, *Bonferroni* adjustment was used as a post-hoc analysis.

3.4.2 Statistical Analysis for RQ 2

The second research question aims to explore the relationship between factors contributing to ORF, which is widely considered a multidimensional construct according to existing theories and empirical studies (Kim et al., 2021a, 2021b). The dimensionality of the reading prosody and TRE indicators were assessed using confirmatory factor analysis (CFA). In case of reading prosody, due to high correlations (r = .97) between intrasentential pause frequencies (IntraP_ratio) and ungrammatical pause frequencies (UGP_ratio), UGP_ratio was excluded, resulting in five indicators for the spectrograph data: overall intonation contour (Pitch_SD), sentence-final F_0 change (Pitch_SF), intrasentential pause ratio (IntraP_ratio), intrasentential pause duration (IntraP_duration) and intersentential pause durations (InterP_duration).

Based on theoretical considerations and previous research findings, two alternative CFA models were evaluated to investigate the dimensionality of the reading prosody and text reading efficiency data. The first model was the unitary construct model, where one factor was indicated by the five variables representing reading prosody and TRE as depicted in the definition of ORF (Kuhn et al., 2010) (see Figure 3.4). The second model (see Figure 3.5) was hypothesized to consist of two dissociable factors, one representing PITCH (i.e., Pitch_SF and Pitch_SD) and the second factor representing TRE and pause variables (i.e., TRE & PAUSE).

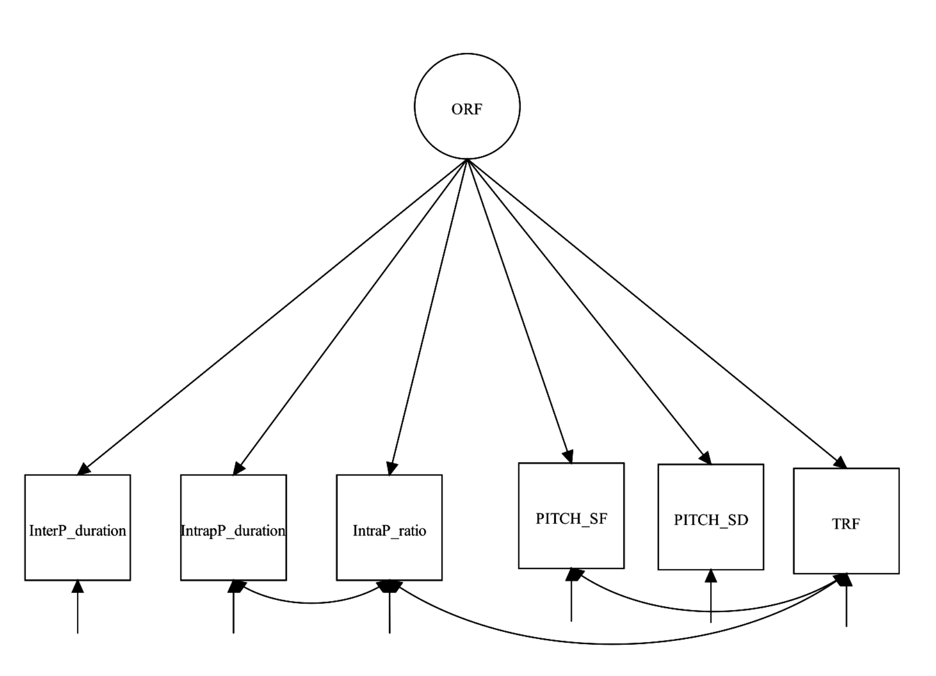


Figure 3.4 Unitary-Construct Model of ORF

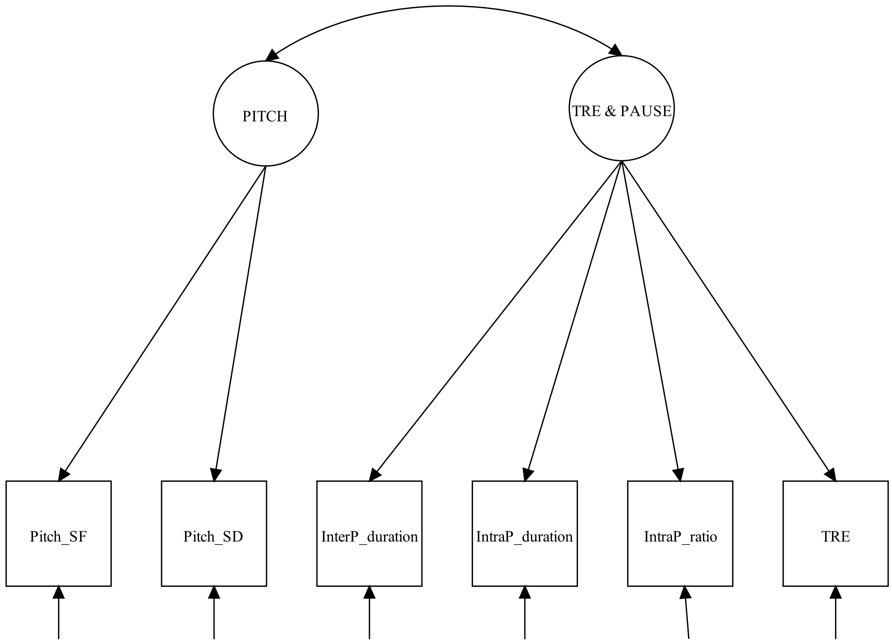


Figure 3.5 Two-Dissociable-Constructs Model of ORF

The rationale for hypothesizing the second model was based on prior research. To date, previous studies have provided evidence that there exists a relationship between TRE and reading prosody (Benjamin et al., 2013; Paige et al., 2014). However, this relationship may differ based on the specific aspects of reading prosody. It seems that pause structure indicators, such as pauses frequencies and durations are moderate to fairly strongly related to TRE (Binder et al., 2013), whereas pitch-related prosody indicators, such as sentence-final pitch changes and overall intonation contour, have little or weak relations with TRE (Binder et al., 2013; Miller & Schwanenflugel, 2008). Consequently, the model was hypothesized and evaluated on the data, wherein one factor (i.e., PITCH) was considered alongside another factor comprising TRE and pause-related features.

To test the model fit, a chi-square test of model fit and four descriptive fit indices were utilized: the root-mean-square error of approximation (RMSEA); the comparative fit index (CFI); the Tucker-Lewis index (TLI); and Standardized Root Mean Square Residual (SRMR). The goodness-of-fit indices are summarized in Table 3.3.

Table 3.3

CFA Goodness - of - fit Indices

Indices	χ^2	RMSEA	CFI/TLI	SRMR	AVE	CR
Cutoff Values	Higher than .05	Lower than .05	Higher than .09	Lower than .08	Higher than .7	Higher than .5

Note. χ^2 = chi-square test; RMSEA = root-mean-square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis index; SRMR = Standardized Root Mean Square Residual; AVE = average variance extracted; CR = composite reliability

The chi-square test of the model fit index evaluates the discrepancy between the unrestricted sample covariance matrix and the restricted covariance matrix (Byrne, 2013). A non-significant probability associated with the chi-square statistics, that is, greater than .05, indicates a good fit. However, the sensitivity of this test to sample size makes it almost always significant for large samples (Harrington, 2009). The RMSEA assesses the degree to which the model reproduces the population covariances with a value below .05 indicating a good fit (Browne & Cudeck, 1992). The CFI and TLI are incremental relative fit indices, measuring the extent to which the model improves over a baseline model in which variables are assumed to be uncorrelated (Kline, 2011). Acceptable model fits are indicated by values greater than .90 (Bentler, 1990). The SRMR is a measure of the discrepancy between the model's predicted covariance matrix and the observed covariance matrix, which takes into account the size of the residuals and the number of degrees of freedom. The SRMR is calculated by dividing the root mean square residual by the square root of the mean squared residual covariance. A lower SRMR value indicates a better model fit, with values less than or equal to .08 generally indicating a good fit.

In addition, the average variance extracted (AVE) and composite reliability (CR) were computed. AVE represents the average amount of shared variance among the indicators that can be accounted for by the latent construct. AVE values range from 0 to 1, with a value higher than 0.5 indicating convergent validity (Fornell & Larcker, 1981). Composite reliability (CR) is a measure of internal consistency reliability in a measurement model. It indicates the extent to which the indicators in the model are consistent in measuring the underlying construct. In general, a CR value of 0.7 or higher is considered acceptable (Raykov, 1997).

3.4.3 Statistical Analysis for RQ 3

To address the third research question, three statistical methods were employed: hierarchical regression analysis, PCA with varimax rotation, and mediation analysis. The hierarchical regression analysis was employed to examine whether reading prosody could explain additional variance in reading comprehension beyond decoding skills (for RQ 3-1) and TRE (for RQ 3-2).

In the hierarchical regression for RQ 3-1, decoding and reading prosody features were entered into the analysis. This model was formulated based on the assumption that reading prosody itself plays a role in reading comprehension (Frazier et al., 2006; Kuhn et al., 2010). The variables were entered into the analysis in a predetermined sequence based on previous research and reading theories (Benjamin & Schwanenflugel, 2010; Hoover & Gough, 1990; LaBerge & Samuels, 1974; Silverman et al., 2013).

Specifically, decoding skills were entered into the first block, pause-related factors into the second block, and pitch-related factors into the third block. This sequence was chosen due to the assumption that minimum levels of decoding skills are necessary for natural and smooth oral reading. Among prosodic features, pause factors were given priority over pitch-related factors in the second block, contrary to most L1 studies which suggest that pitch-related factors are the primary contributors to prosodic development, TRE, and reading comprehension (e.g., Miller & Schwanenflugel, 2008). However, research on L2 English pronunciation and suprasegmental features has indicated that Korean L2 speakers tend to exhibit flatter and more monotonous intonation than L1 speakers (e.g., Kang, 2013), which may suggest a weaker relationship between pitch-related factors and other reading skills. Therefore, pause-related factors were entered in the second block, followed by pitch factors in the third block.

The second model of hierarchical regression was formed to investigate whether reading prosody can account for the unique variance in reading comprehension even beyond what is accounted for by TRE (RQ 3-2). In this model, TRE was entered as the first block, followed by the inclusion of pause-related variables in the second block, and pitch variables in the third block as in the previous hierarchical regression.

Before conducting regression analysis, the assumptions were evaluated based on the different standards. At first, for the evaluation of outliers, standard residuals of each variable were examined if all Z-scores were within +/-3.29 (Field, 2018). Next, in the assumption of collinearity tests, *Tolerance* and *VIF* scores were checked. According to Tabachnick, Fidell, and Ullman (2018), these values of variables should be distributed in a reasonable limit (*Tolerance > .10, VIF <10*). In addition, the assumption of independence was checked by examining *Durbin-Watson* Value. When the value of *Durbin-Watson* approximates 2, it means that the residuals are independent and uncorrelated. However, if the value is less than 1 or greater than 3, it may suggest that the assumption of independent errors is violated, and the residuals are either positively or negatively correlated (Tabachnick et al., 2018). Lastly, linearity and homoscedasticity were also inspected by observing plotted points formed in a straight-line association in the bivariate scatter plots.

To further investigate the hypothesized relationship among decoding skill, prosody, and reading comprehension, a mediation analysis was employed in addition to hierarchical regression. To conduct the mediation analysis, the dimension of reading prosody needed to be reduced, and therefore, PCA with varimax rotation was conducted for reading prosodic features. The result showed that the Kaiser-Meyer-Olkin (KMO) value was .718, suggesting the sampling adequacy was good (Kaiser & Rice, 1974). Bartlett' test of sphericity indicated that the correlation between variables was sufficiently large ($x^2 = 440, p < .001$). The communality values of the variables were all above a cutoff of .45 for the inclusion of a variable in the interpretation of a factor. Eigenvalues revealed the two components in excess of Kaiser's criterion of 1, which accounted for 75.35 % of the variance in the reading prosody features. The two components were labeled as PAUSE and PITCH, and all variables were loaded on them as presented in Table 3.4. The factor scores of the two components were utilized and subsequently entered into the mediation analysis.

Table 3.4

Factor Loading for the Two Components

Item	PAUSE	PITCH
Pitch_SD	.296	**.758**
Pitch_SF	-.198	**.811**
IntraP_ratio	**.948**	.005
UGP_ratio	**.935**	.005
IntraP_duration	**.940**	-.082
InterP_duration	**.705**	.006
Eigenvalues	3.281	1.240
Percentage of variance	54.637	20.708

Note. Pitch_SD=overall intonation contour; Pitch_SF=pitch changes at the sentence-final position; IntraP_ratio=intrasentential pause ratio; UGP_ratio=ungrammatical pause ratio; IntraP_duration= intrasentential pause duration; InterP_duration=intersentential pause duration

Figure 3.6 depicts the proposed relationship between decoding (independent variable), prosody (mediator; M), and reading comprehension (dependent variable). This model is based on previous research that suggests reading prosody plays a mediating role between decoding and reading comprehension (Fernandes et al., 2018; Schwanenflugel et al., 2004; Wolters et al., 2020). The model posits that decoding, as a fundamental component of the reading process, influences the impact of reading prosody on reading comprehension (Kim, 2021a; Wolters et al., 2020). Furthermore, theories and empirical studies on reading prosody suggest that it aids reading comprehension by helping to maintain information in working memory during online processing (Frazier et al., 2006; Koriat et al., 2002; Kuhn et al., 2010). In order to examine the hypothesized relationship among the three components and to confirm the role of reading prosody as a mediator between decoding and reading comprehension in L2 reading, the mediation model was developed and assessed using the data obtained in the current study.

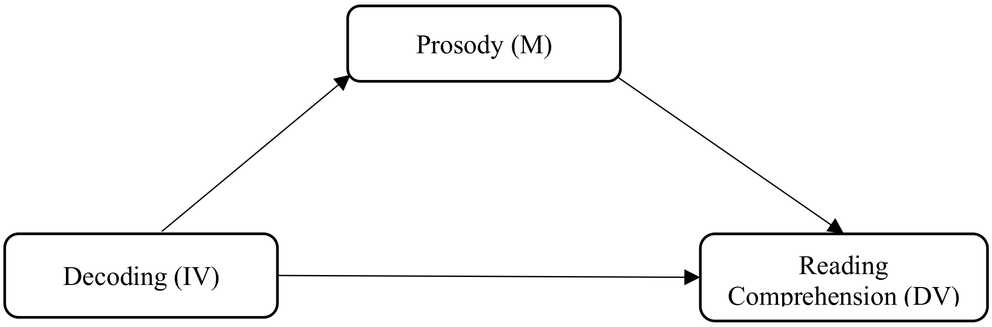


Figure 3.6 The Hypothesized Mediation Model of Reading Prosody

For the mediation analysis, the well-established criteria proposed by Baron and Kenny (1986) were employed to establish a variable as a mediator. First, a significant relationship must exist between the decoding skill (IV) and reading comprehension (DV). Second, a significant relationship must also exist between the decoding skills and the mediator, the prosodic components. Third, the mediator should still predict reading comprehension (RC) even after controlling for decoding skills. Finally, the relationship between the decoding skills and RC should be reduced when a prosodic factor is included in the model. If this reduction results in the complete elimination of the relationship between the decoding skill and the RC, the mediation is considered perfect or complete. If the relationship is merely weakened but not entirely eliminated, the mediation is considered partial. Figure 3.6 illustrates the hypothesized relationship among decoding (IV), prosody (Mediator; M), reading comprehension (DV).

The primary statistical analysis utilized in this study was performed through the use of *SPSS* (IBM Corp, 2019) , especially for conducting ANOVA, PCA, and regression analyses. To investigate the mediating effects of reading prosody on the relationship between decoding skill and reading comprehension, the *PROCESS* macro extension in *SPSS*, developed by Hayes (2017), was employed. For conducting CFA, *M-Plus* (Muthen et al., 2017) was also utilized.

Chapter 4. Results

This chapter reports the research findings based on the three research questions: whether there are differences in reading prosody features depending on fluency skills, whether reading prosody and text reading efficiency constitute the unitary construct of ORF or whether they represent two-dissociable constructs of ORF, and to what extent reading prosody accounts for unique variance in reading comprehension.

4.1 Descriptive Statistics and Correlation

Before conducting one-way ANOVA and regressions, descriptive statistics and bivariate correlation were conducted as preliminary analyses. The results of descriptive statistics are presented in Table 4.1. An examination of the descriptive statistics for reading comprehension found that the mean MCSAT score of the participants was 37.576 and the standard deviation was 13. 506. The minimum and maximum scores were 5 and 64.33 each, showing a wide range of variance. Skewness and kurtosis were within the value of 2, indicating they are almost normally distributed (Field, 2018) (*Skewness=- .237, Kurtosis=- .299*).

Table 4.1

Descriptive Statistics of the Variables (N=90)

Variables	Min	Max	Mean	SD	Skewness	Kurtosis
RC	5	64.33	37.744	13.06	- .237	- .299
TRE[a]	56.92	194.22	105.574	26.535	.782	1.169
Decoding	25	44	35.43	4.332	-. 459	- .010
Pitch_SD[b]	0.86	2.62	1.714	0.383	.057	- .408
Pitch_SF[c]	-.96	4.90	1.996	1.227	.182	- .255
IntraP_ratio	0.05	1.16	0.623	0.238	.198	- .145
UGP_ratio	0.01	0.53	0.260	0.106	.456	.101
IntraP_duration[d]	0.15	6.68	2.620	1.423	.929	.941
InterP_duration[e]	0.42	1.67	0.933	0.302	.620	- .025

Note. TRE=text reading efficiency; Pitch_SD=overall intonation contour; Pitch_SF=pitch changes at the sentence-final position; IntraP_ratio=intrasentential pause ratio; UGP_ratio=ungrammatical pause ratio; IntraP_duration= intrasentential pause duration; InterP_duration=intersentential pause duration
[a] TRE represents the number of correctly read words per minute (WCPM).
[b,c] All pitch variables were measured in the semitone scale.
[d,e] The duration of pauses is presented in seconds.

Next, TRE represents the reading rate and accuracy which was measured in WCPM, and the mean of TRE indicated that the participants read average 105 words correctly in a minute, ranging from 56.92 to 194.22 words per minute and the standard deviation was 26.535 words. TRE scores were normally distributed in that their skewness and kurtosis were within reasonable limits (*Skewness= .782, Kurtosis=1.169*).

The investigation of the participants' decoding skills showed that they read 35.43 words correctly on average with a standard deviation of 4.332. Its highest score was 44 words and the lowest was 25 words, with normal distribution being formed (*Skewness= - .459, Kurtosis= - .010*).

The overall intonation contour drawn from the participants' oral reading was calculated by measuring the mean pitch at the peak of syllables, and the standard deviation of the mean pitch indicated the Pitch_SD values. Measured in the semitone scale, the mean Pitch_SD was 1.714 and the standard deviation was 0.383, showing the normal distribution (*Skewness= .057, Kurtosis=- .408*). Another pitch variable was the changes shown from the last peak to the end of the sentence, named Pitch_SF. The mean of pitch differences was 1.996 semitones and the standard deviation was 1.227. Its skewness (.182) and kurtosis (- .255) suggested that the values of Pitch_SF lie in an almost symmetrical fashion around the mean.

Among the pause factors, pause ratio means the percentage of pauses made in the spaces between words or within words, which were divided by the total number of spaces in a sentence. The mean of IntraP_ratio was .623 and *SD* was .238. The minimum was .05, meaning a student made longer than 100ms of pauses only 5 times out of 100 spaces between words. The maximum value of 1.16 indicated that a participant made pauses longer than 100ms within a word as well as in space between words. The normal distribution of this variable was also confirmed by skewness (.198) and kurtosis (-.145) values. UGP_ratio, another pausal intrusion, was the ratio of pauses made in other than syntactic boundaries to the total number of spaces in a sentence. The mean and the standard deviation of UGP_ratio were .260, and .106, respectively, with its values normally distributed (*Skewness= .456, Kurtosis*= .101).

As for the duration of pauses, the two variables, IntraP_duration and InterP_duration were used. IntraP_duration refers to the total duration of pauses made within sentences while InterP_duration was a time interval between a sentence and the following sentence. The descriptive statistics showed that the mean pause duration within a sentence was 2.620 seconds with a standard deviation of 1.423. The minimum of IntraP_duration was .15 and the maximum was 6.68. Meanwhile, the average InterP_duration was .993 and SD was .032. Its minimum and maximum were .42 and 1.67 each. These two variables showed normal distributions. The histograms illustrating frequencies and distributions of reading prosody variables are presented in Figures 4.1 through 4.6.

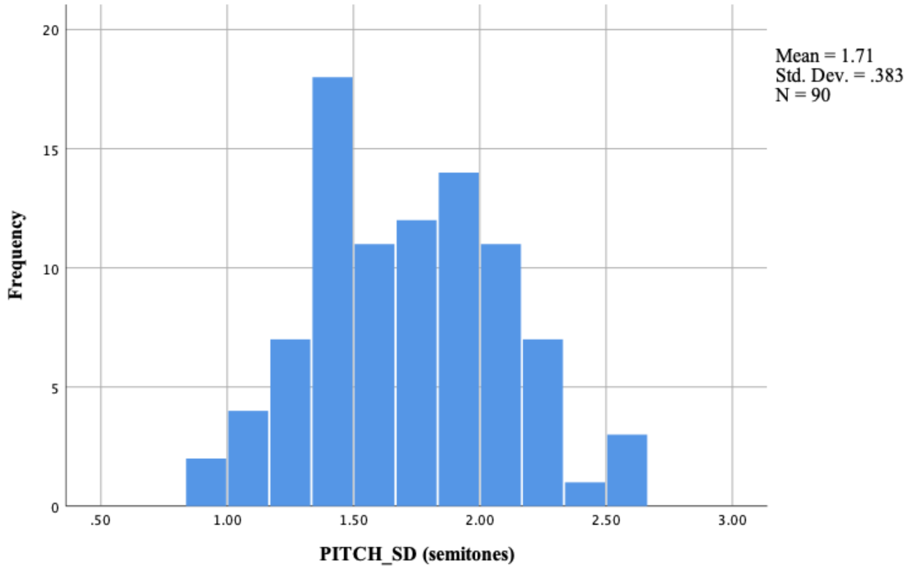


Figure 4.1 Frequency of Pitch_SD

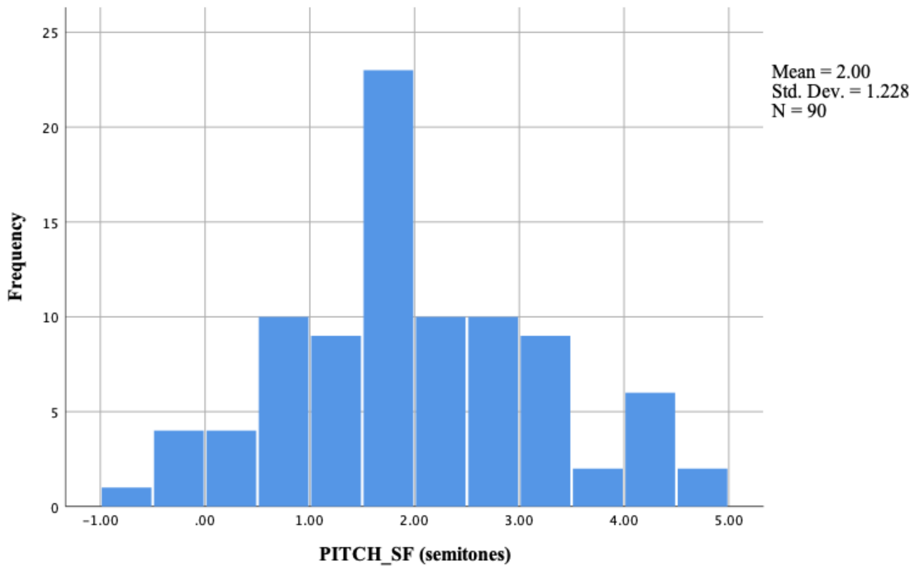


Figure 4.2 Frequency of Pitch_SF

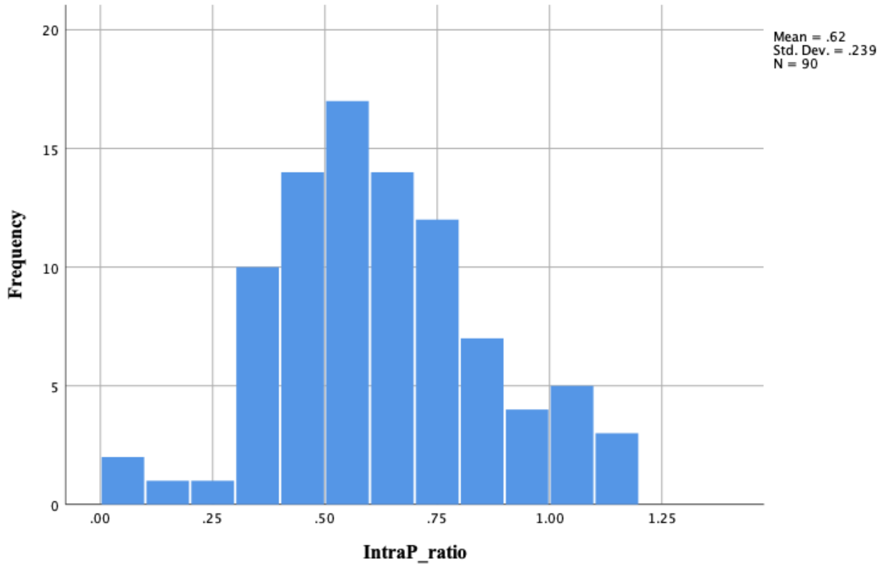


Figure 4.3 Frequency of IntraP_ratio

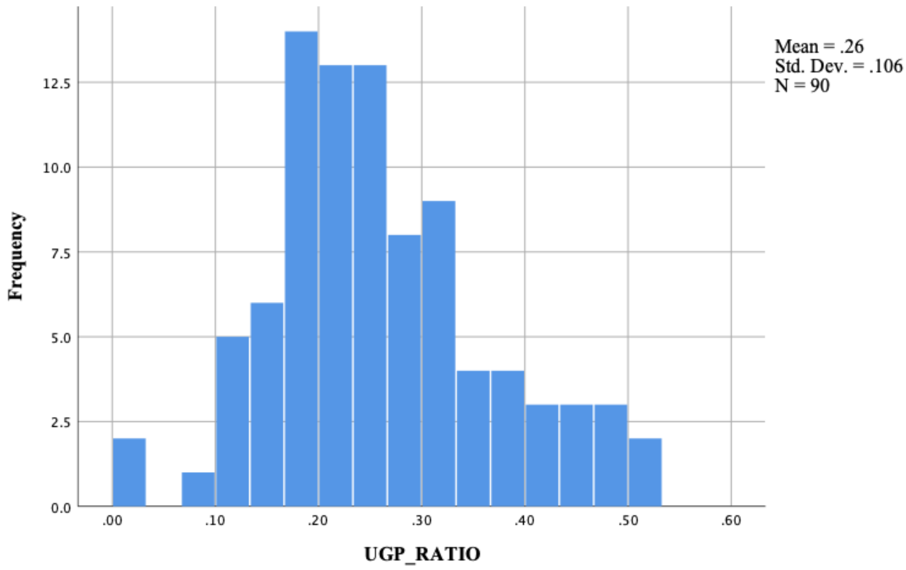


Figure 4.4 Frequency of UGP_ratio

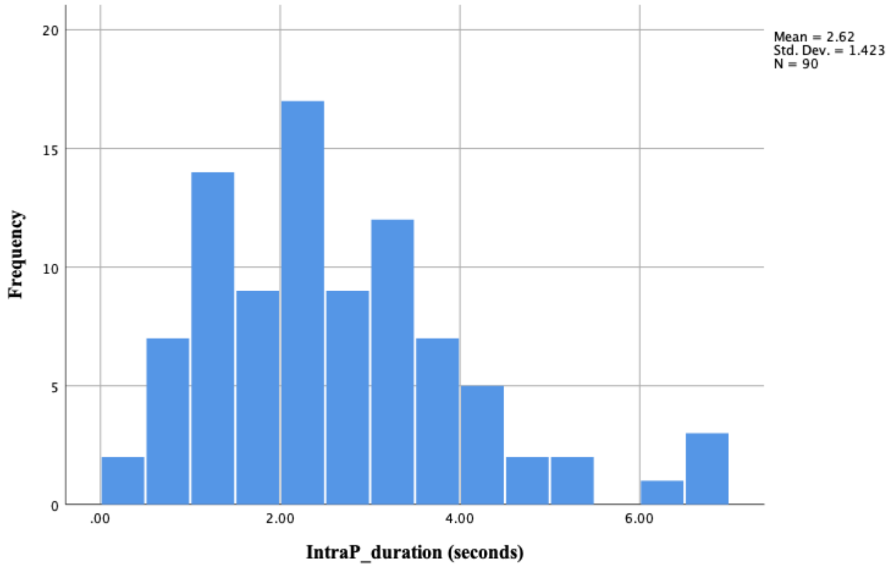


Figure 4.5 Frequency of IntraP_duration

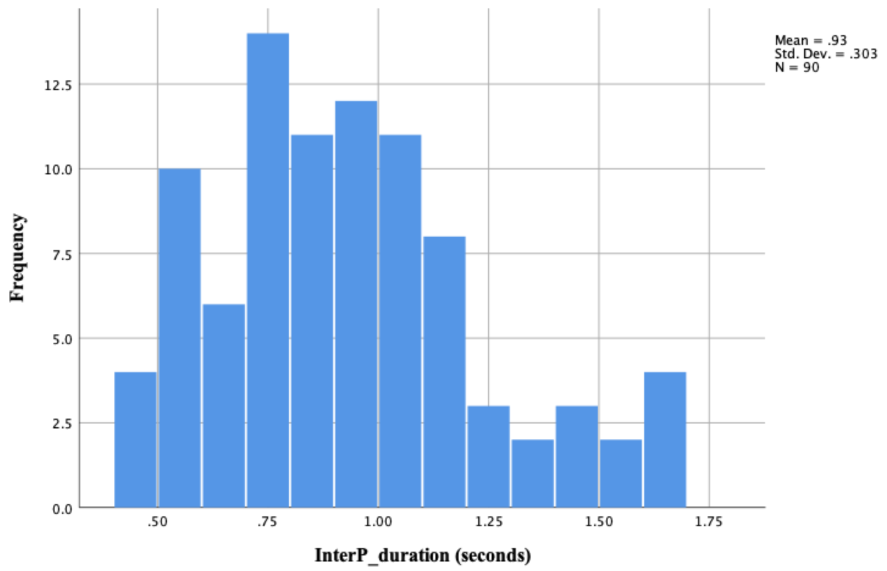


Figure 4.6 Frequency of InterP_duration

Next, in order to examine the relationships among variables, bivariate correlation analysis was conducted. The correlation coefficients and their significances are presented in Table 4.2. Reading comprehension score and TRE were significantly associated with all the other variables except for Pitch_SD. Especially, TRE had a rather strong relationship with pause variables (- .585 to - .799) while the relationships of reading comprehension with the other factors were either moderate or weak, with the coefficients ranging from .261 through .553. Decoding skills were not significantly correlated with pitch-related factors while they had weak to moderate relations with pause factors (-. 239 to -. 541).

Intonation contour (i.e., Pitch_SD) has a significant relationship only with Pitch_SF (r =.238, $p < .5$), UGP_ratio (r =.213, $p < .5$) and InterP_duration (r =.246, $p < .5$). Pitch_SF was significantly correlated with Pitch_SD (r =.238, $p < .5$) and InterP_duration (r =- .172, $p < .5$) although the relationships were weak.

Table 4.2

Bivariate Correlation among the Variables

Variables	TRE	Decoding	Pitch_ SD	Pitch_ SF	IntraP _ratio	UGP _ratio	IntraP _duration	InterP _duration
RC	.553**	.466**	.025	.261*	-.471**	-.452**	-.430**	-.239*
TRE		.564**	-0.049	.310**	-.786**	-.747**	-.799**	-.585**
Decoding			-.097	.120	-.541**	-.528**	-.550**	-.239*
Pitch_SD			-	0.238*	0.196	.213*	0.175	.247*
Pitch_SF			-	-	-0.104	-0.159	-.203	-.172*
IntraP_ratio			-	-	-	.973**	.868**	.506**
UGP_ratio			-	-	-	-	.834**	.450**
IntraP _duration			-	-	-	-	-	.623**

Note. $^*p < .05$, $^{**}p < .01$; TRE=text reading efficiency; Pitch_SD=overall intonation contour; Pitch_SF=pitch changes at the sentence-final position; IntraP_ratio=intrasentential pause ratio; UGP_ratio=ungrammatical pause ratio; IntraP_duration= intrasentential pause duration; InterP_duration=intersentential pause duration

IntraP_ratio had moderate to strong relationships with other variables (-. 471 to .973) while any significant correlation was not found with the two Pitch factors. In particular, the correlation coefficient between UGP_ratio and IntraP_ratio was considerably high ($r = .973$, $p < .01$), suggesting the possibility of multicollinearity in the following regression analyses. In fact, the values of

VIF and *Tolerance* were off the standard cutoff (*VIF > 10, Tolerance < .01*) (Tabachnick et al., 2018) when the two variables were entered into regressions.

IntraP_duration was significantly associated with variables other than pitch-related ones, showing stronger relationships with pause ratio variables (r = .834 - .868, p < .01) and TRE (r = - .799, p < .01) than the other factors. Lastly, inter-sentential pause duration (InterP_duration) showed significant relationships with all the variables, even though the strength of associations was weak to moderate (- .172 to 623).

4.2 Differences in Reading Prosody Features Depending on Reading Skill

This section pertains to the investigation of potential variances in reading prosody characteristics based on fluency skill of the participants. The study classified the participants into three different groups according to their L2 fluency skills. The composite scores for L2 fluency skills were generated using PCA that included decoding scores and TRE. The three groups (i.e., Higher, Intermediate, and Lower) were formed based on the tertile of these composite scores and 30 participants were placed in each group. Table 4.3 presents descriptive statistics of the reading prosody scores for each group.

Table 4.3

Descriptive Statistics of Reading Prosody Features by Groups

Variables		N	M	SE	Min	Max
Pitch_SD[a]	L	30	1.733	.064	.870	2.270
	I	30	1.740	.069	1.00	2.620
	H	30	1.669	.076	.860	2.500
	Total	90	1.714	.040	.860	2.620
Pitch_SF[b]	L	30	1.655	.214	-.960	4.00
	I	30	1.785	.177	-.060	3.83
	H	30	2.549	.245	-.060	4.90
	Total	90	1.996	.129	-.960	4.90
IntraP_ratio	L	30	.813	.039	.380	1.160
	I	30	.614	.028	.320	.910
	H	30	.443	.032	.050	.740
	Total	90	.623	.025	.050	1.160
UGP_ratio	L	30	.342	.019	.130	.530
	I	30	.253	.012	.140	.400
	H	30	.185	.013	.010	.350
	Total	90	.260	.011	.010	.530
IntraP_duration	L	30	3.843	.274	1.250	6.680
	I	30	2.513	.140	.980	3.890
	H	30	1.505	.130	.150	2.850
	Total	90	2.620	.150	.150	6.680
InterP_duration	L	30	1.054	.060	.420	1.670
	I	30	.968	.052	.460	1.670
	H	30	.775	.039	.460	1.190
	Total	90	.932	.031	.420	1.670

Note. M=Mean; SE=Standard Error; Min=Minimum; Max=Maximum; L=lower group, I=Intermediate group, H=Higher group; Pitch_SD=overall intonation contour; Pitch_SF=pitch changes at the sentence-final position; IntraP_ratio=intrasentential pause ratio; UGP_ratio=ungrammatical pause ratio; IntraP_duration=intrasentential pause duration; InterP_duration=intersentential pause duration
[a,b] All pitch variables were measured in the semitone scale.

Table 4.4

Difference of Reading Prosody Features Depending on L2 Fluency Skills

Variables		*SS*	*df*	*MS*	*F*	*p*	η^2
Pitch_SD	Between	.092	2	.046	.308	.736	.007
	Within	12.975	87	.149			
	Total	13.066	89				
Pitch_SF	Between	14.003	2	7.001	5.070	.008**	.104
	Within	120.138	87	1.381			
	Total	134.140	89				
IntraP_ratio	Between	2.058	2	1.029	29.778	.000***	.406
	Within	3.006	87	.035			
	Total	5.063	89				
UGP_ratio[a]	Welch		2		21.310	.000***	
			56				
IntraP_ duration[b]	Welch		2		34.033	.000***	
			54				
InterP_ duration	Between	1.231	2	.616	7.737	.001**	.150
	Within	6.924	87	.080			
	Total	8.155	89				

Note. SS=sum of squares; df=degree of freedom; MS=mean square; Pitch_SD=overall intonation contour; Pitch_SF=pitch changes at the sentence-final position; IntraP_ratio=intrasentential pause ratio; UGP_ratio=ungrammatical pause ratio; IntraP_duration= intrasentential pause duration; InterP_duration=intersentential pause duration
[a,b] The assumption of equality of variances was violated and the Welch test was used.

A series of one-way ANOVAs were carried out to answer the first research question and the results are presented in Table 4.4. As for Pitch_SD, the significant difference was not shown [$F(2,87)= .308, p= .736, \eta^2= .007$], meaning that overall intonation contours were not significantly different across the reading proficiency groups [M_L=1.733, *SE*= .064; M_I= 1.740, *SE*= .069; M_H= 1.669, *SE*= .076]. However, significant differences were found in Pitch_SF [$F(2,87)= 5.070, p=.008, \eta^2= .104$], which means that the pitch changes at the sentence-final position were significantly different among three groups [M_L=1.655, *SE*= .214; M_I= 1.785, *SE*= .177; M_H= 2.549, *SE*= .245]. This result suggested that L2 readers with higher decoding skills and reading fluency tended to show more dramatic pitch changes at the sentence-final position when reading declarative sentences.

Most of the pause factors were shown to have significant differences depending on L2 fluency skills. Firstly, the differences in intrasentential pause ratio across the three groups were statistically significant [$F(2,87)=29.778, p=.000, \eta^2= .406$], which reveals that these groups differ in terms of frequencies of pausing when reading English sentences [$M_L= .813, SE= .039; M_I= .614, SE= .028; M_H= .443, SE= .032$]. The results of the ungrammatical pause ratio showed a similar pattern to those of the pause ratio. There were significant differences across the three groups [$F(2,56)=21.310, p=.000$], meaning that the frequency of pausing which occurs in other than syntactic boundaries was significantly different depending on the participants' fluency skills in English [$M_L= .342, SE= .019; M_I= .253, SE= .012; M_H= .185,$

SE= .013]. These results suggest that as the participants' L2 fluency skills improve, they are likely to make fewer pausal intrusions within a sentence, and the pauses they make are more likely to occur in syntactic boundaries than those made by L2 readers with lower fluency skills.

Next, the index for the duration of intrasentential pauses exhibited significant differences across groups [*F(2,54)=34.033, p=.000*], revealing that there were decreases in pausal duration as the readers' fluency skill was improved [M_L *= 3.843, SE= .274;* M_I *= 2.513, SE= .140;* M_H *= 1.505, SE= .130*]. In the same vein, the intersentential duration of pauses was significantly different across the three groups [*F(2,87)=7.737, p= .001,* η^2 *= .150*]. It means that the duration of pauses that these L2 readers made after finishing reading a sentence and before starting to read the following sentence was likely to be longer as readers were less competent in decoding skills and text reading efficiency [M_L *= 1.054, SE= .060;* M_I *= .968, SE= .052;* M_H *= .775, SE= .039*]. Figures 4.7 through 4.12 illustrated these results.

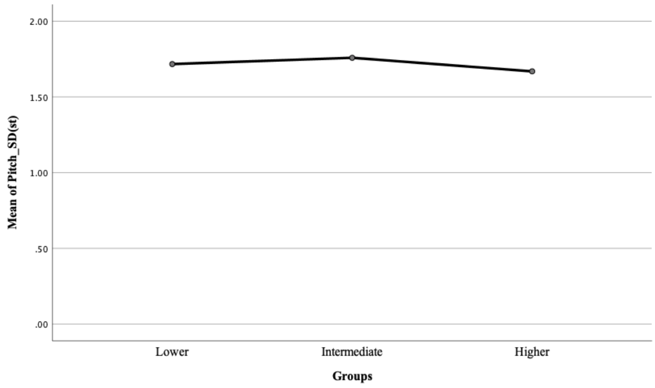


Figure 4.7 Differences of Pitch_SD across Groups

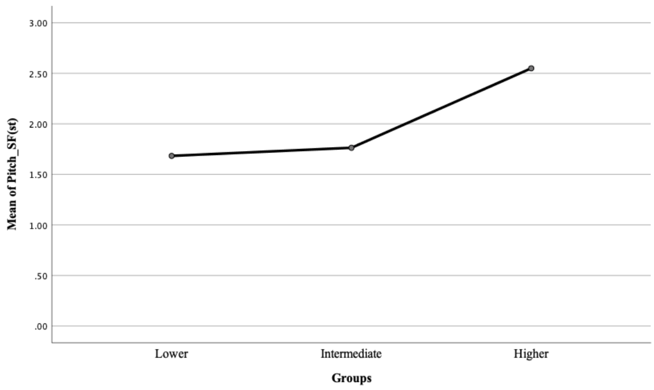


Figure 4.8 Differences of Pitch_SF across Groups

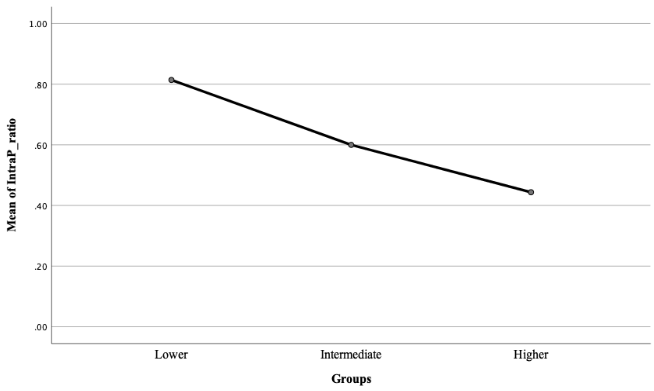


Figure 4.9 Differences of IntraP_ratio across Groups

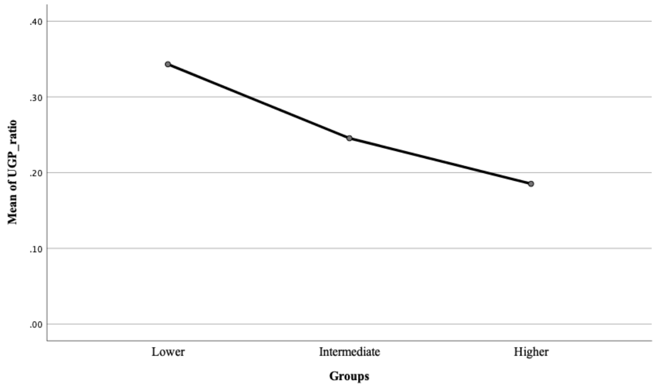


Figure 4.10 Differences of UGP_ratio across Groups

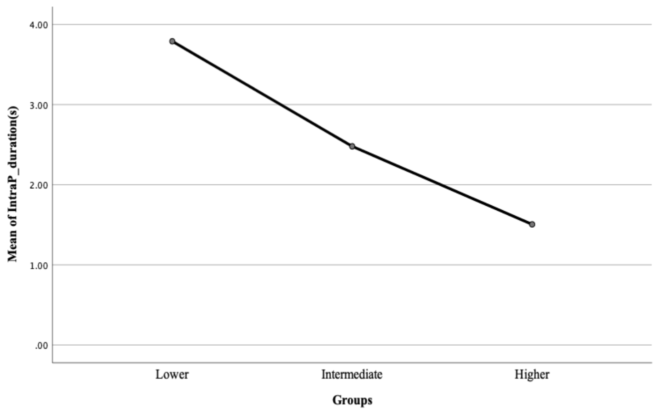


Figure 4.11 Differences of IntraP_duration across Groups

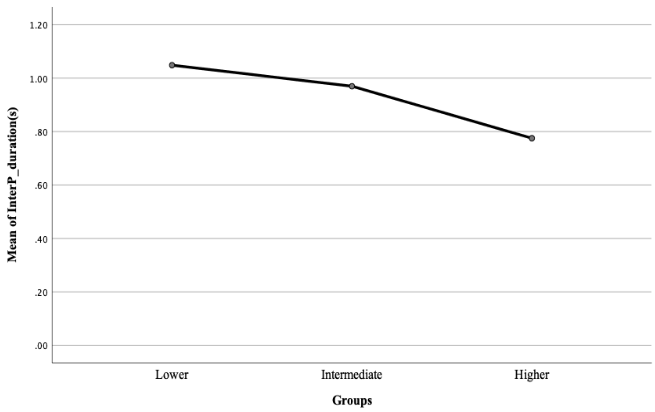


Figure 4.12 Differences of InterP_duration across Groups

Since the one-way ANOVAs revealed significant overall *F*-ratios in Pitch_SF, IntraP_ratio, UGP_ratio, IntraP_duration, and InterP_duration, the post-hoc *Bonferroni* analyses were carried out. Table 4.5 depicts the results of post-hoc comparisons of reading prosody features.

Table 4.5

Post-hoc Comparisons of L2 Reading Prosody Features across the Three Fluency Skill Groups

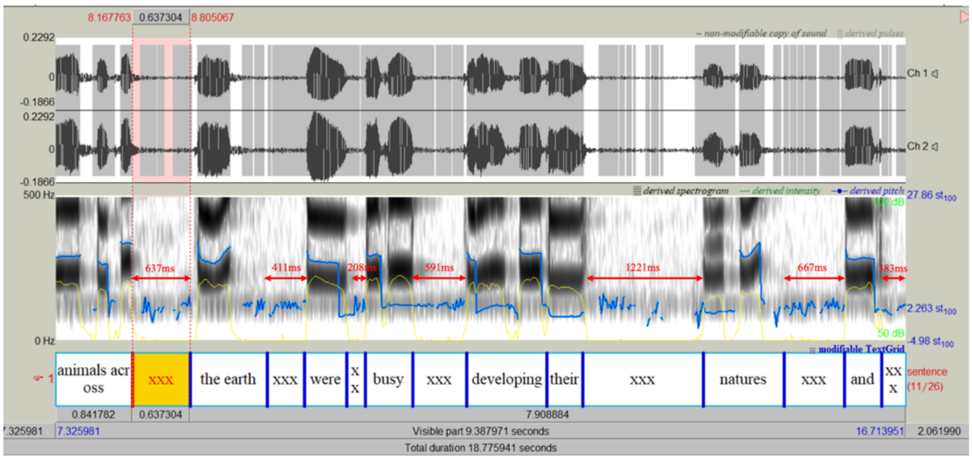
Variables	Group (I)	Group (J)	Mean diff.(I-J)	*SE*	*p*	95% Confidence Interval	
						LB	UB
Pitch_SF	L	I	-.129	.303	1	-.870	.611
		H	-.894*	.303	.012	-1.634	-.153
	I	H	-.764*	.303	.041	-1.505	-.023
IntraP_ratio	L	I	.199**	.047	.000	.082	.316
		H	.370**	.047	.000	.252	.487
	I	H	.170**	.047	.002	.053	.288
UGP_ratio	L	I	.088**	.022	.000	.035	.142
		H	.156**	.022	.000	.102	.210
	I	H	.067**	.022	.008	.013	.121
IntraP_ duration	L	I	1.329**	.273	.000	.661	1.9997
		H	2.337**	.273	.000	1.669	3.005
	I	H	1.008**	.273	.001	.339	1.676
InterP_ duration	L	I	.086	.072	.466	-.091	.264
		H	.279**	.072	.001	.101	.457
	I	H	.193**	.072	.025	.015	.371

Note. $*p < .05$, $**p < .01$; Based on estimated marginal means; Adjustment for multiple comparisons: *Bonferroni*; SE=standard error; LB=lower bound; UB=upper bound; L= lower group, I=Intermediate group, H=Higher group; Pitch_SF=pitch changes at the sentence-final position; IntraP_ratio=intrasentential pause ratio; UGP_ratio=ungrammatical pause ratio; IntraP_duration= intrasentential pause duration; InterP_duration=intersentential pause duration

As shown in Table 4.5, the post-hoc results found that there were no significant differences between the Lower and Intermediate groups in terms of Pitch_SF and InterP_duration. However, the Higher reading group exhibited a discernable difference in these variables compared to the other groups. This suggests that the time interval between sentences becomes considerably shorter, and the pitch drops far more at the end of sentences only when L2 readers are sufficiently fast and accurate in word decoding and text reading.

The post-hoc analysis revealed significant mean differences in pause ratio (InterP_ratio), ungrammatical pause ratio (UGP_ratio), and intrasentential pause duration (IntraP_duration) across the three reading proficiency groups. These difference in pause pattern is clearly illustrated in Figure 4.13, representing a comparison of the pause patterns between a lower-level and a higher-level participant shown on the spectrographic window.

A. Lower-level reader



B. Higher-level reader

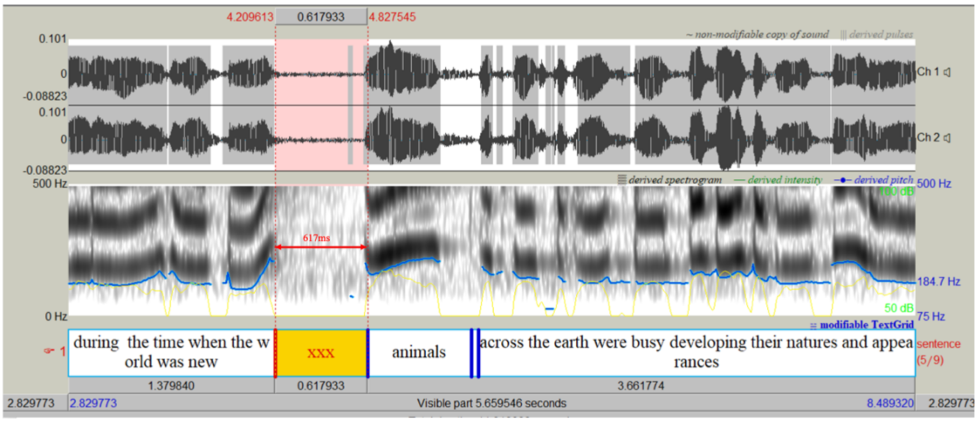


Figure 4.13 Comparison of Pause Patterns between the Two Participants of Different Fluency Skills

According to Figure, a Lower-level reader exhibited a higher frequency and longer duration of pauses compared to the Higher-level reader. Specifically, the Lower-level reader made seven pauses of a total duration of 4118ms while the Higher-level reader paused one time of 617ms. Furthermore, the pauses made by the Lower-level reader often occurred at positions that were not aligned with meaningful units, such as phrasal and clausal boundaries. In contrast, the Higher-level reader demonstrated a single grammatically appropriate pause. For example, the Lower-level reader paused between "*across*" and "*the*", which is within the nominal boundary, whereas the pause the Higher-level reader made between "*new*" and "*animals*" was considered grammatical because it occurred at the clausal boundary. These findings suggest that L2 readers' fluency skills are closely related to their use of pausing when reading.

To sum up the results on L2 learners' reading prosody features, the overall intonation contour (i.e., Pitch_SD) was not significantly different depending on L2 learners' fluency skills. In contrast, most of the pause factors (i.e., the pause ratio, the ungrammatical pause ratio, and the duration of pauses) were shown to have decreasing tendencies as L2 reading skills were more proficient. When it comes to the pitch changes at the sentence-final position and the pause duration across sentences, significant differences were found only between higher reading groups and the other groups. This phenomenon suggests that dramatic changes in the two variables seemed to occur when L2 readers were proficient enough in decoding and fluent reading.

4.3 Dimensionality of Oral Reading Fluency

This section deals with the second research question, whether reading prosody and TRE, as components of ORF, constitute a unitary construct as defined in the L1 literature or they are dissociable but related components. This was explored using confirmatory factor analysis (CFA).

To test whether the hypothesized model of ORF and its relationships to TRE and reading prosody are consistent with the data, CFA was conducted. Two alternative models shown in Figures 3.4 to 3.5 were fit to the data to determine the dimensionality of the prosody and text reading efficiency indicators. The outcomes of the model fit indices are presented in Table 4.6.

Table 4.6

The Model Fit Indices of Two Alternative Models

Model	χ^2	*df*	RMSEA	CFI	TLI	SRMR
Unitary-construct[a]	21.882**	7	.154	.949	.891	.067
Unitary-construct[b]	1.374	2	.000	1.00	1.00	.02
Two-dissociable-construct[c]	20.756**	7	.148	.953	.900	.086

Note. $^{**}p < .01$; χ^2 = chi-square test; RMSEA = root-mean-square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis index; SRMR = Standardized Root Mean Square Residual
[a,c] Observed variables: TRE, Pitch_SD, Pitch_SF, IntraP_ratio, IntraP_duration, InterP_duration
[b] This is the modified model of Unitary-construct[a];
Observed variables: TRE, IntraP_ratio, IntraP_duration, InterP_duration

The Unitary-construct model[a] with all the prosody variables (Pitch_SD, Pitch_SF, IntraP_raio, IntraP_duration, and InterP_duration) and TRE resulted in poor fit [$\chi^2 = 21.882$, $df = 1$, $p < .01$, $RMSEA = .154/CI = .082–.229$, $CFI = .949$, $TLI = .891$, $SRMR = .067$]. When the parameter estimate coefficients were examined, it was discovered that Pitch_SD was an insignificant parameter ($B = -.173$, $p = .09$). Also, another pitch variable, Pitch_SF has low factor loading, suggesting the possibility of compromising the construct validity (Bagozzi & Yi, 1988). Therefore, all the pitch-related variables were excluded from the model. The model without Pitch_SD and Pitch_SF (i.e., Unitary-construct model[b]) showed excellent fit [$\chi^2 = 1.374$, $df = 2$, $p = .503$, $RMSEA = 0/CI = 0 – .187$, $CFI = 1$, $TLI = 1$, $SRMR = .02$]. On the other hand, the two-dissociable-constructs model resulted in poor fit [$\chi^2 = 20.756$, $df = 7$, $p = .004$, $RMSEA = .148/CI = .077–.223$, $CFI = .953$, $TLI = .900$, $SRMR = .900$].

Thus, these statistical results indicate that the most optimal model for the data was the unitary-construct model comprising only TRE and pause-related variables. The parameter estimate coefficients of this Unitary-construct model are presented in Table 4.7 and Figure 4.14. Also, the parameter estimate coefficients of the two-dissociable-constructs model are presented in Appendix C.

Additionally, to assess the reliability and validity of the unitary-construct model, the average variance extracted (AVE) and composite reliability (CR) were computed. The values obtained for AVE and CR were 0.630 and 0.728, respectively. These results suggest that the unitary-construct model, consisting of TRE and pause variables, provides a dependable estimate of L2 ORF (Fornell & Larcker, 1981; Raykov, 1997). To summarize, the results of CFA showed that L2 ORF is a unitary but multidimensional component that is comprised of TRE and pause-related prosodic features.

Table 4.7

Parameter Estimate Coefficients of the Unitary-construct Model

	Variables	*B*	S.E.	Est./S.E.	*p*
Factor Loading					
ORF	TRE	.866	.046	18.948	.000***
	IntraP_ratio	-.750	.073	-10.278	.000***
	IntraP_duration	-.922	.043	-21.569	.000***
	InterP_duration	-.675	.064	-10.476	.000***
Correlation					
	IntraP_ratio with TRE	-.410	.128	-2.338	.01**
	IntraP_ratio with IntraP_duration	.0693	.144	2.407	.016*
Residual Variance					
	TRE	.250	.079	3.252	.001***
	IntraP_ratio	.437	.109	4.190	.000***
	IntraP_duration	.149	.079	1.732	.083
	InterP_duration	.545	.087	6.276	.000***

Model Fit Indices					
χ^2	*df*	RMSEA	CFI	TLI	SRMR
1.374	2	.000	1.00	1.00	.02

Note. **p* < *.05*; ***p* < *.01*;****p* < *.001*; B =standardized factor loading; S.E.=standard error; Est.S.E.=estimated standard error; TRE=text reading efficiency; IntraP_ratio=intrasentential pause ratio; IntraP_duration=intrasentential pause duration; InterP_duration=intersentential pause duration; RMSEA = root mean square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis index; SRMR = standardized root mean square residual

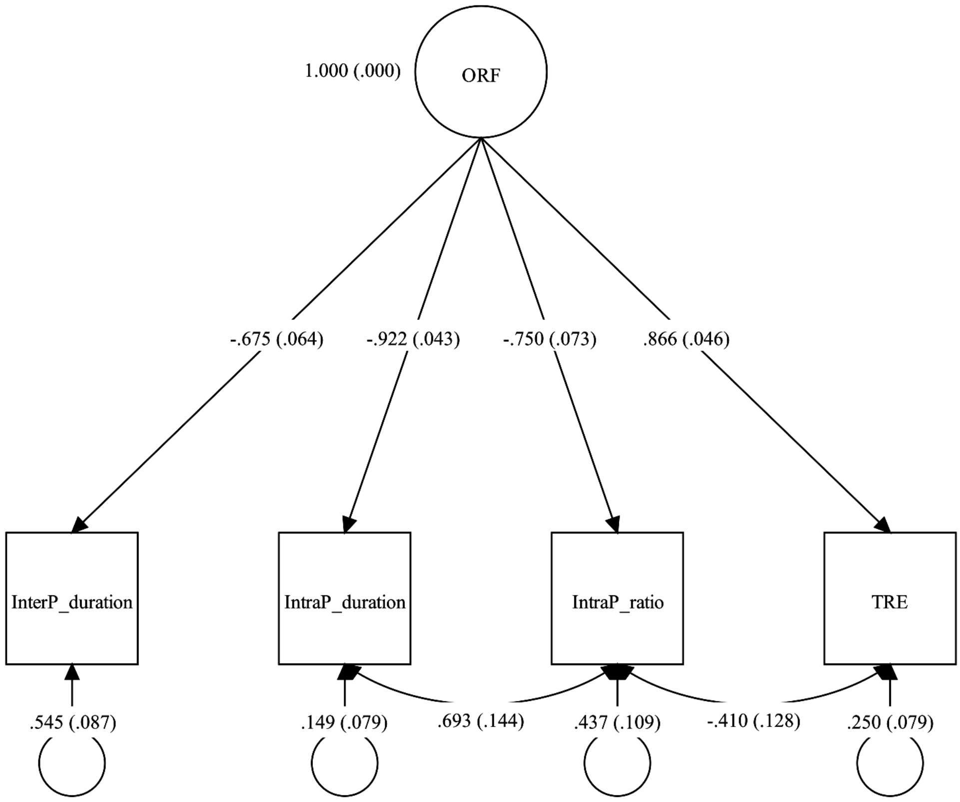


Figure 4.14 The Unitary-construct Model of ORF

4.4 Reading Prosody as a Predictor of Reading Comprehension

This section deals with how much additional variance of reading comprehension reading prosody features account for after controlling for decoding and TRE. Therefore, the results of a couple of hierarchical linear regressions are presented in which decoding skills and TRE are entered into the first block at each regression model. In addition, to examine the role of prosodic features between decoding skills and reading comprehension, the results of mediation analysis are presented.

4.4.1 Unique Contributions of Reading Prosody to Reading Comprehension Beyond the Decoding Skills

In assessing the assumption of the regression, the analysis of standard residuals showed that all the Z-scores were between -2.437 and 3.103, indicating no outliers. However, when it comes to the multicollinearity test, the two highly correlated variables, UGP_ratio (*Tolerance* =.051*, VIF*=19.606) and the IntraP-ratio (*Tolerance = .046, VIF=21.946)* were shown to violate the assumption of multicollinearity. To determine which variable, one of the two options, would be selected, multiple regression analyses were conducted with reading comprehension as the dependent variable. As presented in Table 4.8, the regression model with IntraP-ratio had slightly better predictability than that with UGP_ratio even though the model comparison index AIC and BIC were the same. Therefore, the ungrammatical pause ratio was excluded from the next step and only the intrasentential pause ratio was retained.

Table 4.8

Comparison of the Models with Different Pause Variables

Model	B	SE	*R*	R^2 (*Adj*.)	Model *F*	*AIC/BIC*
IntraP_ratio[a]	-.510**	10.289	.542	.294 (.252)	*F* (5, 84) = 6.953***	10/9.77
UGP_ratio[b]	-.402*	21.897	.525	.275 (.232)	*F* (5, 84) = 6.387***	10/9.77

Note. ** $p < .01$; *** $p < .001$; *df*= degree of freedom; DV= reading comprehension
[a]Predictors: (Constant), Pitch_SD, Pitch_SF, IntraP_ratio, IntraP_duration, InterP_duration
[b]Predictors: (Constant), Pitch_SD, Pitch_SF, UGP_ratio, IntraP_duration, InterP_duration

Except for the two variables, multicollinearity was not problematic [decoding skill, *Tolerance* = .610, *VIF*=1.640; Pitch_SD, *Tolerance* = .846, *VIF* = 1.181; Pitch_SF, *Tolerance* = .854, *VIF* = 1.171; IntraP_ratio, *Tolerance* = .234, *VIF* =4.271; IntraP_duration, *Tolerance* = 0.177, *VI F*=5.657; InterP_duration, *Tolerance* = .520, *VIF* = 1.924]. In addition, the assumption of independence was met (*Durbin-Watson* = 1.964), and a straight-line association in the bivariate scatter plot suggested linearity and homoscedasticity as shown in Figure 4.15.

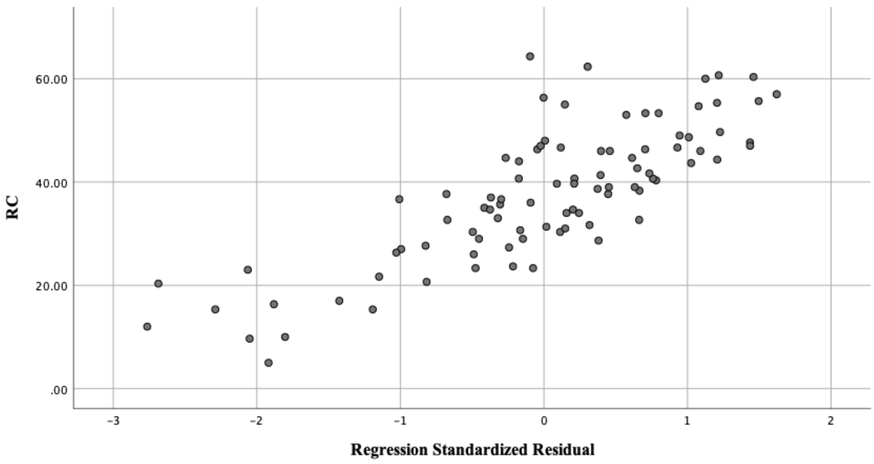


Figure 4.15 RC by Standardized Residual Scatter Plot

The results of HLR are presented in Table 4.9. In *Decoding Model*, R^2-value was .223 [$F(1,88)$ =25.287, p = .000]. With the addition of reading prosody features in the *Pause* model, the R^2-value significantly increased [ΔR^2 = .080, $F(3,85)$ =29.519, p = .026], indicating that pause-related factors made a significant contribution to the equation in this step. At the last step where Pitch-factors were added, R^2-value significantly increased [ΔR^2 = .052, $F(2,83)$ = 32.893, p = .039], meaning pitch factors were playing a role in explaining reading comprehension. Taken together, R^2 = .355 with a 95% confidence interval between .208 to .501 (Soper, 2016; Tabachnick et al., 2018), suggesting that more than a third of the variability of reading comprehension could be explained by knowing decoding skill and reading prosody factors.

Table 4.9

Contributions of Reading Prosody to RC Beyond Decoding Skills

Model Summary

Model	*R*	R^2	Adj.R^2	SE	*F*	ΔR^2	*df*1	*df*2	Sig.*F*
Decoding[a]	.472	.223	.214	11.582	25.287	.223	1	88	.000***
Pause variables[b]	.550	.303	.270	11.165	29.519	.080	3	85	.026*
Pitch variables[c]	.596	.355	.309	10.866	32.893	.052	2	83	.039*

Coefficients

Variables	*β*	SE	B	*t*	Sig.
Decoding	.959	.341	.318	2.815	.006**
IntraP_Ratio	-24.168	9.980	-.441	-2.422	.018*
IntraP_Duration	1.723	1.925	.188	.895	.373
InterP_Duration	-3.955	5.278	-.092	-.749	.456
Pitch_SD	4.450	3.267	.130	1.362	.177
Pitch_SF	1.745	1.015	.164	1.718	.089†

Note. †$p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$; *df*= degree of freedom;
Pitch_SD=overall intonation contour; Pitch_SF=pitch changes at the sentence-final position; IntraP_ratio=intrasentential pause ratio; IntraP_duration= intrasentential pause duration; InterP_duration=intersentential pause duration

[a] Predictors: (Constant), decoding

[b] Predictors: (Constant), decoding, IntraP_ratio, IntraP_duration, InterP_duration,

[c] Predictors: (Constant), decoding, IntraP_ratio, IntraP_duration, InterP_duration, Pitch_SD, Pitch_SF

The coefficients of the final model are also displayed in Table 4.9. As assumed in the HLR model, decoding skills were significant factors with β-value of .959 (p = .006, .281 < CI < 1.631). In a similar way to the results of the previous section, IntraP_ratio was significant in predicting reading comprehension (β = -24.168, p = .018) and its confidence interval was between -44.018 and -4.318. The other reading prosody factors were not significant contributors to reading comprehension [IntraP_duration, β = 1.723, p = .373; InterP_duration, β = -3.955, p = .456; Pitch_SD, β = 4.450, p = .177; Pitch_SF, β =1.745, p = .089].

4.4.2 Reading Prosody as a Mediator to Reading Comprehension

To examine the potential mediating role of the prosodic components (i.e., PAUSE and PITCH) in the relationship between decoding skills and reading comprehension, a mediation analysis was conducted. As discussed in Section 3.4.3, this model was constructed based on the following assumptions: i) when a learner becomes proficient in decoding words, it enables them to allocate cognitive resources toward reading with prosody (Kim et al., 2021a; Schwanenfluegel et al., 2004; Wolters et al., 2020). ii) reading prosody offers extra feedback to the leaner about the significant meaning and structure of the language being read (Frazier et al., 2006; Kuhn et al., 2010). iii) hence, the learner's reading skills are enhanced beyond what can be attributed solely to their decoding speed.

Baron and Kenny’s (1986) criteria were used to determine if reading prosody acts as a mediator. According to their framework, the following conditions must be met: i) there should be a significant relationship between decoding skills and reading comprehension, ii) decoding skills should also have a significant relationship with reading prosody, iii) the reading prosody should be able to predict reading comprehension even after accounting for decoding skills, iv) when the prosodic factor is included in the model, the relationship between decoding skills and reading comprehension should be reduced.

When the mediation analysis was conducted with factor scores of PITCH as a mediator, the result indicated that PITCH did not meet the criteria, failing in accounting for reading comprehension beyond decoding skills. Therefore, only PAUSE was included in the mediation analysis.

Table 4.10 shows that the total effect of decoding skills on reading comprehension was significant (B = .472). The indirect path from decoding skills to PAUSE was also significant, with a standardized beta value of -.510. Additionally, an indirect path from the mediator to reading comprehension was found, indicating that the mediator still predicted reading comprehension even after controlling for the effect of decoding skills (B = .157).

Moreover, the standardized beta coefficient of the direct path between decoding skills and reading comprehension was reduced to .315 when the mediator was entered into the equation. Therefore, the results suggest that the PAUSE factor partially mediates the relationship between decoding skills and reading comprehension. Figure 4.16 depicts the mediation of the PAUSE factor between decoding skills and reading comprehension.

Table 4.10

Model Summary of the Mediation Analysis

					95% CL	
Model	Path	R^2	Std.B	SE	LB	UB
Total effect	Decoding-RC (C)	.223	.472***	.2834	.862	1.989
Indirect effect	Decoding-PAUSE (a)	.260	-.510***	-.012	-.160	-.759
Indirect effect	PAUSE-RC (b)	.293	.157***	1.370	.1307	.8563
Direct effect	Decoding-RC (c)		.315***	3.163	3.236	1.580

Note. $p < .01$; *** $p < .001$; LB =Lower Bound; UB = Upper Bound; PAUSE = Factor score for IntraP_ratio, IntraP_duration, and InterP_duration

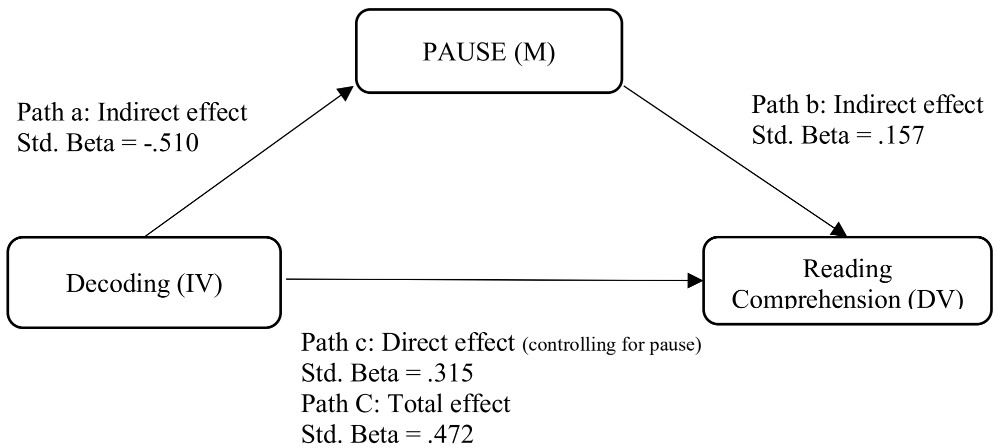


Figure 4.16 The Mediation of PAUSE between Decoding Skills and Reading Comprehension

4.4.3 Predictability of Reading Prosody Beyond Text Reading Efficiency

In order to investigate the potential predictive role of reading prosody, specifically when analyzed with the TRE measure, a hierarchical linear regression analysis was conducted. Before running the regression, its assumption was evaluated. In the analysis of standard residuals of reading comprehension, all Z-scores were within reasonable limits (*Std. Residual Min* = -2.677, *Std. Residual Max* = 2.902). *Tolerance* and *VIF* were within a reasonable limit, meaning that multicollinearity is not problematic [TRE, *Tolerance* = .272, *VIF* = 3.676; Pitch_SD, *Tolerance* = .826, *VIF* = 1.210; Pitch_SF, *Tolerance* = .798, *VIF* = 1.253; IntraP_ratio, *Tolerance* = .201, *VIF* = 4.984; IntraP_ duration, *Tolerance* = .185, *VIF* = 5.416; InterP_duration, *Tolerance* = .543, *VIF* = 1.843]. The *Durbin-Watson* Value was 2.065, satisfying the independence assumption. Plotted points distributed along the straight line in the scatter plot suggested linearity and homoscedasticity.

As shown in Table 4.11, R^2-value was 0.313 [$F(6,83)$ = 40.063, p = .000] when TRE was entered first. However, the addition of pause and pitch factors did not improve R^2 [pause variables, ΔR^2 = .023, $F(3,85)$ = 41.035, p = .410; pitch variables, ΔR^2 = .017, $F(2,83)$ = 42.136, p = .337], thus indicating that there was no significant increase in the prediction of reading comprehension. Similarly, the coefficient table shows that TRE was the only significant predictor of reading proficiency (β = .230, p = .007) with a reasonable confidence interval

(.064 < CI < .396).

Table 4.11

Contributions of Reading Prosody to RC Beyond TRE

Model Summary

Model	*R*	R^2	Adj.R^2	SE	*F*	ΔR^2	*df*1	*df*2	Sig.*F*
TRE[a]	.559	.313	.305	10.894	40.063	.313	1	88	.000***
Pause variables[b]	.579	.336	.304	10.899	41.035	.023	3	85	.410
Pitch variables[c]	.594	.353	.306	10.886	42.136	.017	2	83	.337

Coefficients

Variables	*β*	SE	B	*t*	Sig.
TRE	.230	.083	.466	2.755	.007**
Pitch_SD	-16.094	10.801	-.294	-1.490	.140
Pitch_SF	1.255	1.887	.137	.665	.508
IntraP_ratio	4.419	5.175	.102	.854	.396
IntraP_duration	2.298	3.313	.067	.694	.490
InterP_duration	1.175	1.052	.110	1.117	.267

Note. * $p < .05$; ** $p < .01$; *** $p < .001$; *df*= degree of freedom;
TRE =text reading efficiency; Pitch_SD= overall intonation contour; Pitch_SF= pitch changes at the sentence-final position; IntraP_ratio= intrasentential pause ratio; IntraP_duration= intrasentential pause duration; InterP_duration= intersentential pause duration
[a] Predictors: (Constant), TRE
[b] Predictors: (Constant), TRE, IntraP_ratio, IntraP_duration, InterP_duration,
[c] Predictors: (Constant), TRE, IntraP_ratio, IntraP_duration, InterP_duration, Pitch_SD, Pitch_SF

The findings appeared to suggest that reading prosody features did not contribute to the explanation of additional variance in reading comprehension beyond that was accounted for by TRE. However, previous studies in the L1 setting pointed out that reading prosody and TRE shared substantial variance (Benjamin & Schwanenflugel, 2010; May, 2014; Miller & Schwanenflugel, 2008). If it holds true in the current study, the role of reading prosody would be overshadowed by TRE because TRE was given priority in the order of entry. Therefore, it seemed necessary to examine how much reading prosody and TRE are overlapped with each other in this study.

Following the approach of prior studies (Benjamin & Schwanenflugel, 2010; May, 2014; Miller & Schwanenflugel, 2008), an additional hierarchical regression analysis was performed in the current study. This analysis aimed to examine the relationship between TRE as the DV and reading prosody as IVs. In this regression, decoding was assigned into the first entry and then pause-related variables into the second entry, and lastly, two pitch variables were entered into the third entry. Table 4.12 presents the summary of the analysis of variance for regression. The R-value was statistically significant at each step. Especially, it was shown that pause-related and pitch-related variables, namely reading prosody features together account for 43.8% of TRE.

Table 4.12

Overlap between Reading Prosody and TRE[7]

Model	R	R^2	Adj.R^2	SE	F	ΔR^2	$df1$	$df2$	Sig.F
Decoding[a]	.564	.319	.311	22.028	41.138	.319	1	88	.000***
Pause variables[b]	.848	.718	.705	14.407	81.376	.400	3	85	.000***
Pitch variables[c]	.870	.756	.736	13.561	87.844	.038	2	83	.003**

Note. ** $p < .01$; *** $p < .001$; df= degree of freedom;
[a] Predictors: (Constant), decoding
[b] Predictors: (Constant), decoding, IntraP_ratio, IntraP_duration, InterP_duration,
[c] Predictors: (Constant), decoding, IntraP_ratio, IntraP_duration, InterP_duration, Pitch_SD, Pitch_SF

This result indicated that the shared variance between TRE and reading prosody, particularly in relation to pause variables, would be quite substantial. In other words, these findings suggested that if TRE is treated as the primary factor in the hierarchical analysis, the unique variance of reading comprehension explained by reading prosody would not be captured. Therefore, to obtain a more in-depth understanding of the role of reading prosody in reading comprehension, an additional hierarchical regression analysis was conducted with the order of

[7] The assumption check was also carried out for this analysis. The analysis of standard residuals of TRE demonstrated that there are no outliers (*Std. Residual Min* = -2.14, *Std. Residual Max* =2.885). *Tolerance* and *VIF* of the other variables were within a reasonable limit, indicating that multicollinearity is not problematic [Pitch_SD, *Tolerance* = .878, *VIF* = 1.139; Pitch_SF, *Tolerance* = .876, *VIF* = 1.141; IntraP P_ratio, *Tolerance* = .262, *VIF* = 3.818; IntraP_duration, *Tolerance* = .213, *VIF* = 4.685; InterP_duration, *Tolerance* = .595, *VIF* = 1.680]. In addition, the assumption of independence was met (*Durbin-Watson* Value =2.398), and linearity and homoscedasticity were also detected in the straight-line scatter plot.

entry reversed, with pause variables entered first, pitch variables entered subsequently, and TRE last.

Table 4.13

Alternative Model for the Contribution of Reading Prosody to RC

Model	*R*	R^2	Adj.R^2	SE	*F*	ΔR^2	*df*1	*df*2	Sig.*F*
Pause variables[a]	.490	.240	.213	11.590	9.044	.240	3	86	.000***
Pitch variables[b]	.542	.294	.252	11.305	12.241	.054	2	84	.046*
TRE[c]	.594	.336	.304	10.899	20.833	.059	1	83	.007**

Note. * $p < .05$; ** $p < .01$; *** $p < .001$; *df*= degree of freedom; TRE =text reading efficiency;
[a] Predictors: (Constant), IntraP_ratio, IntraP_duration, InterP_duration
[b] Predictors: (Constant), IntraP_ratio, IntraP_duration, InterP_duration, Pitch_SD, Pitch_SF
[c] Predictors: (Constant), IntraP_ratio, IntraP_duration, InterP_duration, Pitch_SD, Pitch_SF, TRE

The results of the analysis are presented in Table 4.13. The findings indicated that when pause factors were entered as predictors prior to pitch variables and TRE, they accounted for a significant amount of the variance in reading comprehension [R^2 = . 240, $F(3,86)$ = 9.044, p = .000]. Pitch variables also played a significant role in accounting for the variance in reading comprehension, thus increasing the predictability of the regression model. [ΔR^2 = .054, $F(2,84)$ =12.241, p = .046]. Furthermore, the predictability of TRE remained significant, explaining the unique variance of reading comprehension [ΔR^2 = . 059, $F(1,83)$ = 20.833, p = .007]. In summary, reading prosody appeared to play a significant predictor of reading comprehension, even when it was entered into the analysis with TRE in this model.

Despite the significant variance explained by the reading prosody features, the comparison of the two models of reading comprehension as DV showed that the former model whereby TRE was entered before reading prosody has greater predictability of reading comprehension than the latter one. The findings showed that while the initial HLR model displayed better predictive accuracy, the results could differ based on how much the TRE and reading prosody features overlapped.

To encapsulate, in the initial HLR model where TRE was entered earlier than reading prosody features, TRE and reading prosody variables jointly accounted for 35.3% of the variance in reading comprehension. In this model, TRE was the only significant variable of reading comprehension while other reading prosody features were not. However, considerable overlap between reading prosody and TRE was detected, and thus the subsequent HLR model, which prioritized the entry of reading prosody features before TRE was carried out. The results revealed that when the entry was reversed, both TRE and reading prosody variables significantly predicted reading comprehension, with a cumulative predictability of 33.6%.

4.4.4 Summary of the Results on Predictability of Reading Prosody for Reading Comprehension

To summarize the main points, this section was devoted to exploring how reading prosody plays in reading comprehension. Specifically, to what extent reading prosody can account for reading comprehension, and if it could explain a significant amount of variance in reading comprehension, how it acts on reading comprehension were investigated. In addition, it was also investigated whether reading prosody could account for reading comprehension when it was analyzed with TRE in the regression.

The results indicated that reading prosody could account for about 13 percent of unique variance in reading comprehension beyond what is explained by decoding skill. Also, it was revealed that reading prosody acts as a mediator between decoding skill and reading comprehension, facilitating the reading comprehension process.

However, it turned out that when reading prosody was entered into the regression analysis with TRE, reading prosody features were not significant predictors of reading comprehension. The further regression analyses indicated that the insignificant variance was due to the substantial overlap between reading prosody and TRE, masking the effect of reading prosody on reading comprehension in the regression analysis. Hence, when reading prosody was entered into the analysis before TRE, it turned out that reading prosody remained significant. The results indicated that despite the greater predictive performance observed in the first HLR model, the outcome may vary depending on the degree of overlap between TRE and reading prosody features or the developmental stage of reading proficiency in L2 readers.

Chapter 5. Discussion

In this chapter, the major findings of the research presented in Chapter 4 are discussed. Section 5.1 discusses characteristics of reading prosody features which are different depending on L2 readers' word decoding skills and text reading efficiency. The pertinent discussions about how TRE and reading prosody constitute ORF in L2 reading are presented in Section 5.2. Lastly, the contributions of reading prosody features to reading comprehension are discussed in Section 5.3.

5.1 Different Characteristics of Reading Prosody Features as a Function of Fluency Skills

The primary objective of the current study was to investigate the relationship between reading prosody features and fluency skills among L2 readers. The results revealed that there were significant differences in most of the reading prosody features among the three distinct subgroups of fluency skills, indicating the crucial role that fluency skills play in the development of reading prosody in L2 readers.

The study suggests that as L2 readers improve their word and text reading abilities, their reading proficiency evolves towards greater fluency. This is characterized by a reduction in the frequency and duration of pauses, coupled with an improved capacity to employ appropriate phrasing, manifesting as pauses at syntactic junctures and diminished pauses between successive sentences. In addition, skilled L2 readers demonstrate their competence by ending declarative sentences with a recognizable falling pitch. On the other hand, less proficient L2 readers with lower decoding and text-reading skills exhibit hesitant and uneven reading, marked by long and sporadic pauses within a sentence and a word, failing to chunk groups of words into meaningful units based on syntactic structures, resulting in lengthy pauses between sentences, and flat pitch changes at the end of sentences.

The overall characteristics of prosodic reading identified in this study among L2 readers are consistent with those found in previous L1 research (Benjamin & Schwanenflugel, 2010; May, 2014; Miller & Schwanenflugel, 2006, 2008; Schwanenflugel et al., 2004). For instance, Miller and Schwanenflugel (2006) examined individual differences in prosodic qualities among L1 third-graders reading declarative sentences, and found that students who demonstrated quick and accurate oral reading tended to make brief pauses both within a sentence and between sentences. These students also demonstrated a more substantial pitch drop at the end of declarative sentences.

Specifically, the current study revealed significant differences in pause ratio and duration across the three reading fluency groups, as indicated by the effect size of the pause ratio in distinguishing groups with different decoding skills. The pause ratio effect size was substantial, explaining over 40% of the total variance, indicating a strong association between fluency skills and pause ratio. This finding aligns with previous research conducted on L1 reading prosody, which also supports the importance of pause ratio in reading fluency (Arcand et al., 2014; Benjamin & Schwanenflugel, 2010; Benjamin et al., 2013; Binder et al., 2013; Kim et al., 2021a; Miller & Schwanenflugel, 2006, 2008; Schwanenflugel et al., 2015; Álvarez-Cañizo et al., 2018; Álvarez-Cañizo et al., 2020). In a series of studies conducted on L1 reading prosody and its relationship with other reading skills, Miller and Schwanenflugel (2006, 2008) found that frequent and longer pauses can be attributed, in part, to a lack of decoding skills. This means that readers with slow and inaccurate oral reading skills may not have developed automatic word recognition skills and may need to rely on phonological rules to read words, resulting in pauses and longer reading times. In fact, a similar phenomenon was observed in the current study. The participants with lower or intermediate fluency skills tended to struggle with the pronunciation of certain words, resulting in wavering or prolonged pauses, particularly before multisyllabic or low-frequency words (e.g., *appearances, feathers, frightened, opposite, toddling,* and *paddling*).

Another potential explanation for the observed relationship between pause patterns and reading skills is the individual differences in syntactic parsing efficiency. Previous studies have suggested concurrent relationships between reading prosody, particularly pause patterns, and syntactic structures (Cooper &

Paccia-Cooper, 1980; Ferreira, 1993; Goldman-Eisler, 1972; Koriat et al., 2002; Levelt, 1989). Moreover, previous findings on variant pause patterns based on pause types and reading skills (Arcand et al., 2014; Binder et al., 2013; Álvarez-Cañizo et al., 2015, 2018) suggested that fluent readers with efficient parsing skills tend to pause primarily at the phrase or clause boundaries, resulting in fewer and shorter pauses. In contrast, less fluent readers may struggle to process syntactic structures automatically, leading them to pause more frequently and randomly throughout the sentence, irrespective of the grammatical structure. The findings of the current study on the pattern of UGP_ratio seem to lend support to these previous findings. Significant distinctions in the frequency of ungrammatical pauses were observed among the three proficiency groups, with a notable decrease corresponding to higher levels of reading proficiency.

Therefore, based on the aforementioned findings from prior research, it can be inferred that the pause patterns observed in the oral reading of Korean L2 high school students could potentially be attributed to their proficiency in syntactic parsing or decoding abilities.

The current finding that more fluent readers exhibited more drastic pitch declinations at the end of sentences is consistent with previous research on L1 readers. (Benjamin & Schwanenflugel, 2010; Benjamin et al., 2013; Kim et al., 2021a; Miller & Schwanenflugel, 2006, 2008; Schwanenflugel et al., 2015; Álvarez-Cañizo et al., 2015). However, unlike previous studies, the results of the current study showed that only the pitch drops of the higher fluency group were noticeably different from other groups, suggesting that attaining a certain level of good fluency skills is required to show distinctive pitch drops at the sentence-final position.

This result can be interpreted based on earlier theories and research on prosody, which suggest that pitch conveys cognitive and paralinguistic information to listeners. (Benjamin & Schwanenflugel, 2010; Cowie et al., 2002; Gussenhoven, 2004; Himmelmann & Ladd, 2008; Ladd, 1984; Miller & Schwanenflugel, 2006). For example, it has been found that initial pitch rise and final declination in declarative sentences represent the reader's advanced planning and comprehension of the text (Gussenhoven, 2004; Tan et al., 2006). This finding, also, aligns with another research demonstrating that pitch rises are typically employed to convey uncertainty in yes-no questions, and the reverse pattern in reading declarative sentences indicates certainty (Miller & Schwanenflugel, 2006; Tan et al., 2006). Accordingly, the finding of the current study that higher fluency readers exhibited a more pronounced pitch declination at the end of a sentence may show that they had a greater level of certainty in their comprehension of the text, insinuating their higher reading comprehension ability.

An alternative perspective can be derived from the framework of the automaticity theory (LaBerge & Samuels, 1974) and verbal efficiency theory

(Perfetti, 1988). According to the theories, efficient word reading skills are a prerequisite for successful reading comprehension. It follows that readers with lower fluency skills may not attain sufficient efficiency in word reading, which could impede their ability to comprehend the text. This potential source of reading difficulties may have contributed to the relatively flat pitch differences observed at the end of sentences in readers with lower or intermediate fluency skills, and could explain why discernable differences between these groups were not found.

While it may be plausible to tentatively interpret these results as suggesting that reading comprehension affects or determines the characteristics of reading prosody, caution must be exercised given that there is insufficient evidence for the directionality of the relationship between reading prosody and reading comprehension.

Lastly, the overall intonation contour of L2 readers in this study was not significantly different depending on their fluency skills. This finding is contradictory to several L1 studies which reported that overall intonation pattern is indicative of the different degrees of reading skill (Benjamin et al., 2013; Miller & Schwanenflugel, 2006, 2008; Schwanenflugel et al., 2004). This inconsistency may be attributed to the effect of the L1 language (i.e., Korean) on English intonation. The research on Korean EFL students' speaking prosody has reported that Korean English learners tend to have relatively monotonous and narrow intonation contour compared to English native speakers (Kang & Johnson, 2018; Kang & Rhee, 2011; Park & Rhee, 2018; Rhee et al., 2003). Unlike other suprasegmental features, the indices for intonational changes such as F_0 range or slope have failed to discern L2 speaking proficiency (Kang, 2015; Kang & Ahn, 2012) and ORF (Park & Rhee, 2018). The findings of the current research mirror those of these previous L2 research in that the overall intonation variation was insensitive to the L2 readers' decoding skills and reading fluency.

Moreover, upon analyzing the mean Pitch_SD values of three distinct groups, it was found that the original hypothesis regarding the overall intonation contour of L2 readers was not supported. Specifically, a smaller mean Pitch_SD was observed in more proficient readers, contrary to expectations [Pitch_SD, M_L=1.733; M_H= 1.669]. Although the differences among the groups were not statistically significant, this outcome may be attributed to the phenomenon whereby skilled readers tend to read text more smoothly and with greater emphasis on higher-level processes, while making less conscious effort to stress each word appropriately. In contrast, slow and dysfluent readers tend to read in a word-by-word manner, consciously stressing words and thus compromising higher-level processes. As a result, some lower-level students may produce similar or even more fluctuating intonational contours in their reading that higher-level students did.

Taken together, these findings seem to suggest that intonation contour can hardly be an appropriate index for reading fluency or reading comprehension for L2 readers. However, a note of caution is necessary since intonation was not analyzed at the word or phrase level in this study. More in-depth observation of this prosodic factor or different approaches to measuring the data may be needed to figure out the role of intonation or the relation with other reading skills.

To summarize this section, the findings of the study indicate that the pause-related features were more susceptible to the impact of decoding skills and reading fluency skill compared to pitch-related features. Specifically, readers who exhibit higher levels of fluency tend to exhibit shorter and less frequent pauses, while readers who display lower fluency tend to exhibit longer and more frequent pauses. On the other hand, the differences in pitch-related features between higher and lower fluency readers were found to be less significant, potentially due to their varying levels of fluency or the influence of their native language on English intonation, highlighting the need for more in-depth observation of intonation at the word or phrase level.

Furthermore, these findings highlight the role of decoding skills in reading efficiency and more importantly prosodic reading for L2 reading. The prominent role of decoding skills in reading acquisition has been widely acknowledged in theoretical models and empirical studies (e.g., Hoover & Gough, 1990; Kim, 2020a, 2020b). The current study extends this notion to L2 reading by demonstrating its impact on the prosodic aspects of reading aloud.

5.2 Oral Reading Fluency as a Unitary Construct

The present study aimed to ascertain the extent to which the definition of ORF as formulated in L1 literature may be applicable to the L2 ORF. Specifically, the study explored the possibility of a unitary-construct model for ORF in L2 reading, by examining the relationship between text reading efficiency and prosodic variables. Alternatively, the study sought to determine whether ORF could be more accurately represented as two distinct, independent constructs.

The findings of the present study indicate that L2 ORF is a cohesive construct comprising both TRE and prosodic reading features even though only pause-related variables were reliable indicators for ORF among the assessed prosodic features. It suggests that ORF is a unitary construct with widely agreed-upon aspects such as accuracy, rate, and reading prosody in L2 reading as well as L1 reading.

This finding of the present study extends the findings of a prior study that investigated L1 readers' ORF structure. The study conducted by Kim et al. (2021b) showed that ORF is composed of three factors: a general ORF factor, three local factors (i.e., Prosody: Ratings and Pause, Prosody: Pitch, and Text Reading Efficiency), and two specific factors (i.e., Ratings and Pause). Although the model formulated in this previous study is much more complex, and even far from parsimonious, a fundamental similarity shared by this study is that among the three kinds of factors, the general ORF factor was found to be the most dependable in capturing the common ability of the observed variables of TRE and prosodic features.

A primary divergence between the present investigation and the prior study pertained to the consideration of pitch-related factors as meaningful indicators of ORF. This difference may be due to the fact that pitch-related variables, such as overall intonation contour and sentence-final pitch changes, may be affected by the Korean EFL reader's L1. Unlike pause features, which are more universal across languages (Binder et al., 2013; Fernandes et al., 2018; Miller & Schwanenflugel, 2006; Álvarez-Cañizo et al., 2018), pitch factors may vary depending on the language being read as explained in the previous section.

To sum up, the findings of this study demonstrate that in L2 reading as well as L1 reading, ORF is a unitary construct that encompasses various sub-skills in and of itself. Also, the variability of the relationship between reading prosody features and text reading efficiency further underscores the need to consider multiple dimensions of reading prosody when assessing overall reading fluency. These findings resonate with Wolf & Katzir-Cohen's (2001) statement that "the unsettling conclusion is that reading fluency involves every process and subskill involved in reading (p.220)." Both the current study and Wolf & Katzir-Cohen's statement suggest that achieving reading fluency requires proficiency in various areas, including decoding skills and prosodic reading skills, and that all these skills are interconnected and influence each other. Therefore, it is paramount to consider the broader context of reading fluency and the need to develop a comprehensive understanding of the various factors that contribute to it.

5.3 Roles of Reading Prosody in Reading Comprehension

The third research question addressed to what extent reading prosody could account for reading comprehension, and the analyses were carried out threefold to discuss the relationship with reading comprehension under the following three themes: i) the extent to which reading prosody accounts for unique variances in reading comprehension beyond decoding skills, ii) whether reading prosody serves as a mediator between reading comprehension and decoding skills, and iii) whether reading prosody contributes additional variances to reading comprehension even when text reading efficiency was considered.

The present study revealed that a combination of reading prosody and decoding skills accounted for 35.5% of the variance in reading comprehension. Specifically, pause-related variables explained an additional 8% of the variance, while pitch variables accounted for 5.2%. These findings are consistent with previous research (Benjamin & Schwanenflugel, 2010; Binder et al., 2013; Lai et al., 2014; May, 2014) which has consistently suggested a strong relationship between prosodic reading and reading comprehension. For instance, Benjamin and Schwanenflugel (2010) identified a significant relationship between prosodic reading and reading comprehension, accounting for 5.5% of the unique variance in reading comprehension. Similarly, Binder et al. (2013) found that better prosodic reading skills corresponded with higher reading comprehension scores, with prosody measures explaining an additional 13.3% of the variance in reading comprehension. Additionally, Lai et al. (2014) discovered that children with dyslexia commonly lacked prosodic reading skills, and that improvements in these skills were linked to improved reading comprehension abilities.

This finding of the current study seems to suggest that prosodic reading skills play a crucial role in reading comprehension. This perspective was informed by prior research suggesting that prosodic reading could offer crucial syntactic and semantic cues to the reader, potentially improving their comprehension (Frazier et al., 2006; Kuhn et al., 2010; Kuhn & Stahl, 2003). The results of the mediation effect of reading prosody in the present study provide more support for the role of reading prosody between reading comprehension and decoding skills.

The mediation model includes a direct path from decoding skills to reading comprehension. Additionally, the influence of word decoding skills on prosodic reading is acknowledged by incorporating paths between decoding skills and the prosodic factors (i.e., PAUSE). The results of the mediation analysis revealed that reading prosody plays a partial mediating role in the relationship between decoding and reading comprehension.

These findings seem to lend support to the notion that fluent word decoding skills are crucial for freeing up attentional resources that can then be used for prosodic reading (LaBerge & Samuels, 1974). This, in turn, allows prosody to independently contribute to better reading comprehension by providing important linguistic feedback to the L2 readers. However, only pause variables demonstrated a significant indirect relationship with reading comprehension. These findings seem to suggest that the idea of prosody as an additional scaffold for reading comprehension has limited support in L2 reading. Meanwhile, it should be noted that previous studies also have found that only pitch-related variables such as overall intonation contour and sentence-final pitch changes were significant predictors of reading comprehension.

For this phenomenon in which different prosodic variables are more related to reading comprehension in L1 and L2, previous studies shed light on its interpretation. Wolters et al. (2020) hypothesized that the relationship between reading prosody and reading comprehension may vary depending on the specific prosodic features being examined. In fact, empirical studies have shown that prosodic features that primarily reflect decoding skills, such as inappropriate or ungrammatical pauses, are expected to have a stronger association with reading comprehension in the early stages of reading development, while this relationship is likely to attenuate as reading skills become more advanced

(Fernandes et al., 2018; Kim et al., 2021b). Conversely, prosodic features that predominantly reflect semantic processing, such as child-adult pitch match or sentence-final declination, are predicted to have a weaker association with reading comprehension in the early stages of reading development but become more closely linked to reading comprehension as reading skills mature (Binder et al., 2013; Kim et al., 2021a; Miller & Schwanenflugel, 2008; Schwanenflugel et al., 2004). Especially, the study by Kim et al. (2021b) offers insight into how the relationship between prosodic reading and reading comprehension develops over time. The results showed that there was a steady increase in the relationship between the pitch factor and reading comprehension over a period of three years (from the beginning of Grade 1 to the end of Grade 3). This finding suggests that the relationship between prosodic reading and reading comprehension may change as children's decoding skills develop, allowing for more cognitive resources to be devoted to semantic processing (LaBerge & Samuels, 1974). As a result, the relationship between pitch-related prosodic features and reading comprehension may become stronger over time.

Viewed through this scholarly lens, the outcomes of the current investigation propose that the subjects, comprising Korean high school students in the 12th grade, did not manifest an adequate level of decoding fluency, thereby limiting the capacity to allocate cognitive resources to the domain of pitch-related dimensions within the realm of reading comprehension. It is posited that an augmentation in their “verbal efficiency” (Perfetti, 1985) or the attainment of a state of “automaticity in word recognition” (LaBerge & Samuels, 1974) could potentially facilitate the exertion of pitch-related factors on the landscape of their L2 reading comprehension.

Alternatively, as discussed in section 5.1, it is possible that the influence of their L1 may have contributed to overall monotonous intonation, thereby rendering pitch variables insignificant in the present study. However, it is important to exercise caution when interpreting these results, given the limited scope of the study, which only involved 3rd-grade high school students. Further investigation is warranted to gain a more comprehensive understanding of the nature and significance of pitch variables in L2 reading.

Lastly, the present investigation sought to assess the independent contribution of reading prosody to reading comprehension after controlling for TRE. The analyses were carried out in two ways, whereby TRE and reading prosody were entered in different orders.

When TRE was entered before reading prosody, it was found that reading prosody did not significantly predict reading comprehension. This result was derived from the substantial overlap between reading prosody and TRE, which in turn may have masked the unique contribution of reading prosody to reading comprehension in the initial analysis. Therefore, when they were entered into the analysis in the reversed order, the findings revealed that the reading prosody variables, including both pauses and pitches, were statistically significant predictors of reading comprehension. In this analysis, TRE still remained significant even beyond the variance that was accounted for by reading prosody.

Although reading prosody remains significant in predicting reading comprehension in the latter model, the comparison of predictability of the two regression models indicated that the former model where TRE was entered in the first block accounted for a greater variance in reading comprehension. These findings may be construed as suggesting that, when evaluating the ORF of Korean EFL high school students, assessing only the TRE component without taking into account reading prosody could yield sufficient information regarding their reading abilities. However, the cautious interpretation of this finding is imperative, given the observed discernible developmental trend of TRE and its evolving correlations with reading comprehension.

Regarding the developmental trajectory of TRE, Fuchs et al. (2001) have noted that TRE improves during the elementary grades but the rate of improvement decreases over time. Previous literature, such as the works of Hasbrouck and Tindal (2006, 2017), provides evidence supporting this pattern of growth. This developmental trend is also evident in the correlation between reading comprehension and TRE. It has been shown that the correlation between TRE and reading comprehension diminishes as students progress to higher grades (Denton et al., 2011; Schwanenflugel & Knapp, 2015). Additionally, longitudinal studies examining the predictors of reading comprehension have indicated that, for students in higher grades, semantic skills (Vellutino et al., 2007) or language comprehension (Catts et al., 2005) have a stronger influence on reading comprehension compared to TRE.

Based on the existing literature in L1, it is plausible to infer that Korean EFL high school students may not have attained a sufficient level of reading proficiency. Consequently, this could result in a strong relation between TRE and reading comprehension, potentially overshadowing the influence of reading prosody in the regression analysis. Conversely, if they exceed the minimum or a specified threshold of WCPM suggested by Paris and Hamilton (2009), the effect of reading prosody on reading comprehension may be borne out.

Therefore, although these findings revealed that measuring TRE alone could be a sufficient "proxy" of reading comprehension for L2 readers, caution is needed in interpretation. The role of reading prosody may vary depending on the developmental stage of reading prosody features in L2 readers or the degree of overlap between TRE and reading prosody features.

Chapter 6. Conclusion

The present dissertation aimed to investigate the various roles of reading prosody in relation to other reading skills, such as decoding skills, text reading efficiency, and reading comprehension, among Korean EFL high school students. This chapter presents the key findings of the study and discusses their pedagogical implications, as well as limitations and recommendations for further research.

6.1 Major Findings

The present research aimed to investigate the significance of reading prosodic features in L2 reading. The study has three main objectives: firstly, to identify distinct characteristics of different reading prosody features that vary according to the level of L2 fluency skills; secondly, to test the hypothesis regarding the definition of ORF, which comprises both reading prosody and text reading efficiency (i.e., reading rate and accuracy); and thirdly, to explore the extent to which reading prosody contributes to L2 reading comprehension. To accomplish these objectives, the study employs spectrographic analysis as the methodological approach to identify and extract various reading prosody features. The following sections provide a summary of the major findings of the study.

The findings of this study suggest a close relationship between the reading prosody characteristics of L2 readers and their level of L2 fluency skill. Specifically, the study revealed that various pause features, such as pause ratio, ungrammatical pause ratio, and duration of intrasentential pauses, demonstrated a decreasing trend as the reading proficiency of L2 readers improved.

Furthermore, significant variations in pitch changes at the sentence-final position and pause duration across sentences appeared only evident among higher fluency groups, indicating that these changes primarily manifest among L2 readers who have attained a sufficient level of fluency skills. On the other hand, the overall intonation contour (i.e., Pitch_SD) did not exhibit any significant differences based on fluency level, suggesting that while word and text reading skills may be fluent, the overall pitch pattern may remain similar among these various proficient groups of Korean EFL high school students.

With respect to the interplay between the constructs of ORF, which encompasses text reading efficiency and reading prosody, the study revealed that ORF in L2 is a unified construct comprising underlying sub-skills. This result suggests that ORF is a construct that shares commonly agreed-upon aspects such as accuracy, rate, and reading prosody in both L1 and L2 reading.

Finally, the current study demonstrated that reading prosody accounts for a significant proportion of the variance in reading comprehension, even after controlling for decoding skills. Through mediation analysis, it was also found that reading prosody acted as a partial mediator between decoding and reading comprehension. These findings provide evidence that prosodic reading skills play a critical role in reading comprehension, specifically by facilitating comprehension.

However, another finding revealed that the inclusion of reading prosody as a predictor in the hierarchical linear regression did not account for a significant amount of variance in reading comprehension when combined with TRE. The absence of statistical significance was attributed to the substantial convergence between TRE and the dimensions of reading prosody. As a result, a subsequent hierarchical linear regression analysis was undertaken, in which both constituent elements of reading prosody and TRE retained their notable significance.

6.2 Pedagogical Implications

The results of the thesis have significant implications for L2 reading instruction. The assessment of reading prosody is increasingly recognized by reading researchers as an essential component of reading fluency assessment (Dowhower, 1991; Fuchs et al., 2001; Hudson et al., 2005; Hudson et al., 2008; Kuhn et al., 2010). However, the question remains regarding how the assessment of reading prosody can be incorporated into reading fluency practices. In this regard, the study's findings on the spectrographic evidence of reading prosody features have significant pedagogical implications.

The research could contribute to the development of valid ORF assessment tools based on the observation of reading prosody. The Comprehensive Oral Reading Fluency Scale (CORF), developed in the L1 context, is an example of such a tool (Benjamin, 2012; Benjamin et al., 2013). This scale includes items such as "varies tone and pitch to reflect emotions in the text" and "uses pauses and changes in loudness to convey meaning," which were identified using spectrographic analysis. Also, the validation process of the spectrographically grounded scale demonstrated high inter-rater reliability and internal consistency, indicating its potential to be used as a valid and reliable tool for evaluating oral reading prosody. Incorporating reading prosody assessment would require an additional two minutes per reading, but this investment could lead to more accurate assessments of students' reading fluency and comprehension. Therefore, the significance of this study lies in its potential to provide an objective measurement of L2 students' oral reading prosody and contribute to the development of valid ORF assessment tools in educational settings.

The use of ORF reading scales based on reading prosody can provide practical and reliable diagnostic information that can help identify the word callers or gap fillers among L2 readers (Quirk & Beem, 2012). A child who reads with good automaticity but without appropriate expression linked to the text message may be considered a word caller, indicating a discrepancy between their decoding ability and comprehension of the text. This type of reading behavior can be problematic as it can impede a child's reading comprehension and hinder their overall literacy development. However, by observing and using reading prosody, it is possible to provide appropriate intervention and support to these struggling readers. Accurate assessment of reading prosody can help teachers identify areas where students need additional instruction and support and can inform the development of targeted interventions aimed at improving reading comprehension and fluency. In this way, the use of reading prosody assessment can play a critical role in helping L2 readers achieve academic success.

Additionally, in light of the study's finding on the crucial role of reading prosody in reading comprehension, it is recommended that reading instruction should prioritize the development of students' reading prosody. A promising intervention for improving reading prosody is repeated reading, which involves multiple readings of the same passage until a satisfactory level of fluency is achieved. Several studies have demonstrated the efficacy of repeated reading in enhancing reading rate, word recognition, and reading prosody. For instance, studies conducted by Dowhower (1987), Ryu (2020), and Therrien (2004) have reported significant improvements in appropriate intonation patterns and a reduction in inappropriate pauses among readers who engaged in repeated reading. Specifically, Dowhower (1987) noted a decrease in pitch at the end of declarative sentences, highlighting the potential benefits of repeated reading in developing appropriate reading prosody. In addition to repeated reading, there are other instructional approaches that could be effective in enhancing reading prosody. For instance, Rasinski (2006) recommends that teachers model good prosodic reading when reading aloud to students and provide them with ample opportunities to practice reading with expression. Teachers can also guide students to pay attention to punctuation, phrasing, and tone while reading aloud, and provide feedback and praise for expressive reading.

Lastly, in the realm of reading models, such as the DIER model (Kim, 2020a, 2020b) or the multidimensional view of ORF (Hudson et al., 2008), ORF is recognized as a valid and comprehensive construct that encompasses the entirety of the reading process (Wolf & Katzir-Cohen, 2001). It serves not only as a means to assess students' reading progress but also as a fundamental aspect of the act of reading itself. Notably, the present findings indicate that Korean EFL high school students at the tertiary level of secondary education demonstrate the ability to read materials at the 4th to 6th-grade level at an approximate rate of 105 WCPM. This reading rate aligns with the 25th percentile of 5th-grade students' ORF, as determined by the L1 ORF norms (Hasbrouck & Tindal, 2017). In other words, these Korean students have not yet attained a mature stage of reading development. It is evident that these Korean students are still in need of further development in their reading skills to reach a more advanced stage of reading competence. Additional support and targeted instruction in areas such as decoding practice and enhancing ORF are essential to foster their reading growth and promote their progression toward a mature stage of reading proficiency. Therefore, in the context of L2 reading, educational practices aimed at fostering ORF and developing proficient reading prosody assume critical significance.

Taken together, the findings of the thesis have important pedagogical implications for reading instruction. The use of ORF reading scales based on reading prosody can provide practical and reliable diagnostic information, and repeated reading is a promising intervention for improving reading fluency and prosody. By implementing these strategies, teachers can help students become more expressive readers and improve their comprehension.

6.3 Limitations and Suggestions for Future Research

Despite the contribution of the current thesis, it is not without limitations. The results of this study consistently indicate that overall pitch contour is not an indicator of reading comprehension and text reading fluency. As described in the methodology section, the variability of pitch contour proved to be valid in the previous studies (Benjamin et al., 2013; Cowie et al., 2002). However, this variable may not capture the differences in L2 readers' pitch variation because some slow readers could read a text in a word-by-word manner by putting stress on every word they encounter, while fluent readers can read a text fast without much variability of pitch pattern. Alternatively, another approach that could have been taken to further investigate the impact of pitch contour on L2 readers would be to compare more proficient L2 speakers' or native speakers' overall intonation contours with L2 readers' intonation (Miller & Schwanenflugel, 2006). This would provide a useful comparison and allow for a more nuanced understanding of L2 readers' pitch contour.

Also, given the uncertainties exist due to a lack of L2 research on reading prosody, it is also recommended that multiple pitch indicators (e.g., pitch range or pitch slope) be employed to ensure the safety and reliability of the results when examining the relationship between pitch and reading ability. Examining and including a variety of variables could provide more reliable information about L2 readers' pitch changes and their impact on reading ability. Therefore, future studies should consider incorporating additional variables to provide a more comprehensive understanding of the factors that impact reading abilities.

In the present study, the hierarchical regression analysis was used to examine the causal effect of reading prosody on text reading efficiency or reading comprehension. While regression analysis is a powerful statistical technique, it has some limitations when it comes to modeling complex relationships between multiple variables. In this regard, the use of more advanced statistical models, such as structural equation modeling (SEM), could be beneficial in future research. SEM allows for the estimation of multiple regression equations simultaneously and provides information about the direct and indirect effects of each variable on the outcome of interest. This can help to reveal more complex relationships between variables and provide a more comprehensive understanding of the factors that influence reading comprehension.

Another limitation of this study is that more extensive observations and in-depth investigation of each reading prosody feature were needed. Given the paucity of research on reading prosody in L2 reading studies, such research is crucial. However, this study could not accomplish this as a statistical model was required to investigate the relationship between reading prosody and other reading ability variables while ensuring the adequate proportionality of dependent and independent variables. Therefore, it is recommended that future research scrutinize more various aspects of pitch, including pitch differences based on the position in the sentence or sentence type (e.g., interrogative, declarative, and exclamatory), and the roles of various pauses based on sentence type, position, and learner reading ability, akin to the studies by Álvarez-Cañizo et al. (2018, 2020). Although this study centered on the relationship between reading prosody and other reading ability (e.g., TRE or reading comprehension) while compromising a more detailed analysis of prosody, forthcoming research should explore the characteristics of each reading prosody component.

Furthermore, future research could also explore the relationship between reading prosody and other factors that may influence reading comprehension, such as vocabulary knowledge, background knowledge, and working memory. Understanding the complex interplay between these different factors could provide valuable insights into the mechanisms underlying the reading process in L2.

Also, although current research focused on the reading prosody of Korean EFL high school students, expanding the study to include lower levels of graders could be useful in understanding how reading prosody and comprehension skills develop over time. It would allow for a comparison of the progression of these skills across different levels of experience and the identification of any patterns or trends that may emerge. Additionally, a longitudinal design would make it possible to track the same group of students over an extended period of time, providing a more complete picture of their reading prosody development.

References

Adams, M. J. (1990). Beginning to read: Thinking and learning. *Psychological Review*, *65*, 197-208.

Ahn, H., & Kang, Y. (2016). Reading fluency and listening comprehension abilities as predictors of reading comprehension. *English teaching*, *71*(1), 3-24.

Al Otaiba, S., Petscher, Y., Pappamihiel, N. E., Williams, R. S., Dyrlund, A. K., & Connor, C. (2009). Modeling oral reading fluency development in Latino students: A longitudinal study across second and third grade. *Journal of educational psychology*, *101*(2), 315.

Allbritton, D. W., McKoon, G., & Ratcliff, R. (1996). Reliability of prosodic cues for resolving syntactic ambiguity. *Journal of experimental psychology: Learning, Memory, and Cognition*, *22*(3), 714.

Allington, R. L. (1983). Fluency: The neglected reading goal. *The Reading Teacher*, *36*(6), 556-561.

Arcand, M.-S., Dion, E., Lemire-Théberge, L., Guay, M.-H., Barrette, A., Gagnon, V., . . . Fuchs, D. (2014). Segmenting texts into meaningful word groups: Beginning readers' prosody and comprehension. *Scientific Studies of Reading*, *18*(3), 208-223.

Ashby, J. (2006). Prosody in skilled silent reading: Evidence from eye movements. *Journal of Research in Reading*, 29(3), 318-333.

Backman, N. (1979). Intonation errors in second-language pronunciation of eight Spanish-speaking adults learning English. *Interlanguage studies bulletin*, 239-265.

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the academy of marketing science, 16*, 74-94.

Baker, D. L., Park, Y., & Baker, S. K. (2012). The reading performance of English learners in grades 1–3: The role of initial status and growth on reading fluency in Spanish and English. *Reading and Writing*, *25*(1), 251-281.

Baker, S. K., Smolkowski, K., Katz, R., Fien, H., Seeley, J. R., Kame'Enui, E. J., & Beck, C. T. (2008). Reading fluency as a predictor of reading proficiency in low-performing, high-poverty schools. *School Psychology Review*, *37*(1), 18-37.

Barker, T. A., Torgesen, J. K., & Wagner, R. K. (1992). The role of orthographic processing skills on five different reading tasks. *Reading Research Quarterly*, 335-345.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology,*

51(6), 1173.
Benjamin, R. A. (2012). *Development and validation of the comprehensive oral reading fluency scale.* [Unpublished doctoral dissertation]. University of Georgia.
Benjamin, R. G., & Schwanenflugel, P. J. (2010). Text Complexity and Oral Reading Prosody in Young Readers. *Reading research quarterly*, *45*(4), 388-404. https://doi.org/10.1598/RRQ.45.4.2
Benjamin, R. G., Schwanenflugel, P. J., Meisinger, E. B., Groff, C., Kuhn, M. R., & Steiner, L. (2013). A spectrographically grounded scale for evaluating reading expressiveness. *Reading research quarterly*, *48*(2), 105-133.
Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological bulletin, 107*(2), 238.
Binder, K. S., Tighe, E., Jiang, Y., Kaftanski, K., Qi, C., & Ardoin, S. P. (2013). Reading expressively and understanding thoroughly: An examination of prosody in adults with low literacy skills. *Reading and writing*, *26*(5), 665-680.
Boersma, P., & Weenink, D. (2023). Praat, a system for doing phonetics by computer (Version 5.0. 34)[Computer software]. In.
Bolanos, D., Cole, R. A., Ward, W. H., Tindal, G. A., Schwanenflugel, P. J., & Kuhn, M. R. (2013). Automatic assessment of expressive oral reading. *Speech Communication*, *55*(2), 221-236.
Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological methods & research*, *21*(2), 230-258.
Byrne, B. M. (2013). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. routledge.
Carlisle, J. F., & Katz, L. A. (2006). Effects of word and morpheme familiarity on reading of derived words. *Reading and Writing*, *19*, 669-693.
Catts, H. W., Hogan, T. P., & Adlof, S. M. (2005). Developmental changes in reading and reading disabilities. In *The connections between language and reading disabilities* (pp. 38-51). Psychology Press.
Chafe, W. (1988). Punctuation and the prosody of written language. *Written communication*, *5*(4), 395-426.
Chan, J. S., & Wade-Woolley, L. (2018). Explaining phonology and reading in adult learners: Introducing prosodic awareness and executive functions to reading ability. *Journal of Research in Reading*, 41(1), 42-57.
Chang, A. (2012). Improving reading rate activities for EFL students: Timed reading and repeated oral reading. *Reading in a foreign language*, 24 (1), p.56-83.
Chang, A. C. (2019). The Effects of Repeated Oral Reading Practice on EFL Learners' Oral Reading Fluency Development. *The Reading Matrix: An International Online Journal*, *19*(2).
Ching Pey, K., Low, H. M., & Lee, L. W. (2014). Relationship between Oral

Reading Fluency and Reading Comprehension among ESL Students. *Gema Online Journal of Language Studies*, *14*(3). https://doi.org/http://dx.doi.org/10.17576/GEMA-2014-1403-02

Clark, R., Morrison, T. G., & Wilcox, B. (2009). Readers' theater: A process of developing fourth-graders' reading fluency. *Reading Psychology*, *30*(4), 359-385.

Clay, M. M., & Imlach, R. H. (1971). Juncture, pitch, and stress as reading behavior variables. *Journal of Verbal Learning and Verbal Behavior*, 10(2), 133-139.

Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and speech*. Harvard University Press.

Coots, J. H., & Snow, D. P. (1981). *Comprehension Skills and Text Organization Ability in Reading*. Southwest Regional Laboratory for Educational Research and Development.

Cowie, R., Douglas-Cowie, E., & Wichmann, A. (2002). Prosodic characteristics of skilled reading: Fluency and expressiveness in 8-10-year-old readers. *Language and speech*, *45*(1), 47-82.

Crosson, A. C., & Lesaux, N. K. (2010). Revisiting assumptions about the relationship of fluent reading to comprehension: Spanish-speakers' text-reading fluency in English. *Reading and Writing*, *23*(5), 475-494.

Danne, M. C., Campbell, J. R., Grigg, W. S., Goodman, M. J., & Oranje, A. (2005). Fourth-Grade Students Reading Aloud: NAEP 2002 Special Study of Oral Reading. The Nation's Report Card. NCES 2006-469. *National Center for Education Statistics*. http://ies.ed.gov/pubsearch/pubsinfo.asp?pubid=2006469

de Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, *41*(2), 385-390. https://doi.org/10.3758/BRM.41.2.385

Deno, S. L., & Fuchs, L. S. (1987). Developing curriculum-based measurement systems for data-based special education problem solving. *Focus on Exceptional Children*, *19*(8), 1-16.

Denton, C. A., Barth, A. E., Fletcher, J. M., Wexler, J., Vaughn, S., Cirino, P. T., . . . Francis, D. J. (2011). The relations among oral and silent reading fluency and comprehension in middle school: Implications for identification and instruction of students with reading difficulties. *Scientific Studies of Reading*, *15*(2), 109-135.

Dowhower, S. L. (1987). Effects of repeated reading on second-grade transitional readers' fluency and comprehension. *Reading Research Quarterly*, 389-406.

Dowhower, S. L. (1991). Speaking of prosody: Fluency's unattended bedfellow. *Theory into practice*, *30*(3), 165-175.

Dowhower, S. L. (1994). Repeated reading revisited: Research into practice.

Reading & Writing Quarterly: Overcoming Learning Difficulties, *10*(4), 343-358.
Duke, N. K., & Carlisle, J. (2011). The development of comprehension. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. P. Afflerbach (Eds.), *Handbook of reading research, Vol. IV* (pp. 199–228). Routledge.
Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of research in reading*, 18(2), 116-125.
Fernandes, S., Querido, L., Verhaeghe, A., & Araújo, L. (2018). What is the relationship between reading prosody and reading comprehension in European Portuguese? Evidence from grades 2 to 5. *Journal of Research in Reading*, *41*, S102-S129.
Ferreira, F. (1993). Creation of prosody during sentence production. *Psychological review*, *100*(2), 233-253. https://doi.org/10.1037//0033-295X.100.2.233
Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed). Sage.
Fodor, J. D. (1998). Learning to parse? *Journal of psycholinguistic research*, *27*, 285-319.
Fodor, J. D. (2002). *Psycholinguistics cannot escape prosody.* Proceedings of the speech prosody 2002 conference, pp. 83–88.
Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research*, *18*(1), 39-50.
Frazier, L., Carlson, K., & Clifton, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in cognitive sciences*, *10*(6), 244-249.
Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, *5*(3), 239-256.
Grabe, W., & Jiang, X. (2013). Assessing reading. In A. J. Kunnan (Ed.), *The companion to language assessment* (Vol. 1, pp. 185–200). Wiley.
Graddol D. (1986). Discourse specific pitch behavior. In C. Johns-Lewis (Ed), *Intonation in Discourse* (pp. 221-237). Croom Helm.
Godde, E., Bosse, M. L., & Bailly, G. (2020). A review of reading prosody acquisition and development. *Reading and Writing*, 33(2), 399-426.
Good, R.H., & Kaminski, R.A. (Eds.). (2002). *Dynamic Indicators of Basic Early Literacy Skills (6th ed.)*. Eugene, OR: Institute for the Development of Education Achievement.
Goldman-Eisler, F. (1972). Pauses, clauses, sentences. *Language and speech*, *15*(2), 103-113.
Goodman, I., Libenson, A., & Wade-Woolley, L. (2010). Sensitivity to linguistic stress, phonological awareness and early reading ability in preschoolers. *Journal of Research in Reading*, *33*(2), 113-127.
Gorsuch, G., & Taguchi, E. (2008). Repeated reading for developing reading

fluency and reading comprehension: The case of EFL learners in Vietnam. *System*, *36*(2), 253-278.
Gorsuch, G., & Taguchi, E. (2010). Developing reading fluency and comprehension using repeated reading: Evidence from longitudinal student reports. *Language Teaching Research*, *14*(1), 27-59.
Goswami, U., Gerson, D., & Astruc, L. (2010). Amplitude envelope perception, phonology and prosodic sensitivity in children with developmental dyslexia. *Reading and Writing*, *23*, 995-1019.
Groen, M. A., Veenendaal, N. J., & Verhoeven, L. (2019). The role of prosody in reading comprehension: Evidence from poor comprehenders. *Journal of Research in Reading*, *42*(1), 37-57.
Guaïtella, I. (1999). Rhythm in speech: What rhythmic organizations reveal about cognitive processes in spontaneous speech production versus reading aloud. *Journal of Pragmatics*, *31*, 509–523.
Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge University Press.
Hamilton, C., & Shinn, M. R. (2003). Characteristics of word callers: An investigation of the accuracy of teachers' judgments of reading comprehension and oral reading skills. *School Psychology Review*, *32*(2), 228-240.
Harrington, D. (2009). *Confirmatory factor analysis*. Oxford university press.
Hasbrouck, J. E., & Tindal, G. (1992). Curriculum-based oral reading fluency norms for students in grades 2 through 5. *Teaching Exceptional Children*, *24*(3), 41-44.
Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The reading teacher*, *59*(7), 636-644.
Hasbrouck, J., Tindal, G. (2017). *An update to compiled ORF norms (Technical Report No. 1702)*. Behavioral Research and Teaching, University of Oregon. http://www.brtprojects.org/
Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications.
Henton, C. G. (1989). Fact and fiction in the description of female and male pitch. *Language & Communication, 9*(4), 299-311. https://doi.org/10.1016/0271-5309(89)90026-8
Hermes, D.J. and Van Gestel, J. 1991. The frequency scale of speech intonation. *Journal of the Acoustical Society of America* 90, 97–102.
Himmelmann, N. P., & Ladd, D. R. (2008). Prosodic description: An introduction for fieldworkers. *Language Documentation and Conservation,* vol. 2, pp. 244-274.
Hintze, J. M., & Silberglitt, B. (2005). A longitudinal examination of the diagnostic accuracy and predictive validity of R-CBM and high-stakes

testing. *School Psychology Review*, *34*(3), 372-386.
Hirst, D. J., & de Looze, C. (2021) Measuring Speech. Fundamental frequency and pitch. In R. Knight & J. Setter, *Cambridge Handbook of Phonetics* (pp.336-361). Cambridge University Press.
Hirotani, M., Frazier, L., & Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language*, *54*(3), 425-443.
Holliman, A. J., Wood, C., & Sheehy, K. (2008). Sensitivity to speech rhythm explains individual differences in reading ability independently of phonological awareness. *British Journal of Developmental Psychology*, *26*(3), 357-367.
Holliman, A. J., Wood, C., & Sheehy, K. (2010). The contribution of sensitivity to speech rhythm and non-speech rhythm to early reading development. *Educational Psychology*, *30*(3), 247-267.
Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and writing*, *2*(2), 127-160.
Hudson, R. F., Lane, H. B., & Pullen, P. C. (2005). Reading fluency assessment and instruction: What, why, and how?. *The Reading Teacher, 58*(8), 702-714.
Hudson, R. F., Pullen, P. C., Lane, H. B., & Torgesen, J. K. (2008). The Complex Nature of Reading Fluency: A Multidimensional View. *Reading & Writing Quarterly*, *25:1*, 4-32.
IBM Corp, N. (2019). IBM SPSS statistics for windows (Version 26.0.0.2)[Computer software]. Armonk, NY:IBM corp.
Jaccard, J., & Guilamo-Ramos, V. (2002). Analysis of variance frameworks in clinical child and adolescent psychology: Issues and recommendations. *Journal of clinical child and adolescent psychology*, *31*(1), 130-146.
Jarmulowicz, L., Hay, S. E., Taran, V. L., & Ethington, C. A. (2008). Fitting derivational morphophonology into a developmental model of reading. *Reading and Writing*, *21*, 275-297.
Jenkins, J. R., Fuchs, L. S., Espin, C., van den Broek, P., & Deno, S. L. (2000). Effects of task format and performance dimension on word reading measures: Criterion validity, sensitivity to impairment, and context facilitation. Pacific Coast Research Conference, San Diego, CA,
Jenkins, J. R., Fuchs, L. S., Van Den Broek, P., Espin, C., & Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology*, *95*(4), 719.
Jenkins, J. R., Hudson, R. F., & Lee, S. H. (2007). Using CBM-reading assessments to monitor progress. *The International Dyslexia Association. Perspectives on language and literacy*, *33*(2). 11-18.
Jeon, E. H. (2012). Oral reading fluency in second language reading. *Reading in a Foreign Language 24*(2). 186–208.

Kaiser, H. F., & Rice, J. (1974). Little jiffy, mark IV. *Educational and psychological measurement*, *34*(1), 111-117.
Kang, O., & Johnson, D. (2018). The roles of suprasegmental features in predicting English oral proficiency with an automated system. *Language Assessment Quarterly*, *15*(2), 150-168.
Kang, S. (2013). The Effect of Language Immersion on Second Language Intonation. *Journal of Pan-Pacific Association of Applied Linguistics*, *17*(1), 1-27.
Kang, S. (2015). A comparative study on the Korean suprasegmental acquisition for the native Mandarin, English, and Japanese learners of Korean. *Studies in Linguistics*, *37*, 1-23.
Kang, S., & Ahn, H. (2012). Cross-directional development of prosody. *Language Research*, *48*(1), 37-57.
Kang, S.-H., & Rhee, S.-C. (2011). A study on the Suprasegmental Parameters Exerting an Effect on the Judgment of Goodness or Badness on Korean-spoken English. *Phonetics and Speech Sciences*, *3*(2), 3-10.
Kang, Y. (2011). Reading fluency vs. oral language comprehension in predicting reading comprehension of Korean EFL learners. *Korean Journal of Applied Lingsuistics*, *27*(4), 81-103.
Kang, Y. (2020). Relative Contribution of Reading Fluency and Vocabulary Knowledge in Predicting Korean EFL Learners' Reading Comprehension. *Journal of Asia TEFL*, *17*(3), 778-790.
Kang, Y. (2021). Does the Simple View of Reading Explain Korean Elementary EFL Learners' Reading Comprehension? *English Teaching*, *76*(1), 57-78.
Keehn, S. (2003). The effect of instruction and practice through readers theatre on young readers' oral reading fluency. *Literacy Research and Instruction*, *42*(4), 40-61.
Kim, Y.-S. (2012). The relations among L1 (Spanish) literacy skills, L2 (English) language, L2 text reading fluency, and L2 reading comprehension for Spanish-speaking ELL first grade students. *Learning and individual differences*, *22*(6), 690-700.
Kim, Y. S. G. (2020a). Hierarchical and dynamic relations of language and cognitive skills to reading comprehension: Testing the direct and indirect effects model of reading (DIER). *Journal of Educational Psychology*, 112(4), 667.
Kim, Y. S. G. (2020b). Toward integrative reading science: The direct and indirect effects model of reading. *Journal of Learning Disabilities*, 53(6), 469-491.
Kim, Y.-S., Apel, K., & Al Otaiba, S. (2013). The relation of linguistic awareness and vocabulary to word reading and spelling for first-grade students participating in response to intervention.
Kim, Y.-S., Park, C. H., & Wagner, R. K. (2014). Is oral/text reading fluency a

"bridge" to reading comprehension? *Reading and writing*, *27*(1), 79-99.
Kim, Y.-S., Petscher, Y., Schatschneider, C., & Foorman, B. (2010). Does growth rate in oral reading fluency matter in predicting reading comprehension achievement? *Journal of Educational Psychology*, *102*(3), 652.
Kim, Y.-S., Wagner, R. K., & Foster, E. (2011). Relations among oral reading fluency, silent reading fluency, and reading comprehension: A latent variable study of first-grade readers. *Scientific Studies of Reading*, *15*(4), 338-362.
Kim, Y.-S., Wagner, R. K., & Lopez, D. (2012). Developmental relations between reading fluency and reading comprehension: A longitudinal study from Grade 1 to Grade 2. *Journal of experimental child psychology*, *113*(1), 93-111.
Kim, Y.-S. G., & Petscher, Y. (2016). Prosodic sensitivity and reading: An investigation of pathways of relations using a latent variable approach. *Journal of Educational Psychology*, *108*(5), 630.
Kim, Y.-S. G., Quinn, J. M., & Petscher, Y. (2021a). Reading prosody unpacked: A longitudinal investigation of its dimensionality and relation with word reading and listening comprehension for children in primary grades. *Journal of Educational Psychology*, *113*(3), 423.
Kim, Y.-S. G., Quinn, J. M., & Petscher, Y. (2021b). What is text reading fluency and is it a predictor or an outcome of reading comprehension? A longitudinal investigation. *Developmental Psychology*, *57*(5), 718.
Kim, Y.-S. G., & Wagner, R. K. (2015). Text (oral) reading fluency as a construct in reading development: An investigation of its mediating role for children from grades 1 to 4. *Scientific Studies of Reading*, *19*(3), 224-242.
Kim, Y. S. G. (2015). Developmental, component-based model of reading fluency: An investigation of predictors of word-reading fluency, text-reading fluency, and reading comprehension. *Reading research quarterly*, *50*(4), 459-481.
Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge university press.
Kitzen, K. R. (2001). *Prosodic sensitivity, morphological ability, and reading ability in young adults with and without childhood histories of reading difficulty*. Columbia University.
Klauda, S. L., & Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational Psychology*, *100*(2), 310.
Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). Guilford.
Knight-Teague, K., Vanderwood, M. L., & Knight, E. (2014). Empirical Investigation of Word Callers Who Are English Learners. *School*

Psychology Review, *43*(1), 3-18.
Koriat, A., Kreiner, H., & Greenberg, S. N. (2002). The extraction of structure during reading: Evidence from reading prosody. *Memory & cognition*, *30*, 270-280.
Kuhn, M. R., Schwanenflugel, P. J., & Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Reading research quarterly*, *45*(2), 230-251.
Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of educational psychology*, *95*(1), 3.
LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive psychology*, *6*(2), 293-323.
Lai, S. A., George Benjamin, R., Schwanenflugel, P. J., & Kuhn, M. R. (2014). The longitudinal relationship between reading fluency and reading comprehension skills in second-grade children. *Reading & Writing Quarterly, 30*(2), 116-138.
Ladd, D. R. (1984). Declination.: a review and some hypotheses. *Phonology*, *1*, 53-74.
Laver, J. (1994). *Principles of phonetics*. Cambridge university press.
Lee, B. (2014). *A Study on the Effects of the Decoding Skills, Reading Fluency and Interest in English on Reading Comprehension* [Unpublished doctoral dissertation]. Ewha Womans University.
Lee, K., & Chen, X. (2019). An emergent interaction between reading fluency and vocabulary in the prediction of reading comprehension among French immersion elementary students. *Reading and Writing*, *32*(7), 1657-1679.
Lee, Y. (2018). The development of English reading fluency in elementary school student. *Studies in Foreign Language Education*, *32*(1), 139-164.
Lee, Y. (2020). Development of English reading fluency norms for elementary school students. *Studies in Foreign Language Education*, *34*(1), 1-32.
Lekwilai, P. (2014). Reader's Theater: An Alternative Tool to Develop Reading Fluency among Thai EFL Learners. *PASAA: Journal of Language Teaching and Learning in Thailand*, *48*, 89-111.
Leslie, L., & Caldwell, J. S. (2021). *Qualitative Reading Inventory-7*. Pearson.
Levelt, W. J. (1989). *Speaking: From intention to articulation.* MIT Press.
Lo, Y.-y., Cooke, N. L., & Starling, A. L. P. (2011). Using a repeated reading program to improve generalization of oral reading fluency. *Education and Treatment of Children*, *34*(1), 115-140.
Marston, D., Fuchs, L. S., & Deno, S. L. (1986). Measuring pupil progress: A comparison of standardized achievement tests and curriculum-related measures. *Diagnostique*, *11*(2), 77-90.
May, M. C. (2014). *The effects of text complexity on the oral reading prosody of fifth-grade students* (Order No. 3630925). Available from ProQuest

Dissertations & Theses Global. (1564219263). https://www.proquest.com/dissertations-theses/effects-text-complexity-on-oral-reading-prosody/docview/1564219263/se-2
Meisinger, E. B., Bradley, B. A., Schwanenflugel, P. J., & Kuhn, M. R. (2010). Teachers' perceptions of word callers and related literacy concepts. *School Psychology Review*, *39*(1), 54-68.
Meisinger, E. B., Bradley, B. A., Schwanenflugel, P. J., Kuhn, M. R., & Morris, R. D. (2009). Myth and reality of the word caller: The relation between teacher nominations and prevalence among elementary school children. *School Psychology Quarterly*, *24*(3), 147.
Miller, J., & Schwanenflugel, P. J. (2006). Prosody of syntactically complex sentences in the oral reading of young children. *Journal of educational psychology*, *98*(4), 839.
Miller, J., & Schwanenflugel, P. J. (2008). A Longitudinal Study of the Development of Reading Prosody as a Dimension of Oral Reading Fluency in Early Elementary School Children. *Reading Research Quarterly*, *43*(4), 336-354.
Mislevy, R. J., Haertel, G., Riconscente, M., Wise Rutstein, D., Ziker, C.(2017). Evidence-centered assessment design. In Mislevy, R. J., Haertel, G., Riconscente, M., Wise Rutstein, D., Ziker, C (Eds.), *Assessing Model-Based Reasoning using Evidence- Centered Design.* SpringerBriefs in Statistics. https://doi.org/10.1007/978-3-319-52246-3_3
Morris, D., Trathen, W., Frye, E. M., Kucan, L., Ward, D., Schlagal, R., & Hendrix, M. (2013). The role of reading rate in the informal assessment of reading ability. *Literacy Research and Instruction*, *52*(1), 52-64.
Muthen, L., Muthen, B., Asparouhov, T., & Nguyen, T. (2017). Mplus Program (Version 8.0)[Computer Software]. Los Angeles, CA: Muthen & Muthen
Nagy, W., Berninger, V., Abbott, R., Vaughan, K., & Vermeulen, K. (2003). Relationship of morphology and other language skills to literacy skills in at-risk second-grade readers and at-risk fourth-grade writers. *Journal of educational psychology*, *95*(4), 730.
National Reading Panel (US), National Institute of Child Health, & Human Development (US). (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.
Newell, K. W., Codding, R. S., & Fortune, T. W. (2020). Oral reading fluency as a screening tool with English learners: A systematic review. *Psychology in the Schools*, *57*(8), 1208-1239.
Ng, P., & Boucher-Yip, E. (2010). Readers Theater: Improving Oral Proficiency in a Japanese University EFL Course.(2). *ELT World Online*, *2*, 1-18.

Paige, D. D., Rasinski, T., Magpuri-Lavell, T., & Smith, G. S. (2014). Interpreting the relationships among prosody, automaticity, accuracy, and silent reading comprehension in secondary students. *Journal of Literacy Research*, *46*(2), 123-156.
Paige, D. D., Rasinski, T. V., & Magpuri-Lavell, T. (2012). Is fluent, expressive reading important for high school readers? *Journal of Adolescent & Adult Literacy*, *56*(1), 67-76.
Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading research quarterly*, *40*(2), 184-202.
Paris, S. G., & Hamilton, E. E. (2009). The development of children's reading comprehension. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 32-53). Routledge.
Park, H., & Rhee, S.-C. (2018). Suprasegmentals affecting oral reading fluency in Korean elementary students English read speech. *Modern English Education*, *19*(3), 48-56.
Perfetti, C.A. (1985). *Reading ability*. Oxford University Press.
Perfetti, C. A. (1988). Verbal efficiency in reading ability. In M. Daneman, G. E. Mackinnon, & T. G. Waller (Eds.), *Reading research: Advances in theory and practice*, Vol. 6, (pp. 109–143). Academic Press.
Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. *Precursors of functional literacy*, *11*, 67-86.
Pinnell, G. S. & National Assessment of Educational Progress, Princeton, NJ. & Educational Testing Service, Princeton, NJ. Center for the Assessment of Educational Progress. (1995). *Listening to Children Read Aloud Data from NAEP's Integrated Reading Performance Record (IRPR) at Grade 4*. https://eric.ed.gov/?id=ED378550
Proctor, C. P., Carlo, M., August, D., & Snow, C. (2005). Native Spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology*, *97*(2), 246.
Quirk, M., & Beem, S. (2012). Examining the relations between reading fluency and reading comprehension for English language learners. *Psychology in the Schools*, *49*(6), 539-553.
Quirk, M., Schwanenflugel, P. J., & Webb, M.-y. (2009). A short-term longitudinal study of the relationship between motivation to read and reading fluency skill in second grade. *Journal of literacy research*, *41*(2), 196-227.
Rasinski, T. V. (2006). Reading fluency instruction: Moving beyond accuracy, automaticity, and prosody. *The Reading Teacher, 59*(7), 704-706.
Rasinski, T. V., Blachowicz, C. L., & Lems, K. (Eds.). (2012). *Fluency instruction: Research-based best practices.* Guilford Press.
Rasinski, T.V., Rikli, A., & Johnston, S. (2009). Reading fluency: More than automaticity? More than a concern for the primary grades? *Literacy*

Research and Instruction, *48*(4), 350-361.
Rasinski, T. V., Reutzel, D. R., Chard, D., & Linan-Thompson, S. (2011). Reading fluency. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. P. Afflerbach (Eds.), *Handbook of Reading Research* (Vol. IV, pp. 286-319). Routledge.
Ravid, D., & Mashraki, Y. E. (2007). Prosodic reading, reading comprehension and morphological skills in Hebrew-speaking fourth graders. *Journal of Research in Reading, 30*(2), 140-156.
Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement*, *21*(2), 173-184.
Rhee, S.-C., Cho, C.-H., & Moon, S.-Y. (2003). Korean & native speakers' high-low range differences in F0 and its role in pronunciation assessment. *Speech Sciences*, *10*(4), 93-103.
Riedel, B. W. (2007). The relation between DIBELS, reading comprehension, and vocabulary in urban first-grade students. *Reading research quarterly*, *42*(4), 546-567.
Roe, B., & Burns, P. C. (2010). Informal reading inventory: Preprimer to twelfth grade. Cengage Learning.
Roehrig, A. D., Petscher, Y., Nettles, S. M., Hudson, R. F., & Torgesen, J. K. (2008). Accuracy of the DIBELS oral reading fluency measure for predicting third grade reading comprehension outcomes. *Journal of School Psychology*, *46*(3), 343-366.
Ryu, J. (2020). *Processing Changes in Three Levels of Mental Representations of Korean EFL Readers during Repeated Reading and Narrow Reading of L2 Narrative Texts* [Unpublished doctoral dissertation]. Seoul National University.
Ryu, J., & Lee, B. (2021). Diagnosis of Korean EFL High School Students' Reading Fluency Using Informal Reading Inventory. *The Journal of AsiaTEFL*, *18*(2), 489-504.
Sabatini, J., Wang, Z., & O'Reilly, T. (2019). Relating reading comprehension to oral reading performance in the NAEP fourth-grade special study of oral reading. *Reading Research Quarterly*, *54*(2), 253-271.
Saiegh-Haddad, E. (2003). Bilingual oral reading fluency and reading comprehension: The case of Arabic/Hebrew (L1) English (L2) readers. *Reading & writing*, *16*(8), 717-736. https://doi.org/10.1023/A:1027310220036
Samuels, S. J. (2007). The DIBELS tests: Is speed of barking at print what we mean by reading fluency? *Reading Research Quarterly*, 42(4), 563–566.
Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of educational psychology*, *96*(2), 265.
Scherer, K. R. (2000). A cross-cultural investigation of emotion inferences from

voice and speech: implications for speech technology. In: Proc. Internat. Conf. on Spoken Language Processing (ICSLP ’00), Vol. 1, pp. 379–382.
Schreiber, P. A. (1980). On the acquisition of reading fluency. *Journal of reading behavior*, *12*(3), 177-186.
Schreiber, P. A. (1987). Prosody and structure in children's syntactic processing. In R. Horowitz & S. J. Samuels (Eds.), Comprehending oral and written language (pp. 243–270). Academic Press.
Schreiber, P. A. (1991). Understanding prosody's role in reading acquisition. *Theory into practice*, *30*(3), 158-164. https://doi.org/10.1080/00405849109543496
Schwanenflugel, P. J., & Benjamin, R. G. (2016). Lexical prosody as an aspect of oral reading fluency. *Reading and Writing*, *30*(1), 143-162. https://doi.org/10.1007/s11145-016-9667-3
Schwanenflugel, P. J., Hamilton, A. M., Kuhn, M. R., Wisenbaker, J. M., & Stahl, S. A. (2004). Becoming a Fluent Reader: Reading Skill and Prosodic Features in the Oral Reading of Young Readers. *J Educ Psychol*, *96*(1), 119-129. https://doi.org/10.1037/0022-0663.96.1.119
Schwanenflugel, P. J., & Knapp, N. F. (2015). *The psychology of reading: Theory and applications*. Guilford Publications.
Schwanenflugel, P. J., Westmoreland, M. R., & Benjamin, R. G. (2015). Reading fluency skill and the prosodic marking of linguistic focus. *Reading and Writing*, *28*, 9-30.
Seidenberg, M. S. (1997). Language acquisition and use: Learning and applying probabilistic constraints. *Science*, *275*(5306), 1599-1603.
Selkirk, E. O. (1986). *Phonology and syntax: The relationship between sound and structure*. MIT press.
Shin, J. (2021). Completing the Triangle of Reading Fluency Assessment: Accuracy, Speed, and Prosody. In *Challenges in Language Testing Around the World: Insights for language test users* (pp. 307-316). Springer.
Shin, D. J., & Iverson, P. (2011). Individual differences in vowel epenthesis among Korean learners of English. *The Journal of the Acoustical Society of America*, 128(4),1822-1825.
Shin, J., & McMaster, K. (2019). Relations between CBM (oral reading and maze) and reading comprehension on state achievement tests: A meta-analysis. *Journal of School Psychology*, *73*, 131-149.
Shin, K. C. (2018). Reading Fluency and Accuracy for English Language Acquisition in EFL Context. *Journal of the Korea Convergence Society*, *9*(3), 249-256.
Shinn, M. M., & Shinn, M. R. (2002). A*IMSweb training workbook: Administration and scoring of reading curriculum-based measurement (R-CBM) for use in general outcome measurement*. Edformation.

Shinn, M. R., Knutson, N., Good III, R. H., Tilly III, W. D., & Collins, V. L. (1992). Curriculum-based measurement of oral reading fluency: A confirmatory analysis of its relation to reading. *School Psychology Review*, *21*(3), 459-479.

Silverman, R. D., Speece, D. L., Harring, J. R., & Ritchey, K. D. (2013). Fluency has a role in the simple view of reading. *Scientific Studies of Reading*, *17*(2), 108-133.

Smith, G. S., & Paige, D. D. (2019). A study of reliability across multiple raters when using the NAEP and MDFS rubrics to measure oral reading fluency. *Reading Psychology*, *40*(1), 34-69.

Solari, E. J., Aceves, T. C., Higareda, I., Richards-Tutor, C., Filippini, A. L., Gerber, M. M., & Leafstedt, J. (2014). Longitudinal prediction of 1st and 2nd grade English oral reading fluency in English language learners: Which early reading and language skills are better predictors? *Psychology in the Schools*, *51*(2), 126-142.

Soper, D. S. (2016). http://www.danielsoper.statcalc

Speece, D. L., & Ritchey, K. D. (2005). A longitudinal study of the development of oral reading fluency in young children at risk for reading failure. *Journal of learning disabilities*, *38*(5), 387-399.

Stahl, S. A., & Kuhn, M. R. (2002). Center for the Improvement of Early Reading Achievement: Making It Sound like Language: Developing Fluency. *The Reading Teacher*, 55(6), 582-584.

Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading research quarterly*, 32-71.

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2018). *Using multivariate statistics* (Vol. 7). Pearson.

Taguchi, E., Gorsuch, G. J., Takayasu-Maass, M., & Snipp, K. (2012). Assisted repeated reading with an advanced-level Japanese EFL reader: A longitudinal diary study. *Reading in a Foreign Language*, 24(1), 30-55.

Taguchi, E., Takayasu-Maass, M., & Gorsuch, G. J. (2004). Developing reading fluency in EFL: How assisted repeated reading and extensive reading affect fluency development. *Reading in a Foreign Language*, 16(2), 70-96.

Tan, J., Dunham, J., Lee, C., & Vatikiotis-Bateson, E. (2006). Initial pitch informs sentence duration. *The Journal of the Acoustical Society of America*, *119*(5), 3393-3393.

Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading: A meta-analysis. *Remedial and special education, 25*(4), 252-261.

Therrien, W. J., & Kubina Jr, R. M. (2006). Developing reading fluency with repeated reading. *Intervention in school and clinic*, *41*(3), 156-160.

Thienkalaya, C., & Chusanachoti, R. (2020). Improving the English reading prosody of L2 learners through readers theater. *LEARN Journal: Language Education and Acquisition Research Network*, *13*(2), 306-320.

Thomson, J., & Jarmulowicz, L. (2016). *Linguistic rhythm and literacy* (Vol. 17). John Benjamins Publishing Company.

Tilstra, J., McMaster, K., Van den Broek, P., Kendeou, P., & Rapp, D. (2009). Simple but complex: Components of the simple view of reading across grade levels. *Journal of research in reading*, *32*(4), 383-401.

Valencia, S. W., Smith, A. T., Reece, A. M., Li, M., Wixson, K. K., & Newman, H. (2010). Oral reading fluency assessment: Issues of construct, criterion, and consequential validity. *Reading Research Quarterly*, *45*(3), 270-291.

Veenendaal, N. J., Groen, M. A., & Verhoeven, L. (2015). What oral text reading fluency can reveal about reading comprehension. *Journal of Research in Reading*, *38*(3), 213-225.

Veenendaal, N. J., Groen, M. A., & Verhoeven, L. (2016). Bidirectional relations between text reading prosody and reading comprehension in the upper primary school grades: A longitudinal perspective. *Scientific Studies of Reading*, *20*(3), 189-202.

Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific studies of reading*, *11*(1), 3-32.

Wade-Woolley, L., & Heggie, L. (2016). The contributions of prosodic and phonological awareness to reading. *Reading and Writing*, *29*, 1-12.

Wade-Woolley, L., & Wood, C. (2006). Prosodic sensitivity and reading development.

Wade-Woolley, L., Wood, C., Chan, J., & Weidman, S. (2022). Prosodic competence as the missing component of reading processes across languages: Theory, evidence and future research. *Scientific Studies of Reading*, *26*(2), 165-181.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental psychology*, *30*(1), 73.

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

Wiederhold, J. L., & Bryant, B. R. (2012). *Gray Oral Reading Tests-Version 5 (GORT-5)*. PRO-ED.

Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W., Wolf, M., Kuhn, M., . . . Schwanenflugel, P. (2010). The relationship between different measures of oral reading fluency and reading comprehension in second-grade students who evidence different oral reading fluency difficulties. *Lang Speech Hear Serv Sch*. 41(3), 340-348. https://doi.org/ 10.1044/0161-

1461.
Woodcock, R. W. (2011). *Woodcock reading mastery tests: WRMT-III*. Pearson.
Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of educational psychology*, *91*(3), 415.
Wolf, M., & Katzir-Cohen, T. (2001). Reading Fluency and Its Intervention. *Scientific studies of reading*, *5*(3), 211-239. https://doi.org/10.1207/S1532799XSSR0503_2
Wolters, A. P., Kim, Y.-S. G., & Szura, J. W. (2020). Is Reading Prosody Related to Reading Comprehension? A Meta-analysis. *Scientific Studies of Reading*, 1-20.
Woo, K., & Jeong, E. (2018). A Case Study of In- Depth Diagnostic Assessment for English Reading Intervention. *Journal of Language Sciences*, *25*(4), 1-20.
Woo, K., & Jung, E. (2012). A Study on the Correlation between Elementary Students Reading Fluency and ReadingComprehension. *The New Korean Journal of English Language and Literature*, *54*(4), 183-201.
Woo, K., & Jung, E. (2017). An Analysis of Correlation among Reading Speed, Accuracy and Comprehension. *Journal of Language Sciences*, *24*(1), 85-103.
Yeganeh, M. T. (2013). Repeated reading effect on reading fluency and reading comprehension in monolingual and bilingual EFL learners. *Procedia-Social and Behavioral Sciences*, *70*, 1778-1786.
Yeo, S. (2010). Predicting performance on state achievement tests using curriculum-based measurement in reading: A multilevel meta-analysis. *Remedial and Special Education*, *31*(6), 412-422.
Young, C., & Rasinski, T. (2009). Implementing Readers Theatre as an Approach to Classroom Fluency Instruction. *Reading Teacher*, *63*(1), 4-13. https://doi.org/10.1598/rt.63.1.1
Yuasa, I. P. (2010). Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech*, *85*(3), 315-337.
Zutell, J., & Rasinski, T. V. (1991). Training teachers to attend to their students' oral reading fluency. *Theory Into Practice*, *30*(3), 211-217.
Álvarez-Cañizo, M., Suárez-Coalla, P., & Cuetos, F. (2015). The role of reading fluency in children's text comprehension. *Frontiers in psychology*, *6*, 1810.
Álvarez-Cañizo, M., Suárez-Coalla, P., & Cuetos, F. (2018). Reading prosody development in Spanish children. *Reading and Writing*, *31*(1), 35-52.
Álvarez-Cañizo, M., Martínez-García, C., Cuetos, F., & Suárez-Coalla, P. (2020). Development of reading prosody in school-age Spanish children: a longitudinal study. *Journal of research in reading*, *43*(1), 1-18. https://doi.org/10.1111/1467-9817.12286

Appendices

Appendix A: Consent Form

IRB No. 2204/003-015　　　　　　　　　　　　유효기간: 2023년 04월 21일

동　의　서 (연구 참여자 보관용)

연구 과제명 : 한국인 영어학습자의 영어 읽기 낭독에 나타난 운율적 특성의 역할에 대한 탐구

연구 책임자명 : 유지선(서울대학교 사범대학 외국어교육과, 박사과정, 010)

1. 나는 이 설명서를 읽었으며 담당 연구원과 이에 대하여 의논하였습니다.
2. 나는 위험과 이득에 관하여 들었으며 나의 질문에 만족할 만한 답변을 얻었습니다.
3. 나는 이 연구에 참여하는 것에 대하여 자발적으로 동의합니다.
4. 나는 이 연구에서 얻어진 나에 대한 정보를 현행 법률과 생명윤리위원회 규정이 허용하는 범위 내에서 연구자가 수집하고 처리하는 데 동의합니다.
5. 나는 담당 연구자나 위임받은 대리인이 연구를 진행하거나 결과 관리를 하는 경우와 법률이 규정한 국가 기관 및 서울대학교 생명윤리위원회가 실태 조사를 하는 경우에는 비밀로 유지되는 나의 개인 신상 정보를 확인하는 것에 동의합니다.
6. 나는 언제라도 이 연구의 참여를 철회할 수 있고 이러한 결정이 나에게 어떠한 해도 되지 않으리라는 것을 압니다.
7. 나는 수집되는 자료가 본 연구 이외에 연구 책임자 및 다른 연구자의 연구목적으로 사용되는 것에 동의합니다.

   동의함 □　　동의하지 않음 □
8. 나의 서명은 이 동의서를 받았다는 것을 뜻하며 나와 동의받는 연구원의 서명이 포함된 동의서를 보관하겠습니다.
9. 나는 연구를 수행하는 중에 (녹음)이 진행되는 것에 동의합니다.

   동의함 □　　동의하지 않음 □

연구 참여자 성명	서 명	날짜 (년/월/일)
법정 대리인 성명　참여자와의 관계	서 명	날짜 (년/월/일)
동의받는 연구원 성명	서 명	날짜 (년/월/일)

- 1 -

Ver 1.1(2022.04.22.)

SNUIRB

Appendix B

Reading Material A Modified from May (2014)

John and Mary sat on the edge of the pool and watched. **The koala swam and headed for the opposite side of the pool across from Mary and John.** The koala tried to pull himself out, but the tiles were too slippery.

“He can’t get out,” John said. “We’ve got to help him.”

Mary kicked off her sandals and stepped quietly into the pool.

By the time Mary reached the frightened animal, the koala was paddling to the metal ladder. Mary put her arms under the koala and gave him a lift. John helped to lift the animal onto solid ground. **The koala gave one shake, spraying John with water, before toddling across the lawn and up the nearest tree.**

“We’ve never had a koala in the pool before,” Mary said, “but koalas love to swim. My teacher said that backyard pools are a big cause of accidents to koalas in Australia.”

Reading Material B Modified from May (2014)

A long time ago, during the time when the world was new, animals across the earth were busy developing their natures and appearances. The birds of the world were especially active at this time.

Each bird was learning the songs that would be its own and that would identify that particular bird to the other animals. They were also trying on feathers that would mark each type of bird as distinct and beautiful.

One bird, Pi-coo, was having an especially difficult time. **She could not make up her mind about which feathers she should wear.** The more she tried on, the more confused she became. Soon, almost all of the feathers were spoken for, and she was left with almost nothing to cover her naked body. Because she had no feathers, she was very ashamed and refused to come out of her nest.

The other birds felt sorry for her. They gathered together and talked about a way they could help Pi-coo.

Appendix C

Parameter Estimate Coefficients of Two-Dissociable-Constructs Model of ORF

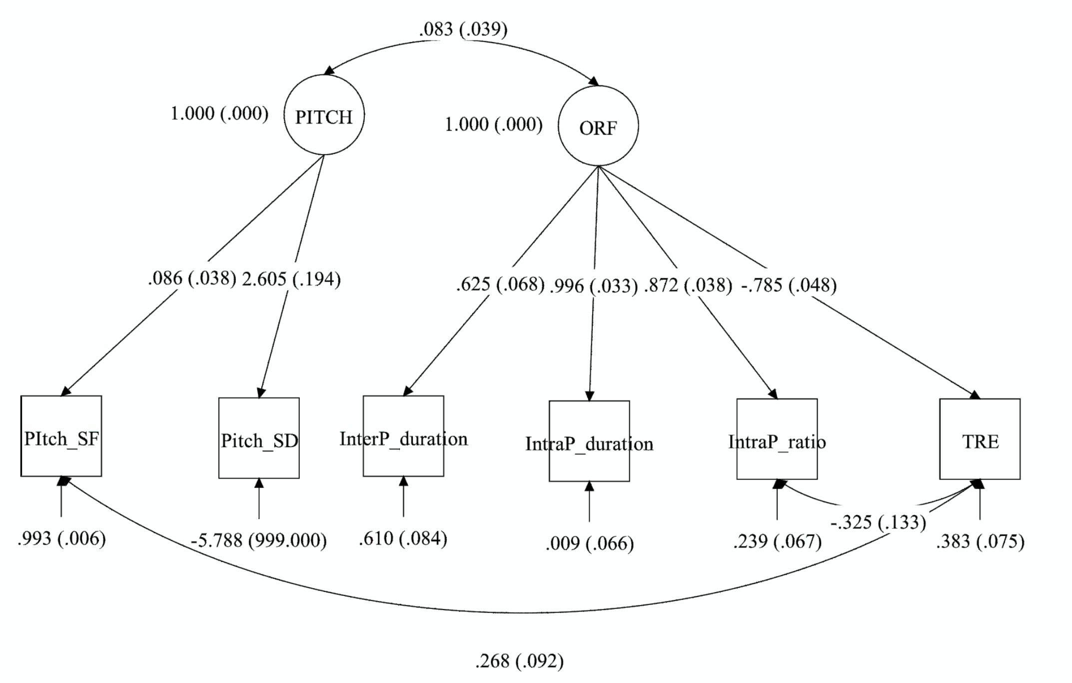
	Variables	B	S.E.	Est./S.E.	*p*
Factor Loading					
ORF	TRE	-.785	.048	-16.479	.000***
	IntraP_ratio	.872	.038	22.685	.000***
	IntraP_duration	.996	.033	30.088	.000***
	InterP_duration	.625	.068	9.251	.000***
PITCH	Pitch_SD	2.605	0.194	13.432	.000***
	Pitch_SF	.086	.038	2.273	.023*
Correlation					
	PITCH with ORF	.083	.039	2.101	.036*
	TRE with IntraP_ratio	-.325	.133	-2.437	.015*
	TRE with Pitch_SF	.268	.092	2.931	.003**
Residual Variance					
	TRE	.383	.075	5.120	.000***
	Pitch_SD	-5.788	999	999	999
	Pitch_SF	.993	.006	152.835	.000***
	IntraP_ratio	.239	.067	3.563	.000***
	IntraP_duration	.009	.066	.133	.894
	InterP_duration	.610	.084	7.220	.000***

Model Fit Indices

χ^2	*df*	RMSEA	CFI	TLI	SRMR
20.756**	7	.148	.953	.900	.086

Note. $^*p < .05$, $^{**}p < .01$; $^{***}p < .001$; TRE=text reading efficiency; Pitch_SD=overall intonation contour; Pitch_SF=pitch changes at the sentence-final position; IntraP_ratio=intrasentential pause ratio; UGP_ratio=ungrammatical pause ratio; IntraP_duration= intrasentential pause duration; InterP_duration=intersentential pause duration; RMSEA = root mean square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis index; SRMR = standardized root mean square residual

Graphic Illustration of CFA Results of Two-Dissociable-Constructs Model of ORF



국 문 초 록

한국인 고등학교 영어 학습자의 영어 읽기 유창성에서
운율적 자질의 역할에 관한 탐구

유 지 선

외국어교육과 영어전공

서울대학교 대학원

읽기 속도, 정확성, 표현력으로 정의되는 읽기 유창성은 읽기 이해에 있어서 중요한 요소로 간주된다. 정보처리이론의 이론적 배경과 더불어 상당수의 실험연구 결과로 읽기 유창성은 연구자들과 교육자들로부터 많은 관심을 받아왔다. 그 중요성이 인정되어, 모국어와 제 2 언어 교육에서 다양한 수업 방법이 개발되고 수업에 이용되었으며, 읽기 유창성을 측정하기 위한 다양한 평가 도구가 개발되었다. 그러나, 이러한 평가도구들은 읽기 속도와 중요성만을 측정하고, 읽기 유창성의 중요한 요인인 운율적 특성을 제외한 것에 대해 끊임없는 비판의 대상이 되어왔다.

읽기의 운율적 특성은 휴지, 억양, 리듬 등과 같이 읽기 낭독의 음악적 특성을 일컫는다. 선행연구들은 읽기에 능숙하고 유창할수록 읽기 낭독이 좀더 부드럽고 구어와 비슷하며, 단어단어를 더듬거리며 읽지 않게 된다는 것을 보여주었다. 또한, 읽기의 운율적 특성은 처리되는 정보를 작동기억에 저장하는 역할을 하거나, 의미 처리를 위한 인지적 비계 역할을

하며 읽기 이해에 있어 중요한 역할을 하는 것으로 알려져 있다(Frazier et al., 2006). 비록 지금까지 읽기의 운율적 특성에 대해서 많은 연구가 행해지지는 않았지만, 모국어의 읽기에 관한 연구에서는 이러한 특성이 읽기 유창성과 이해에 있어 중요한 요인이라는 것이 점차 인정받고 있다.

제 2 언어 연구에서도 읽기 유창성은 읽기에 있어서 중요한 요소로 자리매김하고 있으며, 상당히 많은 연구들로 인해 읽기 유창성과 읽기 이해의 관계도 밝혀지고 있다. 하지만, 여전히 운율적 특성은 읽기 연구에 있어서 간과되어 왔다.

따라서, 본 연구는 운율적 요소들이 제 2 언어 읽기에서 하는 다양한 역할을 탐구하였다. 첫째로, 학습자의 유창성 수준에 따라 운율적 특성이 어떻게 달라지는지를 탐구하였다. 둘째로, 모국어 읽기 연구에서 규정된 읽기 유창성의 정의가 제 2 언어 상황에도 적용되는지, 즉, 읽기 유창성이 읽기 속도와 정확성 그리고 운율적 특성으로 구성된 단일한 구인인지를 검증하고자 하였다. 마지막으로, 읽기의 운율적 특성과 이해 능력 간의 관계를 고찰하였다.

영어를 외국어로 사용하는 총 90 명의 한국 고등학교 학생들이 이 연구에 참여하였다. 참여자들은 단어 읽기 능력, 텍스트 읽기 능력, 그리고 읽기이해 평가를 받았으며, 운율적 요소들은 학생들이 읽기 낭독을 녹음한 후, 이를 스펙트로그래프 분석을 통해 추출하였다. Praat 과 같은 음성분석도구를 사용한 스펙트로그램의 분석은, 소리 파형의 시각 그래프를 만들어 다양한 운율적 특성의 식별과 측정이 가능하며, 정확하고 객관적인

측정이 가능하다는 장점이 있다.

이러한 분석의 결과로 문장 내 휴지의 빈도, 문장 내 휴지의 총 길이, 비문법적 휴지의 빈도, 문장 사이의 휴지 길이, 전반적인 억양 곡선, 문장 말미에서 억양의 변화, 이렇게 총 6 개의 운율적 특성이 추출되었으며 추후 분석의 대상이 되었다. 분석방법으로 일원분산분석, 확인적 요인분석, 위계적 회귀분석이 사용되었다.

본 연구의 결과는 다음과 같다. 첫째로, 휴지와 관련된 변인, 즉, 문장내 휴지 빈도, 문장 내 휴지 길이, 비문법적 휴지의 빈도는 유창성의 정도에 따라 두드러진 차이를 보였다. 그러나, 전반적인 억양 곡선은 유창성의 수준에 따라 유의미한 차이를 보이지 않았으며, 이는 억양곡선은 한국인 제 2 언어 학습자들의 유창성과는 관계없이 큰 차이가 없다는 것을 보여주었다. 또한 문장 말미에서의 억양 변화와 문장 간 휴지의 길이는 상위 학습자만 유의미하게 변별되어, 이러한 차이는 단어와 문장의 읽기 유창성이 충분히 발달해야 분명해진다는 사실을 나타냈다.

읽기 유창성의 정의에 대한 가설을 검증한 결과, 영어를 모국어로 사용하는 맥락과 마찬가지로 제 2 언어 읽기 상황에서도 이 구인이 정확성, 속도, 운율적 특성을 하위 요인으로 가지는 다면적이지만 단일한 속성을 가지고 있다는 사실이 밝혀졌다.

마지막으로 본 연구는 운율적 요소들이 단어 읽기 능력과 읽기 이해에서 매개 역할을 하며, 읽기이해에 중요한 역할을 한다는 것을 발견하였다. 하지만, 위계적 회귀분석에서 텍스트 읽기 효율성이 먼저

분석에 투입될 시 운율적 요소는 읽기 이해에 있어서 의미 있는 예측력은 없는 것으로 나타났으나, 이는 텍스트 읽기 효율성과 운율적 요소의 중첩되는 효과 때문인 것으로 확인되었다. 따라서, 이들 변수의 투입 순서를 바꾸어 다시 위계적 회귀 분석을 실시한 결과 텍스트 읽기 효율성과 운율적 요소 모두 읽기 이해를 설명하는 유의미한 변수임을 발견하였다.

본 논문은 제 2 언어 읽기 수업에 있어서 중요한 교육적 함의를 지닌다. 연구의 결과는 운율적 요인의 발달이 제 2 언어 읽기 능을 향상시키는 데 있어서 중요하다는 것을 시사한다. 교사들은 함께 소리내어 읽기나, 반복 읽기와 같은 방법을 수업에 적용하여 제 2 언어 학습자들이 읽기 유창성을 도모해야 한다. 또한 운율적 요인들이 단어 읽기 능력과 읽기 이해에서 매개역할을 한다는 사실을 고려하면, 읽기 유창성 척도를 사용하여 학생들의 발달을 측정하고 이를 평가에 반영할 수 있는 방법을 고안해야 한다. 선행연구처럼, 본 연구의 결과는 타당하고 신뢰로운 읽기 유창성 척도를 개발하는데 이용될 수 있을 것이다. 전반적으로 본 연구의 결과는 제 2 언어 학습자들의 읽기 이해 능력을 촉진하기 위해 읽기 유창성, 특히 운율적 요소를 수업과 평가에 통합하는 더 포괄적인 접근을 재고해야 할 것을 시사한다.

주요어: 영어 읽기 유창성, 운율적 요소, 읽기 이해, 읽기 평가, 제 2 언어 읽기

학 번: 2018-35246