Diagnosis of Korean EFL High School Students’ Reading Fluency: Using Informal Reading Inventory and Their L2 Reading Experiences

비공식 영어읽기 평가도구와 읽기 경험이 토대로 한 한국 고등학교 학생들의 영어 읽기 유창도 진단

2018 년 2월

서울대학교 대학원
외국어교육과 영어전공
류 지 선
Diagnosis of Korean EFL High School Students’ Reading Fluency: Using Informal Reading Inventory and Their L2 Reading Experiences

by
Jiseon Ryu

A Thesis Submitted to
the Department of Foreign Language Education
in Partial Fulfillment of the Requirements
for the Degree of Master of Arts in Education

At the
Graduate School of Seoul National University

February 2018
Diagnosis of Korean EFL High School Students’ Reading Fluency: Using Informal Reading Inventory and Their L2 Reading Experiences

이 논문을 교육학 석사 학위논문으로 제출함
2018년 2월

서울대학교 대학원
외국어교육과 영어전공
류지선

류지선의 석사학위논문을 인준함
2018년 2월

위원장 ________________________
부위원장 ________________________
위원장 ________________________
Diagnosis of Korean EFL High School Students’ Reading Fluency: Using Informal Reading Inventory and Their L2 Reading Experiences

APPROVED BY THESIS COMMITTEE:

__________________________
JIN-WAN KIM, COMMITTEE CHAIR

__________________________
YOUNGSOON SO

__________________________
BYUNGMIN LEE
ABSTRACT

The present thesis attempts to examine and diagnose Korean EFL high school students’ English reading fluency using the informal reading inventory (IRI), combined with their L2 reading experiences.

Fluent reading skills can make L2 learners read for an extended period of time. It can be defined as overall reading competence accompanying the automatic execution of lower-level skills and the higher-level reading comprehension process. That is, reading comprehension is critically dependent on the reader’s fluency in recognizing words, and it is represented in the readers’ accuracy and speed in identifying words in a running passage as well as in the ability to read with appropriate expression. Based on the definition of reading fluency, this thesis explores the questions as to how fluent Korean-speaking L2 readers are.

This study involved 68 eleventh grade high school students in Gyeonggi province. In the oral reading fluency (ORF) measure, the participants were asked to read aloud the graded texts imported directly from IRI (Burns & Roe, 2011), and answer seven to nine questions pertaining to the passage. After completing oral reading tasks, a post semi-structured interview was conducted in order to elicit comprehensive information as to which L2 reading experiential factors were involved. The detailed analysis of the students’ performances revealed that the students’ ORF showed wide dispersion, ranging from level below-preprimer (PP) to level 7 at the ORF/Independent Level (ORF-IL) and from level pp to
level 9 at the ORF/Frustration Level (ORF-FL). Notably, about 40 percent of the students seemed to read the text of the level 2 and 3 independently, while approximately 50 percent of the students showed frustration reading passages at the level 3 through level 5. The subsequent statistical analysis revealed that the English proficiency level drawn from the mock-CSAT and ORF level are significantly correlated. Second, there was less variability in word recognition accuracy across ORF-IL and ORF-FL. The participants can read texts at both independent level and frustration level with high accuracy. This word recognition pattern at the frustration level was consistent with the previous studies and the result confirmed that the ‘word calling’ phenomenon is prevalent among L2 readers. Third, there were significant differences in the number of the decoding errors across the reading fluency levels. The lower the students’ reading fluency was, the more decoding errors were detected. Fourth, there was an increasing pattern in the reading rate depending on the ORF level even though significant variances were detected in the same fluency group. Lastly, regarding the L2 reading experiences that contributed to English reading fluency, the results revealed that there were significant differences in the total amount of L2 exposure and reading experiences across the reading fluency levels. This indicates that continuous and extensive L2 reading practice with the appropriate text allows L2 readers to gain automaticity of basic component skills, which enhances reading comprehension.
Key Words: L2 reading fluency, reading rate, reading accuracy, decoding errors, L2 reading experience, informal reading inventory.

Student Number: 2016-21792
# TABLE OF CONTENTS

ABSTRACT ................................................................................................................i

TABLE OF CONTENTS ............................................................................................. iv

LIST OF TABLES ....................................................................................................... vi

LIST OF FIGURES .................................................................................................... viii

CHAPTER 1 INTRODUCTION .................................................................................... 1
  1.1 The Purpose of the Study .................................................................................... 1
  1.2 Research Questions ............................................................................................ 7
  1.3 Organization of the Thesis ................................................................................ 8

CHAPTER 2 LITERATURE REVIEW ........................................................................ 9
  2.1 Definition and Theoretical Bases for Reading Fluency .................................... 10
  2.2 The Construct of Reading Fluency ..................................................................... 12
    2.2.1 Reading rate ................................................................................................. 13
    2.2.2 Word recognition accuracy ......................................................................... 16
    2.2.3 Reading comprehension ............................................................................... 21
    2.2.4 Research in L2 reading fluency ................................................................. 23
  2.3 Reading Assessment .......................................................................................... 26
    2.3.1 Standardized testing ..................................................................................... 26
    2.3.2 Informal reading inventory .......................................................................... 28
  2.4 L2 Experiential Factors for Reading Fluency ................................................... 31

CHAPTER 3 METHODOLOGY .................................................................................... 35
  3.1 Participants ....................................................................................................... 35
LIST OF TABLES

Table 2.1 The synthesis of the L1 readers’ decoding errors.................................20
Table 2.2 Performance criteria for reading accuracy and comprehension (Betts,
1946) .........................................................................................................................29
Table 3.1 The participants’ reading proficiency based on the mock-CSAT...........37
Table 3.2 The readability indices of Form A .............................................................41
Table 3.3 Sample comprehension questions (Level 3) ...........................................43
Table 3.4 The readability indices of the MCSAT reading passages .......................44
Table 3.5 Descriptive statistics of the MCSAT reading passages ............................45
Table 3.6 Performance criteria for WRA and ORC ...............................................47
Table 3.7 Reading fluency group by the ORF level ..............................................50
Table 4.1 The distribution of the 11th grade high school students across the ORF
levels (N=68) ........................................................................................................53
Table 4.2 The correlation between ORF-IL and MCSAT level ..............................57
Table 4.3 The WRA scores of the participants across the ORF-FL .......................61
Table 4.4 The frequency and types of the decoding errors .................................64
Table 4.5 Word list for the substitution errors across the fluency groups ............68
Table 4.6 The comparison of the original words and the students’ responses in the
context across the fluency groups .......................................................................69
Table 4.7 The ORR of the fluency groups across the ORF levels

........................................................................................................................................75
LIST OF FIGURES

Figure 4.1 The distribution of the participants across the ORF levels ................................................................. 54

Figure 4.2 The percentage and the types of the decoding errors across the reading fluency groups ................................................................. 65
CHAPTER 1
INTRODUCTION

1.1 The Purpose of the Study

In the modern society where print is all around, reading is the activity which is required throughout the day. Many people are engaged in reading, consciously or unconsciously, according to their own purposes (i.e., to obtain information or for entertainment). As it is one of the typical goal-oriented human behaviors, the act of reading is taken for granted in many societies (Grabe, 2009). Indeed, with little effort or little planning, good and skilled readers can read many different types of text.

However, to become such a skilled reader especially in a second language (L2) is not easy since reading is a complex process in which synchronization and coordination of multifaceted components are involved (Alderson, 2000; Grabe, 2009; Pikulski, 2006; Rasinski, 2012). In order to define and describe fluent reading skill, a wealth of research has been conducted (Adams, 1990; Biemiller, 1997; Fuchs & Deno, 1991; Fuchs, Fuchs, Hops & Jenkins, 2001; Huey, 1908; LaBerge & Samuels, 1974; Pikulski, 2006). Notably, the definition and the nature of fluent reading came from a significant book about reading written by Huey (1908) approximately a century ago. Huey (1908) described students’ progress from the beginner whose close attention to each word is required for recognition, to the fluent reading stage where words could
be recognized automatically with speed and accuracy. This notion was theorized into the automaticity theory by LaBerge and Samuel (1974). It was the theory of reading explaining the development of automaticity in word recognition. This theory postulates that if the learner is not automatic in word recognition, the reading process has to be done in two stages. First, the student has to pay attention to word recognition. As this process is not automatic, the available attentional resources are used for the decoding process and then the student switches the attention to reading comprehension. As this two-stage process is inefficient and places a high demand on the working memory, the subsequent stage of the reading process (i.e., reading comprehension), which is the ultimate goal of reading is impeded. With regard to this theory, Pikulski (2006) describes reading fluency as the “bridge” between word recognition and reading comprehension.

Therefore, the most salient feature of fluent reading is the automatic word recognition, which can be represented in the rate of reading words. Reading scholars (e.g., Adams, 1990; Fuchs et al., 2001) proposed that the speed with which the text is reproduced with spoken language may serve as an index of overall reading competence. Therefore, the term, oral reading fluency (ORF), was proposed to represent a complicated performance that entails a reader’s cognitive process of the text reading including lower-level and higher-level processing. For example, when an individual reader translates the text into the spoken version, s/he changes the letters into sound representation, utilizes this phonological information to access oral vocabulary in lexical memory, and
integrates lexical and syntactic information at the intra- and inter-sentence level. In other words, when a reader reads aloud, s/he quickly orchestrates these skills in a seemingly effortless and unconscious way. Because ORF reflects these complex cognitive processes, it can be used to characterize reading competence.

Recognizing the importance of ORF as overall reading competence, researchers and educators in the L1 setting viewed it as an important goal of reading instruction. Thus, considerable attempts have been made to examine how L1 readers develop ORF (Deno, Marson, Shinn, & Tindal, 1983; Deno, Mirkin, & Chiang, 1982; Fuchs, Fuchs, & Maxwell, 1988; Hintze, Shapiro, Conte, & Basile, 1997; Kuhn & Stahl, 2003; Rasinski, Samuels, Hiebert, Petscher, & Feller, 2011; Shinn, Good, Knutson, Tilly, & Collins, 1992), and there has been a great deal of effort to incorporate ORF instruction in reading classes. Moreover, the attempts to assess ORF and develop appropriate assessment tools have been made in the L1 context.

Nevertheless, in L2 reading research reading fluency has been underdeveloped as a prominent research topic until very recently. Consequently, it has not become an important part of the reading instruction and curriculum. A possible explanation for this is that other reading-related problems have taken priority over fluency or reading fluency has been assumed to develop naturally as the language proficiency enhanced (Rasinski, Blachowicz, & Lems, 2012). A tendency among foreign language learners to regard a text as an object for language studies and not as an object for factual information, literary experience or simply pleasure, can be another explanation (Simensen, 1987). Particularly, in
the Korean context where students’ English reading experience is restrained to in-class reading, priorities in learning and teaching have been given to grammar-translation and test-taking strategies based on standardized tests such as the Korean College Scholastic Ability Test (CSAT) (Song, 1998; Yang, Oh, & Kim, 2009). In fact, as the CSAT is a high-stake assessment, English classroom has been transformed into test preparation.

To make the matters worse, the test results also have not informed what a test-taker really knows or can do. In the paradigm of norm-referenced testing on which the CSAT used to be based, the meaning of students’ reading score on the test had been derived from the relative standing of a student’s score in relation to others’ (Fulcher, 2010). In other words, this type of test identifies whether the test taker perform better or worse than other test takers, not whether the test taker has the proper degree of the reading ability which the test intends to measure.

In addition, even though classroom assessment is fundamentally different from large-scale assessment primarily in that it focuses on enhancing students’ learning (Fulcher, 2010; Fulcher & Davison, 2007), the paradigm of large-scale assessment has operated in classroom assessment as well (Bailey, 1998; Kwak, 2008; Lee, 2002; Moss, 2003; Park, 2010; Seo & Choe, 2017). The current classroom assessment administered in Korean secondary schools has been designed to differentiate students from each other, and multiple-choice questions have been commonly used for the sake of efficiency (Lee, 2002; Seo & Choe, 2017). Thus, the meaning of the score of the classroom assessment could
not provide adequate information on how fluently a student can read and what level of reading competency s/he has. More importantly, it could not provide any meaningful information to teachers who make inferences to students’ reading ability from the test score. As Fulcher and Davison (2010) assert, the test validity can be established “when there is a strong link between inferences and decisions, and when the test has a positive impact on people and institutions” (p. 20). Thus, the validity and the usefulness of the current classroom-based assessment used in a Korean context appear problematic to be justified.

While the classroom-based assessment makes use of test types found in the standardized assessment, many L1 researchers have proposed to use informal assessments to enhance the validity of the classroom assessment (Carlisle & Rice, 2004; Glaser & Silver, 1994; Grabe, 2009; Grabe & Jiang, 2014; Leslie & Caldwell, 2009; Snyder, Caccamise, & Wise, 2005). Grabe (2009), for example, suggests several types of informal reading assessment options which are available to teachers. Among these informal assessment choices, the usefulness of the oral reading fluency measure as a strong predictor of general comprehension has been discussed (Grabe & Jiang, 2014; Fuchs et al., 2001; Shinn et al., 1992; Valencia et al., 2010). For instance, Grabe and Jiang (2014) state that by simply having students read aloud a text, teachers can obtain information on the constructs of oral reading fluency such as rate, accuracy, prosody, and comprehension. Furthermore, they emphasize that if these aspects of reading performance are incorporated in reading assessment, teachers can gain “a fine-grained understanding of students’ reading ability” (p. 192). As a reading
assessment tool based on the practice of reading aloud, the informal reading inventory (IRI) is one of the widely-used oral reading fluency assessments and has a long history in informal comprehension assessment in the L1 setting. In addition, L1 reading researchers and educators have enumerated benefits of using informal reading inventories in the classroom (Paris & Carpenter, 2003; Rubin, 2002). For example, Rubin (2002) states that IRI is one of the most valuable diagnostic aids in the sense that it can convey the considerable amount of information on students’ reading abilities. Paris and Carpenter (2003) also mention that IRIs are “authentic, daily, quick, immediate, flexible, teacher-controlled, and student centered—all positive characteristics of classroom assessments” (p. 578).

To recapitulate, accumulated research evidence has suggested that for L1 readers, ORF can be a proxy for overall reading competence and its assessment using IRI can provide valuable information on students’ weakness and strength in reading. Nevertheless, research on fluency assessment has scarcely been carried out in L2 reading contexts (Grabe & Jiang, 2014). Particularly, Korean L2 readers’ fluency has been neglected as a research topic (Chae, 2016) and an attempt to examine their reading fluency has rarely been made.

Hence, the present study aims to examine and diagnose English reading abilities of Korean high school students by employing IRI as the ORF measure. Drawing upon theoretical and empirical research, the study also explores the extent to which students’ L2 reading experience varies depending on the students’ reading ability. This is important as the literature suggests that the
influence of experiential factors on L2 reading ability differs among the different ability groups and it provides further insights into the nature of L2 reading problems that struggling adolescent L2 readers in Korea may have.

1.2 Research Questions

The current study investigates Korean high school EFL learners’ English reading fluency from two perspectives. Firstly, this study diagnoses Korean EFL high school students’ oral reading fluency by employing IRI. Specifically, scores yielded from the IRI administration, namely, students’ reading fluency level, reading rate, reading accuracy, and decoding errors are examined. Secondly, this study looks into which experiential differences are involved in the development of L2 reading among the readers of various reading fluency levels. Thus, the following two research questions are raised for this study.

1) How fluently do Korean EFL high school students read based on the measurement of the informal reading inventory?

2) What kinds of L2 reading experiences do students have who are at different reading fluency levels?
1.3. Organization of the Thesis

The present study is organized into five chapters. Chapter 1 introduces the purpose of the present study with research questions. Chapter 2 presents an overview of extant theoretical and empirical studies on oral reading fluency with L2 experiential factors. Chapter 3 describes research method, including participants, instruments, data collection and analysis. Chapter 4 reports the results of the study, and discusses central issues, exploring the research questions. Chapter 5 summarizes major findings of the study and concludes the study with pedagogical implications, limitations and suggestions for future studies.
CHAPTER 2
LITERATURE REVIEW

This chapter reviews the literature associated with the main areas of interest of this study. These areas are first, the definition and theoretical bases for ORF; second, the constructs of ORF; third, reading assessment; and finally, research studies concerned with L2 reading experience.

The first section identifies the literature that explains the definition and theoretical bases for ORF. The studies of fluency have been based on the cognitive approach to information processing. ORF is defined based on LaBerge and Samuel’s (1974) automaticity theory, which is the dominant theory in conceptualizing reading fluency. The second section delineates constructs of ORF on the basis of its definition. Among the constructs, two important performance attributes (i.e., reading rate and word recognition) and reading comprehension are mainly discussed. In the third section, reading assessment issues are dealt with. Two types of the reading assessments, such as standardized testing and informal reading inventory are described. The last section reviews the literature on L2 reading experience factors involved in the reading fluency development.
2.1 Definition and Theoretical Bases for Oral Reading Fluency

Reading is a complex process that requires simultaneous synchronization of different lower-level and higher-level skills to build reading comprehension (Fuchs et al., 2001; Grabe, 2009, 2010). To achieve simultaneous coordination of component skills, those skills needed to be executed instantly and automatized. With the instantaneous execution of the skills, reading fluency is achieved and thus, reading is “speeded, seemingly effortless, autonomous, and achieved without much consciousness or awareness” (Fuchs et al., 2001, p. 239).

Although the definition of reading fluency varies across researchers, most agree that it includes reading speed, reading accuracy, and in reading a connected text, prosodic phrasing and contours of the text (Grabe, 2009; Rasinski & Samuels, 2011; Walker, Mokhtari, & Sargent, 2006). Furthermore, whether oral or silent, the significance of reading fluency is that it marks successful orchestration of lower-level skills (e.g., decoding, word recognition, syntactic processing and semantic-proposition encoding) necessary for comprehension (Fuchs et al., 2001; Grabe, 2009, 2010; Grabe & Stoller, 2013; Koda, 2005; Potter & Wamre, 1990; Rasinski, Reutzel, Chard, & Linan-Thompson, 2011). According to LaBerge and Samuel’s (1974) automaticity model of reading, successful reading comprehension is made possible if enough components are executed automatically within tolerable limits of attention. Conversely, if each component skill requires attention, its execution would
exceed the attentional capacity, leading to poor reading comprehension. Therefore, automaticity is the key construct of reading fluency (Fuchs et al., 2001; Grabe, 2009). Specifically, the theory postulates that lower-level and higher-level processing occur subsequently as skilled reading involves the reallocation of attentional capacity from lower-level word identification processing to resource-demanding comprehension functions. However, with regard to these processes, Stanovich (1980) posed an interactive and compensatory mechanism, in which the higher-level process does not wait for the completion of the lower-level processes, but compensates for the inefficiency of lower-level processing (Stanovich, 1980, 2000). Although the model of LaBerge and Samuels (1974) is different from the Interactive and Compensatory Model of reading (Stanovich, 2000) in explaining how lower- and higher-level processing occur, they share the assumption that efficient lower-level word recognition skill frees up the capacity for higher-level, integrative comprehension processing of text. Within these theoretical frames, the reading fluency with which a reader transforms texts into spoken forms means that s/he has attention-saving word recognition skill. Hence, fluent oral reading of the text serves as a performance indicator of overall reading competence in that the reading development presumes integrative comprehension processing as a result of the increasing automaticity of lower-level word recognition skills.
2.2 The Construct of Reading Fluency

Fluency has been viewed as an oral reading phenomenon (Chard, Pikulski, & McDonagh, 2006). However, because most readers spend a small amount of time engaging in oral reading compared to silent reading, a number of scholars have tried to extend a definition of fluency to encompass more than oral reading (Pinnell, 1995; Samuel, 2002; Stecker, Roger, & Marinez, 1998). While rapid and accurate oral reading is regarded as the major characteristic of fluent reading, most have agreed that reading fluency is part of a developmental process of building oral language and decoding skills that form a bridge to reading comprehension (Armbruster, Lehr, & Osborn, 2001; Kuhn & Stahl, 2003; Rasinski, 2004b). This view of reading fluency focuses on the reciprocal relationship between fluency and comprehension. More specifically, this view considers four dimensions of fluency: oral reading rate, oral reading (decoding) accuracy, quality of oral reading (prosody) and reading comprehension.

Among the constructs of ORF, prosody means the expressiveness of the language (Dowhower, 1991; Kuhn & Stahl, 2003; Schwanenflugel & Benjamin, 2012). As there has been growing agreement among researchers that fluent readers can read with phrasing and expression, accuracy and smoothness, and pacing, prosody has been emphasized as the characteristic of fluency. While prosody is surely linked to other aspects of reading fluency as well as reading comprehension in developing readers (Miller & Schwanenflugel, 2006, 2008; Schwanenflugel & Benjamin, 2012), there have also been ongoing arguments as
to whether prosody should be a part of the assessment of ORF (Kuhn, Schwanenflugel, & Meisinger, 2010). Furthermore, the previous studies on the measurement of ORF both in L1 and L2 contexts reveal that prosody does not have explanatory power on reading comprehension (Fuchs et al., 2001; Lems, 2006). For these reasons, the examination and the literature review of prosody are excluded in the current study.

Therefore, the literature regarding the reading rate, word recognition accuracy, and reading comprehension is reviewed. Additionally, studies on ORF in the L2 setting are presented. What follows is a detailed discussion of the constructs of reading fluency and it starts with an introduction to the reading rate.

### 2.2.1 Reading rate

Reading rate is a crucial component of reading fluency because it indicates the ease or efficiency with which a student can process the text of different difficulty levels. As fast reading rate is only possible when readers can recognize words efficiently and automatically, it is sometimes defined as automaticity or word recognition efficiency. Educators looking for a way to assess students’ reading fluency have considered reading rate as an important index of it. Thus, in the L1 setting, there have been attempts to establish reading performance norms based on the reading rate (Daane, Campbell, Grigg, Goodman, & Oranje, 2005; Hanbrouk & Tindal, 2006). In ORF measure, reading rate was calculated based on the number of Words read Per Minute (WPM) or on
the number of Words Correct Per Minute (WCPM), and the percentile scores of elementary students’ reading rate were provided to the stakeholders. In particular, students’ reading rate has been measured using the standardized Curriculum-Based Measurement (CBM) and the accumulated data of ORF are used for decision-making purposes, i.e., screening, diagnostic, progress monitoring, and outcome measure (Kame’enui, 2002).

The systematic and experimental research on reading rate has been conducted by tracking the eye movement of skilled L1 readers. Eye cameras precisely measure the time required to identify words in context and isolation (Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982). These studies assumed that eye fixations and patterns of saccades carry important information that reveals readers’ cognitive processing. Some of the findings discovered thus far by the eye-movement studies on reading rate are as follows. First, it was reported that in silent reading, skilled L1 readers can read 300 words per minute or 5 words per second and their word recognition in context is assumed to be highly correlated with recognizing words in isolation. This means that as the rate of word recognition increases, the text reading rate will become faster. In addition, with regard to oral reading rate, Rayner and Clifton (2009) found that oral reading rates are generally slower than silent reading rates with the reading rate in the range of 150-200 words per minute, which approximates normal speech rates. Second, the average fixation duration during reading among skilled L1 readers of English is likely to range from 225 to 250ms (Rayner, 2009), and the fixation duration may fluctuate depending on the properties of the fixated word,
such as the length or the difficulty. Third, the level of text difficulty, the text genre or readers’ reading skill may affect the fixation duration. For instance, as a text becomes more difficult to understand, fixations become longer, saccades become shorter, and regression occur more frequently (Rayner, 1998). Hence, multiple fixations and long durations are often associated with comprehension difficulties (Paulson, 2000; Rayner, Pollatsek, Ashby, & Clifton, 2012). Lastly, the difference of reading rate can be explained in terms of a reader’s perceptual span. The perceptual span is reported to be asymmetric, 15 letter distances to the right from the fixation location, and the right side of the fixation is called the parafoveal vision. According to researchers (Rayner, Well, & Pollatsek, 1980; Rayner, Well, Pollatsek, & Bertera, 1982), more skilled readers have the wider perceptual span and they can obtain information of the partial words from parafoveal vision. Specifically, when a reader processes the phonological or lexical information of the word (n) being fixated, s/he can also check the familiarity of the next word (n+1), which is within the parafoveal vision. As the word n+1 is processed orthographically in the parafoveal vision before it is fixated, the time for processing it is saved, leading to the faster speed of reading. Such a tight linkage between word identification and eye fixation was attempted to be explained by the EZ Reader model (Rayner, Li, & Pollatsek, 2007; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Pollatsek, & Rayner, 2006; Reichle, Ryner, & Pollatsek, 2003).

From these research bodies, it can be summarized that the reading rate is closely related to word recognition skills which perceive a word and processes
its information. In addition, the reading rate can be slower or faster depending on the difficulty of the words or the genre of the text. More importantly, it can be an indicator for reading comprehension in that the slower passage reading rate means the slower rate of word recognition, which can be the symptom of the difficulty of reading comprehension (Morris et al., 2013, 2017).

2.2.2 Word recognition accuracy

Oral reading accuracy, which represents the accurate word recognition or the phonological decoding skill, is another essential construct of reading fluency (Fuchs, Fuchs, & Deno, 1982; Rasinski, 2004a, 2006; Samuels & Farstrup, 2011). Obviously, accurate word recognition is reported to have a strong relationship with the development of reading fluency (Nation, 2009). Morris et al. (2011) described in the empirical research that reading accuracy is the second significant predictor of the reading rate, following word-recognition speed.

Understanding the significant role of effortless and accurate word recognition in successful reading comprehension, reading researchers in the L1 context have attempted to delineate the components of word recognition (Coltheart, 2005; Harm & Seidenberg, 1999, 2004; Perfetti, 1999, 2007; Perfetti & Hart, 2001). Words can be segmented into their components at multiple levels—such as grapheme, phonemes, and morphemes (Koda, 2005). Specifically, readers should first become aware of the correspondence of written symbols with speech units, activate appropriate semantic and syntactic resources, and access
her or his mental lexicon (Dehaene et al., 2010; Ehri, 1992, 1995, 2005, 2014; Koda, 2005; Perfetti, 1999, 2007). Each of different knowledge operates in association with the others so that readers are able to read the words successfully.

To reiterate, as echoed in the automaticity theory (LaBerge & Samuel, 1974), reading cannot proceed effectively without accurate and automatic word recognition skills. For this reason, L1 researchers and educators have observed how readers develop word recognition skills. As far as the reading development stage related to word recognition is concerned, Ehri’s (1992, 1995, 2005) Phase theory is noteworthy. Ehri (2005) had developed a carefully researched theory of the stages through which readers systematically progress in order to achieve fluency, starting from the pre-alphabetic phase, through to becoming partial alphabetic readers, further to becoming full alphabetic readers, and finally moving on to become consolidated alphabetic readers. This theory concerning knowledge acquisition of grapheme-phoneme correspondence postulates phonological awareness facilitates the formation of connections between letters, sounds, and meaning of words in memory. In actuality, L1 readers who reach the consolidated alphabetic stage can recognize more words accurately and accumulate a larger lexicon of sight vocabulary. Attempts to verify this theory has been made in the L2 context and it turned out that this developmental phase is plausible in explaining the L2 word recognition development. For instance, Jeon (2016) and Yin, Anderson and Zhu (2007) who explored the possibility of application of Ehri’s model to L2 learners concluded that L2 readers follow the four phases of English word recognition and those at more advanced phase were
able to recognize more words accurately and quickly.

In addition, what kind of decoding errors developing L1 readers made has been observed in order to reflect this information for planning the reading instruction. Notably, many researchers have committed to analyzing oral errors students make during reading and tried to draw inferences about the reasons for making the errors (Beck & Juel, 1995; Goodman, 1965, 1967; Hosp & Fuchs, 2005; LaBerge & Samules, 1974). The most influential model was from Goodman (1967). In a key study, Goodman (1965) viewed the reader as a user of language who processes three kinds of information, graphophonic, syntactic, and semantic, as s/he reacts to the graphic display on the page. He found that children identified words more accurately in context than in isolation, and the psycholinguistic reading process can be revealed in comparing unexpected responses in the oral reading context to expected responses. Hence, Burke and Goodman (1968) first coined the term, *miscue*, to refer to oral reading errors which are “mistakenly cued by the cognitive and linguistic systems of readers as they interact with the text (Leu, 1982, p. 423). The frame of miscue analysis was presented for the depth analysis of oral reading phenomena. By using the taxonomy developed by Goodman and his colleagues (1969), educators and researchers observed the learner read passages and attempted to diagnose and identify the learner’s reading problems. Published informal reading inventories contained the frame of error analysis, known for miscue analysis, which reflects Goodman’s theory. Even though at present Goodman’s reading theory is presumed to be inaccurate, the frame of miscue analysis has provided valuable
information regarding readers’ word recognition development. The synthesis of
L1 readers’ decoding errors coming from the studies of the miscue analysis was
provided by McKenna and Stahl (2009, p. 52), which is presented in Table 2.1.

Nevertheless, since there is a paucity of research on L2 decoding errors,
it is not clear whether L2 readers show the same pattern of decoding errors which
L1 readers commit. However, as it turned out from the previous research that L2
readers follow the developmental phases of L1 word recognition (Jeon, 2016;
Yin et al., 2007), certain patterns of decoding errors are presumed to emerge.
Thus, the present study explores what types of decoding errors are shown and
whether there are variances according to students’ reading abilities.
Table 2.1
The synthesis of the L1 readers’ decoding errors

| 1. | About half of first graders’ miscues are omissions. |
| 2. | As children mature, substitutions become more common. |
| 3. | Initially, substitutions may amount to mere guesses, often with a word beginning with the same letter. |
| 4. | Substitutions gradually become more sophisticated, bearing greater phonic resemblance to the actual word. |
| 5. | Substitutions begin as real words (for beginning readers) but often become nonword (i.e., phonic approximations). |
| 6. | Insertions are always rare (10% or fewer of all miscues). |
| 7. | Letter reversals (such as ‘was’ for ‘saw’) are common in first grade but typically disappear during second grade. |
| 8. | Reliance on context is heaviest for beginning readers and older poor readers. |
| 9. | Repetitions can be a mark of comprehension monitoring— that is, of a readers’ attempt to make print make sense. |
| 10. | Word-perfect reading does not always mean good comprehension. |

*Note. These decoding errors were synthesized by McKenna and Stahl (2009).*
2.2.3 Reading comprehension

As noted thus far, ORF has a close relationship with reading comprehension (Fuchs et al., 2001; Jenkins, Fuchs, Van Den Broek, Espin, & Deno, 2003; Schwanenflugel et al., 2006). Accordingly, there has been growing consensus that the comprehensive assessment of ORF must encompass measures of reading rate, accuracy, and quality, and these constructs of fluency must be measured in the context of reading comprehension.

Even though the widely utilized assessment for measuring ORF in the L1 setting is the Curriculum-Based Measurement (CBM), in which ORF is assessed while a teacher listens to a student’s reading aloud from unpracticed passages for one minute, it has been criticized for focusing on only two components of fluency: rate and accuracy. In fact, several studies conducted to investigate the ‘word caller’ have raised the questions regarding the utilization of CBM for the high-stakes decision-making purposes.

A word caller refers to a learner who efficiently decodes words but does so without comprehension (Meisinger, Bradley, Schwanenflugel, & Kuhn, 2010; Stanovich, 1986). In other words, they call out words in the text without understanding the meaning of the text. While previous studies have suggested that there are only a small portion of word callers among L1 elementary students (Hamilton & Shinn, 2003; Meisinger, Bradley, Schwanenflugel, Kuhn, & Morris, 2009; Schwanenflugel & Knapp, 2015), a small number of studies reported that this is not the case in L2 reading. For instance, Quirk and Beem (2012) reported that word callers may exist in appreciable numbers regardless of the grade level.
in L2 reader populations. In their study, they identified a substantial portion of word callers (15.8%) within the sample of the participants.

With regard to the word callers in L2 reading, there are issues on the relationship between the reading rate and reading comprehension. Research on L2 reading fluency found that reading fluency (represented by reading rate) develops faster than reading comprehension (Geva & Wang, 2001; Nakamoto, Lindsey, & Manis, 2007; Quirk & Beem, 2012). For instance, Quirk and Beem (2012) found that L2 readers exhibit the significant gap between reading fluency and reading comprehension. Therefore, they attempted to broaden the definition of the word caller by including readers whose reading comprehension lags behind the reading fluency. It was reported from their studies that 55.5% of the total participants had reading fluency scores that were significantly higher than their reading comprehension scores.

In a similar vein, Meisinger et al. (2009) address that in the L1 context the number of word callers tends to increase to nearly 10% among older elementary school children. They point out that fluent reading does not guarantee reading comprehension as reading texts of upper elementary grades requires more sophisticated vocabulary, background knowledge, and metacognitive skill.

Given the inherent complexity of reading fluency (Schwanenflugel et al, 2006; Stanovich, 2000) and the multitude of subskills that affect reading comprehension, it is difficult to sort out the specific factors that might contribute to the gap between reading comprehension and reading fluency. However, as the previous research shows, if L2 readers are assumed to have the different
developmental profile of reading from L1 readers, the gap between reading comprehension and the reading rate is worth being noted.

2.2.4 Research in L2 reading fluency

To date, reading fluency has been receiving substantial attention from both researchers and practitioners in L1 settings. However, in the L2 context, a dearth of research addressing L2 readers’ reading fluency has been conducted. Among influential studies on L2 fluency are those conducted by Taguchi and his colleagues (Gorsuch & Taguchi, 2008; Gorsuch & Taguchi, 2010; Taguchi, Takayasu-Maass, & Gorsuch, 2004) investigating the training effects of fluency instruction on the reading rate and comprehension development in a series of studies. Other researchers, on the other hand, examined the relationship between different aspects of ORF among adult L2 readers (Jeon, 2012; Jiang, Sawaki, & Sabatini, 2012; Lems, 2006, 2012).

Most notably, Lems (2006) made the first significant start in the study of L2 reading fluency by exploring validation of ORF as a proficiency measure among L2 readers who had different English proficiency levels and various L1 backgrounds. The result of the study revealed that English proficiency score and oral passage reading score were in the positive relationship. However, the experiment revealed that although there is a significant relationship between oral reading fluency and reading comprehension, the correlations between the variables were small or moderate (0.27 - 0.41) in size. Thus, Lems (2006)
concluded ORF can become a good index of silent reading fluency after some requisite skills are over a certain threshold: listening comprehension and oral language proficiency.

Likewise, Jiang et al. (2012) examined the relationship between word reading efficiency, text reading fluency, and reading comprehension for adult Chinese EFL learners preparing to take the Test of English as a Foreign Language (TOEFL). The finding showed a strong relationship between text oral reading fluency and reading comprehension, and text oral reading fluency alone explained 26% of silent reading variance.

Another study on the role of ORF in L2 reading was conducted by Jeon (2012). In the study, 255 Korean high school students were assessed on three ORF variables and six other reading predictors. The results indicated that the three ORF variables collectively explained 21.2% of the variance in silent reading comprehension. Oral passage reading fluency alone explained 20.9% of the variance in silent reading comprehension.

Even though the three above studies concluded that passage reading fluency was a much stronger predictor of reading comprehension than word reading efficiency variables, none of the studies offered a strong support for oral reading fluency as a proxy for reading comprehension. However, these results of the studies need to be reexamined in terms of the fluency measure used in their studies. The studies only measured passage reading speed and accuracy among ORF constructs, which has been criticized by reading researchers. In particular, Alt and Samuels (2011) strongly argued that the field had misunderstood the
definition of fluency and it led to invalid assessment of fluency that actually measured reading speed only. They argued that ORF measure must not only assess the characteristics of speed and accuracy but also explicitly assess prosody and reading comprehension. Similarly, Morris et al. (2011, 2013, 2017) contended reading fluency needs to be explored in the reading-for-meaning context and only then, reading fluency can be a proxy for reading comprehension.

Furthermore, as Quirk and Beem (2012) put it, there exists the significant gap between reading fluency and reading comprehension in L2 reading. Accordingly, if ORF is measured when L2 readers focused on the reading comprehension, different results can be drawn. It is necessary to examine the L2 reading fluency by employing different research methodology in order to expand understanding of L2 readers’ actual reading fluency. Amid the problems and necessities mentioned above, further studies on L2 reading fluency are called for.
2.3 Reading Assessment

In this section, two reading assessments are reviewed. The review begins by describing standardized testing, followed by the reviews of informal reading inventory, which is closely related to the current study.

2.3.1 Standardized testing

Standardized testing is historically and presently the most common form of comprehension assessment in the L2 reading test. Test developers of the standardized tests have made an effort to reflect the construct of reading comprehension abilities and to capture crucial aspects of the component abilities of reading comprehension (Davis, 1944, 1968). The typical format of the test comes in the form of a short passage accompanied by several multiple-choice questions with a single predetermined correct answer. While this test format makes it possible to assess the language learners’ level of reading comprehension within the time and cost constraints, a number of problems have been posed as well.

One of the criticisms of the standardized testing is that it is less straightforward to capture the reading abilities which test-takers are assumed to have to complete the test task. For example, Field (2011) points out “the dangers of relying exclusively on an approach that attempts to track back from a product or outcome to the process that gave rise to it” (p. 79). Indeed, such test by its
nature says little about what is really happening when a reader processes the text under the test conditions. Furthermore, the actual processes which a reader might be expected to apply in real-world reading are not the concern of the standardized test (Weir & Khalifa, 2008). The primary concern of the standardized test is the capabilities that a reader is assumed to need in order to tackle certain test items. In other words, this test format has not elicited the cognitive processing involved in the contexts beyond the test itself, which is the non-test reading context. Thus, little reference has been made to the cognitive processing in the test development and the interpretation of the test score.

With regard to the L2 reading assessment and instruction, these limitations associated with standardized testing have posed serious problems. For a long time, there have been few attempts to examine and assess how L2 readers actually undergo the comprehension processes while reading. Furthermore, as the standardized testing often represents high-stakes testing affecting students’ future educational goals, teaching the test items or test-wisenedness strategies has been taken for granted in the majority of English reading classes in the L2 setting. Accordingly, the process of students’ meaning construction has been overlooked in the English reading classroom.

These posed problems call for an alternative reading assessment to provide information regarding actual reading abilities that L2 readers have. Hence, the current study investigates L2 reading using informal reading inventory which is widely used for a diagnostic purpose in the L1 educational context.
2.3.2 Informal reading inventory

As a reading assessment method, the measurement of ORF has many strengths (Fuchs et al., 1988). First, oral passage reading can be performed using any reading material and does not require extensive tester-training for administering it. Second, scoring can be done promptly and with a higher inter-rater reliability (Jeon, 2012; Jiang et al., 2012; Lems, 2006; Taguchi et al., 2004). Third, unlike silent reading where the test-takers’ reading performance is invisible, oral reading can provide a transparent observation of the test-takers’ reading process (Jeon, 2012).

Aside from the CBM, a commonly used method for measuring ORF is to employ informal reading inventories (IRI). The IRI which is informal, criterion-referenced assessment— as opposed to standardized, norm-referenced assessment—has a long history and a prominent place in the field of L1 reading education for more than 60 years (Morris et al., 2011). It is the individually administered diagnostic assessment designed to evaluate different aspects of students’ reading performances and many published IRIs are practically used in the L1 reading class. The essential function of IRI is to help the teacher determine students’ independent, instructional, and frustration levels of reading. Identifying students’ reading level is important as to make a proper match between the student and the book s/he reads. Also, it helps the teacher to learn about the student’s reading strength and weakness. From listening to the student’s reading aloud, the teacher can hypothesize possible skill deficiencies such as word recognition, reading speed, or comprehension.
Three levels of IRI (i.e., independent, instructional and frustration level) are determined by the criteria established by Betts (1946), and many IRIs still use these criteria. Table 2.2 displays the performance criteria for reading accuracy and reading comprehension.

Table 2.2
Performance criteria for reading accuracy and comprehension
(Betts, 1946)

<table>
<thead>
<tr>
<th></th>
<th>Reading Accuracy</th>
<th>Reading Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent level</td>
<td>98%-100%</td>
<td>and 90%-100%</td>
</tr>
<tr>
<td>Instructional level</td>
<td>95%-97%</td>
<td>and 75%-89%</td>
</tr>
<tr>
<td>Frustration level</td>
<td>90% or below</td>
<td>or Below 50%</td>
</tr>
</tbody>
</table>

The independent level is the highest level at which an individual can read on his or her own without assistance. When the student reads the text at this level, s/he can read the text with 99% of accurate pronunciation and answer minimum 90% of comprehension questions. This level is important as students need to be instructed to choose books at this level to read on their own. The instructional level is the teaching level at which the reading text should be neither challenging nor easy. At this level, the student should obtain a minimum
75% of comprehension score and 95% of reading accuracy score. While there is only one level of independent and frustration level, the instructional level can span one to three levels depending on the content area of the text or students’ interests, background information, and experience. In addition, there could be the area called the buffer zone that falls between the instructional level and the frustration level (Rubin, 2002). To demonstrate, if students’ word recognition score ranges between 94% and 91% and comprehension score ranges between 74% and 51%, they are judged to be in the buffer zone. At the frustration level, the student reads with many word recognition errors (10%) and with less than 50% of comprehension. It is the lowest level and the reading material at this level should be avoided as the text difficulty could frustrate the student’s willingness to read.

As such, IRI administration is specifically matched with the needs of individual readers. With regard to this, Paris and Carpenter (2003) describe IRIs as all positive characteristics of classroom assessment. Furthermore, several research bodies have been conducted to verify the reliability of IRIs (Paris, Paris, & Carpenter, 2002; Paris, Pearson, Carpenter, Siebenthal, & Laier, 2002; Spector, 2005). In the study conducted by Paris et al. (2002), the reading passages of the identical level from two IRIs were verified and both IRIs showed positive and significant test-retest reliability. The same evaluation project also revealed acceptable concurrent validity in the correlations \( (r = .48 - .90) \) between a reading comprehension test and two IRIs. Since most commercial IRIs are produced based on acceptable levels of reliability and validity, it is possible to
say that the data derived from IRIs can be used to assess individual growth and to document reading achievement.

In L2 reading research, Taguchi et al. (2004) measured reading rate by utilizing the Informal Reading Inventory (Burns & Roe, 2011) to investigate the effect of repeated reading on L2 reading fluency. Nevertheless, there is general lack of research in measuring L2 ORF using informal reading inventories. Therefore, the current research investigates how L2 oral reading fluency should be assessed or whether the IRI made in L1 setting could be applied in order to examine or diagnose the development of L2 reading skill.

2.4 L2 Experiential Factors for Reading Fluency

There are many differences between L1 and L2 reading contexts, resulting in complexities of L2 reading comprehension. In order to illuminate the unique nature of L2 reading processing, qualitative and quantitative variations in L2 reading experience need to be incorporated into L2 reading research.

One of the important learner-specific factors which are crucial to L2 reading proficiency is the sheer amount of L2 exposure to print. With respect to L2 word recognition, Koda (1996) states that the amount of L2 processing experience is one of the most critical factors for the word recognition development. Word recognition studies, in fact, demonstrate that with increased experience, processing speed improves (Favreau & Segalowitz, 1982; Haynes &
Carr, 1990) and the error rate decreases (Segalowitz & Segalowitz, 1993). These studies make it plain that the performance improvement of reading is achieved through increased L2 processing experiences.

Similarly, Logan (1997) suggests that automaticity develops as a consequence of the ‘Power Law’. According to Logan’s (1997) theory, as students read words, they lay down traces for each word in their memories. If the word is read frequently enough, the cumulative practice with that word results in an increased likelihood that the word becomes recognized upon further exposures and that the speed with which it will be recognized will increase. Stanovich (1980) also contributed to the contemporary focus on reading fluency. He demonstrated a clear relationship between fluency and the amount of reading in which a reader is engaged. Thus, readers who achieved fluency are likely to read more extensively than those who lack fluency.

Aside from specific linguistic and individual differences, a number of socio-cultural factors affect students’ learning (Grabe & Stoller, 2013). Especially, home literacy environments such as parental intervention and accessibility to the reading material are known to contribute to the literacy development (e.g., Evans & Shaw, 2008; Snow, Burns, & Griffin, 1998). Although most of the studies on the home environmental factors have been conducted in the L1 reading context, their results indicate that these factors all contribute to L2 literacy development, by directly or indirectly affecting the opportunities for extensive L2 reading practice, or providing motivation toward learning to read in L2 (Grabe, 2009). Previous studies concerned with Korean
context also back up the argument that supportive home literacy environment strongly is correlated with children’s literacy development (An, 2013; Han, 2002; Yu, 2008).

In learning to read in L2, a wide range of experiential factors are involved. It is, thus, necessary to look into crucial factors contributing to literacy development (i.e., L2 literacy environment, reading engagement in L2, the amount of reading and print exposure to L2, and the length of residency in L2-speaking countries) in order to investigate the experiential differences involved in reading fluency.

In sum, reading comprehension is critically dependent on the reader’s fluency in identifying printed words, as reflected in the readers’ accuracy and speed in identifying words in a passage as well as in his or her ability to read smoothly with appropriate expression. As the definition of fluency encompasses all levels of reading processes from word recognition to reading comprehension, fluency is the prerequisite if learners are to succeed at the primary purpose of reading. Nevertheless, it has not drawn much attention from L2 researchers unlike the L1 research where reading fluency has been a very ‘hot topic’ (Cassidy & Cassidy, 2009) since the Report of National Reading Panel (2000) was published. Therefore, it is necessary to investigate how fluently L2 students can read, how their accumulated reading experience affects reading fluency. Specifically, the basic information about how fast and accurately L2 readers could read with comprehension and how L1 reading fluency measures can be applied to L2 students need to be investigated. By doing so, more specific
empirical evidence can be obtained about how fluent L2 readers are and where difficulties arise in their reading processes.
CHAPTER 3
METHODOLOGY

This chapter introduces the methodology adopted in the present study. It begins with the detailed description of the participants (3.1), followed by the procedures to collect L2 oral reading fluency data (3.2) and L2 reading experiences (3.3). Finally, the data analysis method (3.4) is presented.

3.1 Participants

A total of 68 students from four high schools in Gyeonggi province participated in this study. Out of the four schools, one high school was located in Yeoju-city and the other three schools were in the different districts of Seognam-city. As the students volunteered to participate in this study, the factors such as the number of participants from each school, the English proficiency and the gender were not considered. Out of the 68 participants, 26 students from school A and school B respectively participated in the study. Ten students were from school C and six from school D. Among 68 participants, 34 are male and 34 are female and all of them are Koreans.

They were 16 years old when the study was conducted and were attending their second year of high school (equivalent to U.S. eleventh grade). They had studied English as a foreign language for approximately nine years mostly by
formal schooling as most Korean students are required to learn English as a mandatory school subject from third grade of their elementary school. They had four hours of English education per week in school at the time of data collection. Only two students had experience of studying abroad in English speaking countries. One student had experience of living in an English-speaking country for four years and had a formal education in the elementary and the middle school. The other student reported that she had attended the elementary school for a year in an English-speaking country. Despite their overseas experiences, they were not excluded from the data analysis, since they received English education mostly in Korea.

The information regarding the participants’ English reading proficiency level was obtained from their homeroom teachers after receiving consent from the participants in order to investigate how their English proficiency levels were placed based on standardized, multiple-choice reading assessments. The information regarding the participants’ English proficiency level was the result of the mock College Scholastic Ability Test (MCSAT) administered in June 2017. The detailed English proficiency levels are presented in Table 3.1.
Table 3.1

<table>
<thead>
<tr>
<th>Levels</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>The number</td>
<td>14</td>
<td>7</td>
<td>9</td>
<td>9</td>
<td>16</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>68</td>
</tr>
</tbody>
</table>

Even though the participating students were not controlled in terms of their English proficiency levels, their levels were in a diverse range from high-achieving students to low-achieving ones.

### 3.2 Oral Reading Fluency Measure

In this section, the research instrument and the procedure for measuring ORF are described. In order to collect data for ORF, an informal reading inventory (IRI) was employed. A description of the task is described as well.

#### 3.2.1 Test instrument

In order to investigate and diagnose the participants’ ORF, the present

---

1 This mock-CSAT was hosted by Busan Metropolitan City Office of Education and administered on June 2017. According to Korea Institute for Curriculum and Evaluation (2017), the grade in the MCSAT is assigned as a range from 1 to 9 by a percentage of a total number of questions answered correctly. For example, 90 points or higher fall into the first grade; scores between 80 and 89 points belong to the second grade.
study employed IRI. Among well-known published informal reading inventories, the Informal Reading Inventory (Burns & Roe, 2011) was selected for this study. This IRI contains a series of carefully graded test forms for all reading grade levels from preprimer (PP) through twelfth grade, and four forms are provided at each reading level. These test forms are considered to be effective in measuring reading fluency of students in the U.S. According to the publisher’s manual, the reading passages in the IRI are selected primarily from graded materials in basic readers and literature books used in schools with students at each grade level. Pictures are not accompanied in this IRI.

For the selection of this IRI, following things were taken into consideration: interchangeability of test forms, the range of grade levels and the length of the passages. As informal reading inventories are produced for L1 readers and have not been validated for measuring L2 reading fluency (Grabe & Jiang, 2014), some of the reading tasks and passages needed to be excluded because the content of some passages required the culture-specific background knowledge for reading comprehension. Moreover, since reading comprehension was to be measured based on the recall protocol (without students’ referring to the passage), the number of words in each passage was controlled so that the reading task of the lower grade level does not have longer passages than that of the higher grade level. For these reasons, test forms for each grade needed to be reassembled. Among the available IRIs, this particular IRI has four parallel test forms, each of which can be used interchangeably (Pikulski, 1990). Furthermore, this IRI covered materials ranging most widely in terms of difficulty from
preprimer (PP) through the twelfth grade (Pikulski, 1990). Therefore, this IRI was selected among published IRIs.

In order to measure the participants’ ORF, the researcher reassembled the reading tasks of the IRI and made two sets of test forms: Form A and Form B. When the reading tasks for each test form were selected, the passage factors contributing the reading difficulty, such as the number of words, the content, and the genre (Morris et al., 2013) were considered. Each form was comprised of fourteen grade-level reading tasks. Form A was primarily used for measuring ORF and Form B was used as a supplementary reading task only when the participants’ reading performance at a certain grade level was needed to be verified again.

The passages of reading tasks in Form A hierarchically ranged in terms of length and difficulty from the grade PP passage that contains 60 words to the twelfth-grade passage that contains 195 words. According to the publishers’ manual, the reading level of the passages was assessed by readability formulas and during the field testing, these passages were readjusted for increasing difficulty in word recognition and comprehension. The Spache Readability Formula was used to assess the difficulty of the selected texts for grade PP through grade 3, and the Fry Readability Formula was used to check selections for grade 4 through grade 12. Even though the readability of texts was verified by the publishers, both Flesch-Kincaid formula and Lexile measure were utilized to provide more reliable information about the grade level readability and supplement the limitations of the readability formula. The detailed information
on the reading passages is presented in Table 3.2.

As shown in Table 3.2, there were inconsistencies in the readability indices of Form A among the readability formulae. For example, the passage of level 1 had higher Lexile measure than the passage of level 2 and 3, and the passage of level 5 had lower readability indices than level 4. However, the increasing difficulty was assured in the actual administration of the tests. In fact, the passages of the higher grade level in the IRI appeared more difficult for the students to read than those of the lower grade level.
Table 3.2

The readability indices of Form A

<table>
<thead>
<tr>
<th>Grade Level</th>
<th>Source</th>
<th>The Number of Words</th>
<th>Spache / Fry</th>
<th>Flesch-Kincaid Grade Level</th>
<th>Lexile Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preprimer</td>
<td>-</td>
<td>60</td>
<td>1.2</td>
<td>-</td>
<td>100L-200L</td>
</tr>
<tr>
<td>Primer</td>
<td>-</td>
<td>94</td>
<td>1.4</td>
<td>-</td>
<td>100L-200L</td>
</tr>
<tr>
<td>Level 1</td>
<td>“City Magic”</td>
<td>108</td>
<td>1.8</td>
<td>-</td>
<td>400L-500L</td>
</tr>
<tr>
<td>Level 2</td>
<td>“I love Gram”</td>
<td>120</td>
<td>2.1</td>
<td>-</td>
<td>300L-400L</td>
</tr>
<tr>
<td>Level 3</td>
<td>“Go to Hawaii”</td>
<td>133</td>
<td>2</td>
<td>-</td>
<td>300L-400L</td>
</tr>
<tr>
<td>Level 4</td>
<td>“Pets”</td>
<td>142</td>
<td>5.5</td>
<td>4.0</td>
<td>800L-900L</td>
</tr>
<tr>
<td>Level 5</td>
<td>“How to Wreck a Building”</td>
<td>131</td>
<td>4.5</td>
<td>4.0</td>
<td>700L-800L</td>
</tr>
<tr>
<td>Level 6</td>
<td>“The Real Me”</td>
<td>143</td>
<td>6.0</td>
<td>6.3</td>
<td>900L-1000L</td>
</tr>
<tr>
<td>Level 7</td>
<td>“The Slave Who Bought His Freedom”</td>
<td>155</td>
<td>7.0</td>
<td>7.7</td>
<td>1000L-1100L</td>
</tr>
<tr>
<td>Level 8</td>
<td>“Across Five Aprils”</td>
<td>159</td>
<td>8.0</td>
<td>6.3</td>
<td>800L-900L</td>
</tr>
<tr>
<td>Level 9</td>
<td>“The Day the Sea Went Down the Drain”</td>
<td>170</td>
<td>7.0</td>
<td>8.1</td>
<td>1100L-1200L</td>
</tr>
<tr>
<td>Level 10</td>
<td>“The Buried Treasure of Oak Island”</td>
<td>184</td>
<td>8.0</td>
<td>6.1</td>
<td>900L-1000L</td>
</tr>
<tr>
<td>Level 11</td>
<td>“Number One Son”</td>
<td>190</td>
<td>9.0</td>
<td>8.9</td>
<td>1000L-1100L</td>
</tr>
<tr>
<td>Level 12</td>
<td>“An Astrologer’s Day”</td>
<td>195</td>
<td>10.0</td>
<td>10.9</td>
<td>1100L-1200L</td>
</tr>
</tbody>
</table>
The participants’ reading comprehension was assessed with reading comprehension questions after reading the passages. Each reading passage from grade PP through grade 3 contained eight comprehension questions, while the other graded passages included ten comprehension questions. The comprehension questions were open-ended questions and designed to elicit readers’ comprehension of main ideas, specific details, inferences, sequence, cause-and-effect, and vocabulary. Even though the original version of questions were utilized for the comprehension check, vocabulary questions of the IRI were excluded according to the previous study (Taguchi et al., 2004), reporting that vocabulary questions of the IRI are not suitable for L2 readers because there is a possible difference in vocabulary repertoire between L1 and L2 readers. A set of sample questions of the level 3 is presented in Table 3.3.
Table 3.3

Sample comprehension questions (Level 3)

<table>
<thead>
<tr>
<th>Question Type</th>
<th>Sample Comprehension Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>main idea</td>
<td>What is this story about?</td>
</tr>
<tr>
<td>detail</td>
<td>In what season does this story take place?</td>
</tr>
<tr>
<td>cause and effect</td>
<td>What caused Dunk not to want to play basketball?</td>
</tr>
<tr>
<td>inference</td>
<td>What have they done for the last six days?</td>
</tr>
<tr>
<td>inference</td>
<td>What part of the house were Slam and Dunk in at the beginning of the story?</td>
</tr>
<tr>
<td>sequence</td>
<td>What room do Slam and Dunk go to after Dunk suggests going swimming?</td>
</tr>
<tr>
<td>inference</td>
<td>How does Grandpa feel when Slam and Dunk see him?</td>
</tr>
<tr>
<td>inference</td>
<td>Which guess about what Grandpa had in his hands was best? Why do you say so?</td>
</tr>
<tr>
<td>inference</td>
<td>What did Grandpa have in his hands?</td>
</tr>
</tbody>
</table>

In order to compare participants’ reading performances on ORF with the mock-CSAT (MCSAT), the passages of the MCSAT were assessed with the Flesch-Kincaid readability formula and the Lexile Measure. The data and the descriptive statistics are provided in Tables 3.4 and 3.5.
Table 3.4

The readability indices of the MCSAT reading passages

<table>
<thead>
<tr>
<th>Question Number</th>
<th>The Number of Words</th>
<th>Flesch-Kincaid Grade Level</th>
<th>Lexile Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>130</td>
<td>8.1</td>
<td>1000-1100</td>
</tr>
<tr>
<td>19</td>
<td>122</td>
<td>8.3</td>
<td>800-900</td>
</tr>
<tr>
<td>20</td>
<td>147</td>
<td>5.4</td>
<td>600-700</td>
</tr>
<tr>
<td>21</td>
<td>152</td>
<td>14.9</td>
<td>1400-1500</td>
</tr>
<tr>
<td>22</td>
<td>131</td>
<td>10.6</td>
<td>1100-1200</td>
</tr>
<tr>
<td>23</td>
<td>141</td>
<td>5.6</td>
<td>800-900</td>
</tr>
<tr>
<td>24</td>
<td>134</td>
<td>11.4</td>
<td>1200-1300</td>
</tr>
<tr>
<td>25</td>
<td>98</td>
<td>-</td>
<td>600-700</td>
</tr>
<tr>
<td>26</td>
<td>101</td>
<td>-</td>
<td>700-800</td>
</tr>
<tr>
<td>27</td>
<td>131</td>
<td>8.4</td>
<td>1000-1100</td>
</tr>
<tr>
<td>28</td>
<td>137</td>
<td>12.1</td>
<td>1000-1100</td>
</tr>
<tr>
<td>29</td>
<td>148</td>
<td>7.5</td>
<td>1000-1100</td>
</tr>
<tr>
<td>30</td>
<td>160</td>
<td>1.5</td>
<td>400-500</td>
</tr>
<tr>
<td>31</td>
<td>147</td>
<td>10.2</td>
<td>900-1000</td>
</tr>
<tr>
<td>32</td>
<td>139</td>
<td>9.3</td>
<td>1000-1100</td>
</tr>
<tr>
<td>33</td>
<td>153</td>
<td>10.7</td>
<td>1000-1100</td>
</tr>
<tr>
<td>34</td>
<td>147</td>
<td>8.4</td>
<td>1000-1100</td>
</tr>
<tr>
<td>35</td>
<td>152</td>
<td>5.8</td>
<td>900-1000</td>
</tr>
</tbody>
</table>
Table 3.5

Descriptive statistics of the MCSAT reading passages

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Range</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Means</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flesch-Kincaid Grade Level</td>
<td>23</td>
<td>13.40</td>
<td>1.5</td>
<td>14.90</td>
<td>8.5</td>
<td>2.86</td>
</tr>
<tr>
<td>Lexile Measure</td>
<td>25</td>
<td>1000</td>
<td>400-500</td>
<td>1400-1500</td>
<td>936-1036</td>
<td>244.74</td>
</tr>
</tbody>
</table>

The passage levels of MCSAT were widely dispersed based on the two readability formulae. When calculated with the Flesch-Kincaid Grade level formula, the mean grade level was 8.5, ranging from grades 1 to 14. The mean Lexile was about 936-1036, which is almost equivalent to the grade level 7.
3.2.2 Procedures

The test was administered on a one-on-one basis and the whole process of the measurement was recorded by the researcher. Before the participants began to read the first reading passage, the researcher instructed each participant to read at their normal speed, because comprehension questions will be asked about passages that they read. Next, the researcher provided a brief one-sentence introduction to the passage in Korean (e.g., “Read this to find out what two youngsters decide to do.”). As the participant commenced reading aloud the passage, the researcher checked the recorder to mark the starting time. The researcher followed along as the participants read orally to mark the decoding errors. When the last word in the passage was read, the time was marked again, and the researcher proceeded to ask reading comprehension questions. The participants had to answer the comprehension questions without referring to the passages. They continued to read additional passages until they reached a frustration level.

3.2.3 Scoring

Based on the collected data, the participants’ ORF level was determined. In general, a student’s three reading levels (i.e., independent level, instructional level and frustration level) were decided by the IRI criteria. Although the chief assessment concern of the IRI administration is to identify a students’ instructional level (Burn & Roe, 2011; Morris et al., 2013), the current research
aimed to identify the ORF/Independent Level (ORF-IL) and the ORF/Frustration Level (ORF-FL) as the instructional level can range across several grade levels depending on the content of the passages or readers’ background knowledge. ORF levels were determined by word recognition accuracy (WRA) scores and oral reading comprehension (ORC) scores. The criteria were imported from the IRI, but the score of ORC at the independent level was calibrated in order to ensure the minimum score. When a student reads the text with 99% of accurate pronunciation and answer minimum 85% of comprehension questions, that level was identified as his/her ORF-IL. In order to judge ORF-FL, either one of these WRA and ORC condition is satisfied. The criteria for WRA and ORC are presented in Table 3.6.

<table>
<thead>
<tr>
<th>Performance criteria for WRA and ORC</th>
</tr>
</thead>
<tbody>
<tr>
<td>WRA</td>
</tr>
<tr>
<td>Independent level</td>
</tr>
<tr>
<td>Instructional level</td>
</tr>
<tr>
<td>Frustration level</td>
</tr>
</tbody>
</table>

The passage reading yielded three scores: word recognition accuracy (WRA), oral reading rate (ORR), and oral reading comprehension (ORC). WRA was the percentage of words read correctly. The percentage of accuracy
depended upon the percentage of the decoding error. The percentage of the error was computed and subtracted from 100%, yielding WRA score for a given passage. In the previous research (Morris et al., 2013, 2017), the decoding errors were categorized into five types: substitution, omission, insertions, examiner-helps and self-correction. However, the present research excluded self-correction according to McKenna and Stahl (2009) and included mispronunciation in the error types. Mispronunciation was included as the error type since the participants were L2 readers, some of whom might not yet have developed phonological awareness necessary for word recognition. The decoding errors were collected at the independent level and two or more subsequent levels including the frustration level. In order to verify and categorize errors, the participants were asked whether they knew the meaning and the pronunciation of words after the assessment was completed.

ORR, expressed in Word read Per Minute (WPM), was computed for each passage read. The formula for computing the reading rate (WPM) was 60 multiplied by the number of words in a passage, which is divided by number of seconds.

\[
WPM = \frac{60 \times \text{the number of words in a passage}}{\text{seconds}}
\]

ORC score was the percentage of questions answered correctly. Immediately after finishing reading a passage, the participants were asked to
answer seven to nine questions on each passage. ORC score was only utilized for identifying the ORF level, but excluded from the data analysis.

Two raters were involved in the scoring procedure. They are English teachers who have twelve years of teaching experiences in high schools. The raters scored the students’ responses to reading comprehension questions, and identified and then classified the decoding errors. They scored individually and then met to compare scores and discuss scoring issues. The average rater agreement rates were 92% for reading comprehension, 96% for the word recognition score, and 89% for the classification of decoding errors.

The ORF levels were collapsed to form the following groups: High Fluency (HF), Intermediate Fluency (IF), and Low Fluency (LF) readers. These groups were formed to capture the participants at three distinct levels of English language development. When the fluency groups were formed, the distribution of the participants’ English proficiency levels across the ORF levels was considered. For example, the students whose English proficiency levels were levels 1 and 2 were mostly concentrated in the ORF level 5 through level 7. Likewise, the majority of the students who received the levels 3 to 5 in MCSAT were distributed at the ORF level 2 to 4. Among the students whose proficiency levels ranges 6 to 9 in MCSAT, only one student read the passage of higher than level 1 independently. As a result, the HF group included the readers whose ORF-IL was for upwards of level 5, the IF readers’ ORF-IL ranged from Levels 2 to 4, and that of the LF group was below Level 1. The summary for the distribution of the participants by the fluency level is provided in Table 3.7.
Table 3.7

Reading fluency group by the ORF level

<table>
<thead>
<tr>
<th>Reading Fluency</th>
<th>ORF Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Fluency</td>
<td>Level 5 and above</td>
</tr>
<tr>
<td>Intermediate Fluency</td>
<td>Level 2 ~ Level 4</td>
</tr>
<tr>
<td>Low Fluency</td>
<td>Level 1 and below</td>
</tr>
</tbody>
</table>

3.3 Semi-structured Oral Interview

The semi-structured oral interviews were administered to gain in-depth insight into which individual and experiential differences are involved in the development of L2 reading fluency. The interview was guided by Lindlof and Taylor (2002). By reviewing L2 reading development, a list of questions was generated in advance (see, Appendix D). Every student at each level was interviewed immediately after the test was performed and the interview lasted for about 10 to 15 minutes. For analysis and interpretation, all verbal report was transcribed firstly in Korean and then, translated into English.

3.4 Data Analysis

This section provides the description of the data analysis procedure. The data gathered through the IRI test administration and semi-structured oral
For the first research question, the data collected from the ORF measure were descriptively and statistically analyzed. Firstly, the dispersion of ORF levels is presented. In addition, in order to investigate the students’ distribution in ORF levels across the different proficiency level of the MCSAT, the correlation of the ORF-IL with the MCSAT level was calculated, using the Spearman correlation coefficient. Secondly, WRA and ORR were descriptively analyzed. Lastly, the decoding errors were analyzed quantitatively and qualitatively.

For the second research question, what kind of L2 reading experiences the L2 readers have based on their reading fluency levels, qualitative analyses for the semi-structured oral interviews were conducted for an in-depth investigation of their L2 reading experiences. The data collection and the analyses were conducted based on the methodology described in this chapter, and the following chapter presents the results of the study.
CHAPTER 4
RESULTS AND DISCUSSION

The analyses of research data gathered from the IRI administration and the following interview is presented in this chapter. Quantitative results from the data collected by the means of the face-to-face measurement of oral reading fluency (ORF) are examined, followed by the presentation of the findings from the interviews with the participants. These interviews were conducted on the basis of the qualitative content analysis.

4.1 Oral Reading Fluency Measure

The first research question was directed to the examination of L2 readers’ ORF. The results of the ORF measurement using the IRI are described and discussed in terms of the ORF level, word recognition accuracy (WRA), the decoding errors and oral reading rate (ORR).

4.1.1 The distribution of the ORF level

In order to investigate reading fluency levels of the high school students at eleventh grade, the oral reading tests included in the IRI were administered. As a result, the participants’ ORF/Independent Level (ORF-IL) and ORF/Frustration Level (ORF-FL) were identified and then the correlation of the participants’ ORF-IL with the MCSAT level was analyzed. First of all, the
distribution of the participants across the grade levels of ORF-IL and ORF-FL is summarized in Table 4.1 and presented graphically in Figure 4.1.

**Table 4.1**

**The distribution of the 11th-grade high school students across the ORF levels (N=68)**

<table>
<thead>
<tr>
<th>Grade Level</th>
<th>ORF-IL</th>
<th>N</th>
<th>Percent (%)</th>
<th>ORF-FL</th>
<th>N</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below PP</td>
<td>7</td>
<td></td>
<td>10.29</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP</td>
<td>6</td>
<td></td>
<td>8.82</td>
<td>7</td>
<td></td>
<td>10.29</td>
</tr>
<tr>
<td>P</td>
<td>5</td>
<td></td>
<td>7.35</td>
<td>5</td>
<td></td>
<td>7.35</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td></td>
<td>8.82</td>
<td>1</td>
<td></td>
<td>1.47</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td></td>
<td>17.64</td>
<td>8</td>
<td></td>
<td>11.76</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td></td>
<td>22.05</td>
<td>9</td>
<td></td>
<td>13.23</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td></td>
<td>7.35</td>
<td>12</td>
<td></td>
<td>17.64</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td></td>
<td>7.35</td>
<td>13</td>
<td></td>
<td>19.11</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td></td>
<td>7.35</td>
<td>4</td>
<td></td>
<td>5.88</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td></td>
<td>2.94</td>
<td>5</td>
<td></td>
<td>7.35</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td>4.41</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>1.47</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>68</td>
<td></td>
<td>99.96</td>
<td>68</td>
<td></td>
<td>99.96</td>
</tr>
</tbody>
</table>
The participants’ ORF-IL ranged across 10 levels from level Below-PP to level 7. Although level PP was the lowest level, seven participants could not satisfy the criteria of the independent level in reading the passages of level PP. In contrast, there was no participant whose ORF-IL was above level 7. In general, the largest proportion of the participants’ ORF-IL was concentrated at levels 2 and 3. Level 3 accounted for the highest portion of ORF-IL (22.05%), followed by level 2 (17.64%). At levels Below-PP through 1, the number of the
participants amounted to 24, occupying 35.2%. The number of the participants ranging from ORF-IL 2 to 4 was 32, occupying the largest proportion (47%). The smallest number of the participants was above ORF-IL 5, accounting for 17.8%.

The participants’ ORF-FL was reported to be one through three levels higher than their ORF-IL. ORF-FL ranged across 11 levels from level PP to level 9. When the number of the participants at ORF-FL 3 through 5 was combined, it accounted for about half of the entire participants. Level 4 and level 5 accounted for the highest proportion of ORF-FL, occupying 17.65 % and 19.11% respectively. Cumulative percentage within level 5 and within level 8 was 81% and 98% each.

These results drawn from the IRI administration demonstrated that the participants’ ORF level was widely dispersed even though they were all the same grade in the high school. Their independent reading level ranged from level PP to level 7 and the frustration level were from level PP to level 9. Specifically, 37.69% of the participants seemed to be able to read English books of level 2 and level 3 on their own, while only about 25% of the participants seemed able to independently read the books which have higher than level 4 difficulty. On the other hand, there is a higher probability that 50% of the participants would show frustration in reading texts ranging from level 3 to level 5.

The above findings seemed to reveal meaningful and controversial points with regard to high school students’ L2 reading ability, the selection for the reading material and the test passages in Korean educational contexts. As
presented in Table 3.3, the readability indices of the reading passages used for MCSAT were analyzed. The mean Flesch-Kincaid readability index of the reading passages of MCSAT is 8.5, ranging from 1.5 to 14.9. It is only two out of 23 passages (8.87%) that are lower than level 5 readability. In other words, 91.3% of the passages have higher than level 5 difficulty. This means that 80% of the participants are not likely to read 91.3% of the test passages with successful reading comprehension based on the results that they were frustrated with reading the texts of higher than level 5 difficulty. Furthermore, 16 out of the 23 MCSAT-passages (69.5%) have the difficulty level of higher than level 8. It could be inferred that any student participating in this study do not understand these texts without assistance as these texts are beyond their independent reading level. Therefore, it could be concluded that the reading passages of MSCAT may have been too difficult and challenging for the participants to read and comprehend independently.

It is also likely that most of the participants have difficulty in reading their English textbooks. According to Lee (2013) who analyzed the readability indices of the texts included in the current English textbooks (i.e., English I and English II written for high school students), the mean Flesch-Kincaid readability level of expository texts included in these textbooks were 7.06 (1.63) and 7.68 (1.96), respectively. Compared to the participants’ distribution of ORF levels, only two (2.94%) participants seemed to be able to read the textbooks independently, and 94.12% of them are likely to struggle to read the textbook with less than 50% of comprehension. Given the fact that the textbooks are
written for the instructional purposes in the high school, the readability level of the textbook is deemed to be too challenging to read with instruction assistance.

Next, on the basis of the participants’ ORF level, the Spearman correlation analysis was conducted to identify the relationship between the ORF level and the English proficiency level drawn from MCSAT. Table 4.2 illustrated the results of the correlation coefficient analysis. As shown in Table 4.2, it is clear that the two tests were statistically correlated (p < .000).

Table 4.2

<table>
<thead>
<tr>
<th>Variable</th>
<th>ORF-IL</th>
<th>MCSAT Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORF-IL</td>
<td>1.00</td>
<td>-.709**</td>
</tr>
<tr>
<td>MCSAT Level</td>
<td>-.709**</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Significance level: p < .000

The finding regarding the relationship between the ORF level and the English proficiency level of MCSAT appeared to corroborate the results of the previous study (Paris et al., 2002). Paris et al. (2002) reported that there is a significant relationship between the reading comprehension test and scores from the Qualitative Reading Inventory and Development Reading Assessment on word identification and reading comprehension. These correlations ranged from .48 - .90, indicating acceptable concurrent validity. Therefore, the statistical
analysis of the current study implies that ORF-IL can be a reliable indicator of
the students’ reading comprehension ability. Nevertheless, this statistical result
needs to be viewed from a different perspective. Since the two measures tested
the same construct, the correlation should be higher with the \( r \) in the .80 to 1
range (Hatch & Larzaraton, 1991). In fact, this finding is consistent with the
previous studies that have compared two or more reading comprehension tests
(Bowey, 1986; Cutting & Scarborough, 2006; Francis, Fletcher, Catts, &
Tomblin, 2005; Keenan, Betjemann, & Roth, 2005; Nation & Snowling, 1997;
Spear-Swerling, 2004). For example, Cutting and Scarborough (2006) who
compared three reading comprehension tests raise a concern that commonly used
tests of reading comprehension do not necessarily examine the same array of
cognitive processes.

Even though it is not clear what is responsible for the high proportion of
unexplained variances, a possible explanation for this is that the listening skill
was measured in MCSAT. Even though listening comprehension ability is known
to be an important factor in explaining L2 reading comprehension (Gernbacher,
1990; Joshi, William, & Wood, 1998), it is possible that listening comprehension
skill might have served as another variable in the correlation analysis of the
present study. It is also possible that the disparities of the passage levels and the
test item types between the IRI and MCSAT might have an influence on the
students’ reading comprehension processes. With regard to this, the validity
issues regarding the CSAT and the IRI need to be considered in detail.
The previous studies on the CSAT provide insight in this aspect. Heo and Yoon (2004), who investigated the teachers’ and students’ perceptions of the CSAT, questioned the validity of the English section of the CSAT in measuring students’ English proficiency. In the same way, the problem related to the concurrent validity of the CSAT was posed by Lee (2010). Lee (2010), who looked into the college students’ private English-learning experiences, demonstrated that their scores of the CSAT were not consistent with those of TEPS and, accordingly, pointed out that the English section of the CSAT may not capture the students’ English ability.

In fact, the multiple-choice formats which MCSAT is based on have been criticized for “creating an environment that permits students to get an answer instead of actually reading and considering text” (Farrall, 2012, p. 247). Additionally, since this test format is less straightforward to capture reading abilities the test-takers are assumed to have, its cognitive validity seems problematic to be justified (Field, 2011; Weir, 2005; Weir & Khalifa, 2008). On the other hand, as the previous studies confirmed (Fuchs et al., 1988; Jeon, 2012; Jiang et al., 2012; Lems, 2006; Taguchi et al., 2004), IRI measuring ORF have many advantages as a diagnostic tool and its cognitive validity can be argued in that IRI can transparently capture the overall core reading processes ranging from word recognition to reading comprehension which are involved in the real-life reading process (Weir & Khalifa, 2008). Still, IRI should be further examined in terms of reliability and validity in the L2 context, and in particular, the comparison with widely accepted English proficiency tests, such as TEPS,
TOEFL and TOEIC, will provide more well-founded information on the results of the IRI administration.

### 4.1.2 Word recognition accuracy

Based on the data collected from the measurement of ORF, the participants’ WRA was scored. A WRA score for a given passage was calculated by subtracting the percentage of the decoding errors from 100%. WRA was analyzed in terms of whether the scores differed according to the ORF levels. The participants’ ORF-IL was determined by the criteria for WRA (98-100%) and ORC (over 85%) whereas ORF-FL was the level at which the WRA score fell below 90% or ORC was under 50%. Accordingly, at ORF-IL, the participants’ WRA scores were all over 99%. Interestingly, the variances in WRA scores were observed only at the lower ORF-FL. Table 4.3 illustrates the means, the range, and standard deviations of the WRA scores across the ORF-FL.

As shown in Table 4.3, the vast majority of the participants could read the passages almost without committing errors. It means that decoding errors were rarely found in most of the levels except for level PP (90.75%), level P (96.33%) and level 4 (98.91%). Aside from these levels, the mean WRA at ORF-FL was as high as that of ORF-IL.
Table 4.3

The WRA scores of the participants across the ORF-FL

<table>
<thead>
<tr>
<th>ORF-FL</th>
<th>N</th>
<th>Mean</th>
<th>Range</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>PP</td>
<td>7</td>
<td>90.75</td>
<td>16</td>
<td>5.99</td>
</tr>
<tr>
<td>P</td>
<td>5</td>
<td>96.33</td>
<td>9</td>
<td>3.76</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>100</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>99</td>
<td>5</td>
<td>1.65</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>98.91</td>
<td>4</td>
<td>1.44</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>99.53</td>
<td>1</td>
<td>0.51</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>99.75</td>
<td>1</td>
<td>0.57</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>99.5</td>
<td>1</td>
<td>0.57</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The findings seemed to reveal the different developmental profile of L2 word recognition process as echoed by the previous research (Jeon, 2012; Jiang et al, 2012; Lems, 2006; Quirk & Beem, 2012). For example, Jeon (2012) claimed that L2 readers’ word recognition accuracy develops before reading fluency (i.e., reading rate). In the research on ORF of Jeon (2012), the word recognition accuracy scores of Korean EFL high school students were reported to reach the near-ceiling level (89%) in reading passages of an average readability
level of nine (Flesch-Kincaid) even though their average reading speed was 62.5WPM. Similarly, Lems (2006) and Jiang et al. (2012) found that adult L2 readers had good passage reading accuracy, 97% and over 99% respectively.

The findings of the current study seem to lend a support for the claim that there are differences in the word recognition process between L1 and L2 readers. Word recognition is the complex process in which phonological decoding and the semantic access are involved (Koda, 2005). In order for accurate and fluent word recognition to occur, a reader must activate orthographic, phonological, semantic and syntactic knowledge of the word. While skilled L1 readers can activate the links between these components (Perfetti, 2007), this is not the case in L2 reading. That is, L2 readers may be able to match the word form with its sound, but not to make a semantic match to the word’s meaning due to lack of vocabulary knowledge (Lems, 2006). This phenomenon is called ‘word calling’ (Meisinger et al., 2010; Stanovich, 1986).

In the reflection of the word calling phenomenon, the participants of the current study seemed to call out words without extracting the meaning of the word especially at ORF-FL, even though the participants read for comprehension. Therefore, it could be concluded tentatively in the current study that at the ORF-FL, the failure of the meaning extraction of words may cause the participants to comprehend the texts insufficiently in spite of their high WRA scores. A significant finding is that the participants at levels PP and P of ORF-FL appeared to fail in activating the sound information as well as the words’ meaning. It means that these levels of high school students even seem to be struggling to
recognize and pronounce words appropriately, despite almost 10 years of school-based English learning.

Furthermore, the different developmental profiles of the L2 reading acquisition indicates that the criteria established on the data of the reading performances of L1 readers do not seem to be appropriate for determining L2 reading level. According to the criteria, the participants should be able to read the text with 99% accuracy at the independent level, 95% at the instructional level, and when the accuracy rate fell below 90%, they are judged to reach at the frustration level (Betts, 1946; McKenna & Stahl, 2009). However, the data in the present study presented considerably different results from these criteria. These findings indicate the word recognition score cannot be the standard for determining students’ reading level and the criteria established in the L1 setting needs to be examined and validated. In addition, it is suggested that in L2 reading, ORF be assessed in the context where the readers are focused on the meaning of the text, since there could be a significant portion of word callers among L2 readers (Nakamoto et al., 2007; Quirk & Beem, 2012).

4.1.3 The pattern of the decoding error

In order to examine what types of the errors were found in each ORF level, the decoding errors were counted, and then categorized into five types: mispronunciation, substitution, insertion, omission, and examiner-help. For the data analysis, the participants were divided into three groups depending on their
ORF levels: High Fluency (HF), Intermediate Fluency (IF), and Low Fluency (LF). Table 4.4 and Figure 4.2 show the proportion of the decoding errors by the fluency groups and the error types.

### Table 4.4

The frequency and types of the decoding errors

<table>
<thead>
<tr>
<th>Decoding Error</th>
<th>Fluency Groups</th>
<th>LF</th>
<th>Percent (%)</th>
<th>IF</th>
<th>Percent (%)</th>
<th>HF</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mispronunciation</td>
<td></td>
<td>84</td>
<td>71.18</td>
<td>22</td>
<td>59.45</td>
<td>3</td>
<td>42.8</td>
</tr>
<tr>
<td>Substitution</td>
<td></td>
<td>17</td>
<td>8.84</td>
<td>8</td>
<td>21.26</td>
<td>3</td>
<td>42.8</td>
</tr>
<tr>
<td>Insertion</td>
<td></td>
<td>1</td>
<td>0.8</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1.4</td>
</tr>
<tr>
<td>Omission</td>
<td></td>
<td>4</td>
<td>3.38</td>
<td>4</td>
<td>10.8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Examiner-help</td>
<td></td>
<td>12</td>
<td>10.16</td>
<td>3</td>
<td>8.1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>118</td>
<td>100</td>
<td>37</td>
<td>100</td>
<td>7</td>
<td>100</td>
</tr>
</tbody>
</table>

The number of the decoding errors varies according to the fluency groups and error types. Among 162 decoding errors, most of them were found in the LF group and it accounted for almost 72%. In the IF and HF groups, the number of the decoding errors amounted to 37 (22.8%) and 7 (5.9%) respectively.
Figure 4.2
The percentage and the types of the decoding errors across the reading fluency groups

LF Readers
- Mispronunciation
- Substitution
- Insertion
- Omission
- Examiner help

IF Readers

HF Readers

- 65 -
Among the decoding errors detected in the LF readers, mispronunciation (71.18%) marked remarkably high among the five types of the errors, followed by examiner-help (10.16%), substitution (8.84%), omission (3.38%) and insertion (0.8%). The decoding errors found in IF readers follow the similar pattern of those of the LF readers, although the total number of errors was almost a third compared to the LF group. Mispronunciation accounted for 59.45% of the entire errors, substitution for 21.26% and omission for 10.8% and the examiner-help for 8.1%. Meanwhile, the HF readers exhibited almost perfect performances in recognizing words, committing only seven errors when they read their independent and frustration levels of the passages. Among them, three mispronunciation and substitution errors (42.8%) were observed and there was only one insertion error.

Among these findings, the frequency of mispronunciation was noticeable. Especially, this result indicates that the participants belonging to the LF group might not have sufficient phonological awareness needed for word recognition. In other words, lack of phonological knowledge, essential for word recognition, was the main cause of inaccurate word recognition. Furthermore, it was likely that the LF readers’ insufficient phonological knowledge in decoding words caused slow reading speed and low level of reading comprehension. Regarding the decoding skill, Koda (2005) points out that decoding creates a threshold for exploiting the comprehension competence readers bring to their reading acquisition processes. Therefore, it could be concluded that the struggling readers in the LF group may not yet approach an adequate threshold
level of automatic word recognition even when they read their reading level of text and as a result, reading comprehension appeared to be negatively affected. Meanwhile, the participants in the IF group made a relatively small amount of decoding errors compared to the LF group. Still, the high proportion of mispronunciation errors implied that they also might not have developed the attention-saving decoding skill.

Next, among the five types of the decoding errors, the different patterns of the substitution errors were detected depending on the participants’ fluency level. Table 4.5 describes the word list of the substitution errors across the fluency groups.
Table 4.5

Word list for the substitution errors across the fluency groups

<table>
<thead>
<tr>
<th>Error type</th>
<th>LF</th>
<th>IF</th>
<th>HF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Word</td>
<td>Error</td>
<td>Word</td>
</tr>
<tr>
<td>Substitution</td>
<td>follow</td>
<td>flower</td>
<td>clearly</td>
</tr>
<tr>
<td></td>
<td>change</td>
<td>children</td>
<td>by</td>
</tr>
<tr>
<td></td>
<td>barked</td>
<td>broke</td>
<td>never</td>
</tr>
<tr>
<td></td>
<td>small</td>
<td>smell</td>
<td>this</td>
</tr>
<tr>
<td></td>
<td>many</td>
<td>very</td>
<td>cried</td>
</tr>
<tr>
<td></td>
<td>was</td>
<td>saw</td>
<td>be</td>
</tr>
<tr>
<td></td>
<td>we</td>
<td>he</td>
<td>Checkers</td>
</tr>
<tr>
<td></td>
<td>buy</td>
<td>but</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a</td>
<td>the</td>
<td></td>
</tr>
<tr>
<td></td>
<td>me</td>
<td>my (2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>that</td>
<td>what</td>
<td></td>
</tr>
<tr>
<td></td>
<td>her</td>
<td>here</td>
<td></td>
</tr>
<tr>
<td></td>
<td>laugh</td>
<td>another</td>
<td></td>
</tr>
<tr>
<td></td>
<td>late</td>
<td>red</td>
<td></td>
</tr>
<tr>
<td></td>
<td>didn’t</td>
<td>don’t</td>
<td></td>
</tr>
<tr>
<td></td>
<td>to</td>
<td>a</td>
<td></td>
</tr>
</tbody>
</table>

Note. The number in the parenthesis indicates the frequency of the error.

As shown in Table 4.5, the substitution errors of the LF readers showed the noticeable differences from those of the IF readers and the HF readers. To capture the different features of the substitution errors in detail, the specific contexts where some of the errors were found are presented in Table 4.6.
### Table 4.6

The comparison of the original words and the students’ responses in the context across the fluency groups

<table>
<thead>
<tr>
<th>Group</th>
<th>Original Texts</th>
<th>Students’ responses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LF</strong> Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>The dog barked</td>
<td>The dog break.</td>
</tr>
<tr>
<td>(2)</td>
<td>I’ll follow rules.</td>
<td>I’ll flower rules.</td>
</tr>
<tr>
<td>(3)</td>
<td>She will change her mind.</td>
<td>She will children her mind.</td>
</tr>
<tr>
<td>(4)</td>
<td>They are small and can fit well in my small room.</td>
<td>They are smell and can fit well in my smell room.</td>
</tr>
<tr>
<td>(5)</td>
<td>Dad says I can buy some fish.</td>
<td>Dad says I can but some fish.</td>
</tr>
</tbody>
</table>

| **IF** Group | | |
| (6) | Gram was never sick. | Gram was very sick. |
| (7) | We did that for the last six days. | We didn’t that for the last six days. |
| (8) | Not until they are about three weeks old can they see clearly. | Not until they are about three weeks old can they see clear. |
| (9) | By this time they can walk in a tumbling fashion… | But this time they can walk in a tumbling fashion… |

| **HF** Group | | |
| (10) | … soldiers who made forays on chicken coops, piggens, and smokehouses… | …soldiers who made forays on chicken coops, pigeon, and smokehouses… |
| (11) | I couldn’t hear the story but went back to some shopkeepers… | I couldn’t hear the story just went back to some shopkeepers… |

From Table 4.6, it was clear that the majority of the errors which were found in the LF reader group were syntactically and semantically distinct from the expected recognition of words. The substitution errors were graphically
similar with the original words in the text even though a substantial number of
the errors were not syntactically appropriate and caused the meaning change of
the sentence. In particular, when the LF readers who made the errors were asked
questions for comprehension check, they responded to the questions based on the
meaning of the substituting words. Two samples of the answers, which the LF
readers, Students (34) and Student (10), gave to the comprehension questions,
are presented below. The questions and the answers were all given in Korean.

Example (1); Student (34)
Question: “What caused Jane to be sad?”
Answer: “She was sad because the dog’s legs were broken”

Example (4); Student (10)
Question: “What probably caused him to suggest fish for her pet?”
Answer: “I don’t know, she said they are smelly”

From the answers they replied, it can be inferred that they knew the
meaning of the words they pronounced (i.e., break and smell). However, when
they were asked to pronounce the original words in the text (i.e., barked and
small), they had difficulty pronouncing the words correctly. In other words, it
was highly likely that the LF readers replaced the words in the text with the
graphically similar words which they had known, and the substituting words did
not share the common features with the original words in terms of the grammatical category and the meaning.

However, these patterns of the substitution errors were not observed in the word recognition of the IF readers. The substitution errors made by the IF reader were generally graphically similar with the original words and syntactically acceptable in the context. However, some words were not semantically appropriate. For example, in the case of the substitution errors such as very for never in the example (6) and didn’t for did in the example (7), the participants recognized the meanings in the opposite way. Aside from these errors, the substitution errors shown in the samples did not seriously impair the meaning of the sentences. For example, even though to replace clearly with clear in the example (8) was syntactically less acceptable, it did not cause a problem in grasping the overall meaning of the sentence. Similarly, as shown in the example (9), the substitution error (i.e., but for by) did not alter the meaning of sentence even though it is slightly inaccurate in the grammatical sense.

Meanwhile, the HF readers made the fewest substitution errors. Among the two errors, one error was made since the participants replaced pigpen with pigeon, which is described in the example (10). The error in this example was graphically similar to the original word in the text as well as syntactically acceptable in that both words belonged to the same grammatical category. Semantically, the error has the similar qualities in that both words involve the meaning of animal. The other error was observed when the participant substituted but for just as shown in the example (11). This was the only error
which was graphically different from the original word. Even though they differed in the visual display, the substituting word (i.e., *just*) was not syntactically and semantically problematic in the reading context.

When it comes to the substitution errors, these finding appeared to show the different developmental patterns of word recognition according to the L2 readers’ fluency level. First of all, the substitution errors made by the LF reader were distinctive in terms of the amount and the characteristics. In particular, the majority of the errors were deviated grammatically or semantically from the context of the sentences. The meaning of the sentences and the passages were altered by their substitution errors. In addition, there was little consideration concerning the grammatical structures of the sentences which the substituted words belonged to. Given the fact that they had difficulty in pronouncing the original words in the text, it seemed that the insufficient phonological knowledge of the LF readers is attributable to the substitution errors, as well.

In fact, their word recognition patterns shown in the substitution errors were detected among the L1 readers who are still in the developmental stage of word recognition. According to the Phase theory which was developed by Ehri (2005), L1 readers systematically progress through the four stages in order to achieve fluency. Among these four stages, the readers at the *pre-alphabetic stage*, the second stage, tend to focus on the most salient part of a word and consequently use initial and, at later stage, final letters as the clues to a printed word’s pronunciation. The same patterns of word recognition were detected in the LF reader group in that the substantial portion of the substitution errors were
made based on the first letter of the words. In fact, these patterns of the errors suggested that these L2 readers were not able to deal with the full complexity of the sounds in the words and were unable to make complete use of the letter-sound relationships.

On the contrary, the IF readers and the HF readers substituted the words in the text for the words which were syntactically acceptable and semantically interchangeable. Even though it is hard to distinguish which of the information sources were actually involved in the error production, the IF readers and the HF readers seemed to be more aware of the grammatical structure of the sentence and more focused on the meaning construction of the sentences while they were reading a text. In particular, the errors found in the HF reader group did not have much effect on the continuity of the meaning of the passage.

4.1.4 Reading rate

While the participants read aloud the IRI passages, their oral reading rate (ORR) was measured. The ORR score was calculated based on the number of words read per minute (WPM). The ORR of each participant was calculated at their own ORF-IL and ORF-FL and then they were divided into three groups for the data analysis. Among the 68 participants, the seven participants were excluded from ORF-IL as they could not understand the lowest level of the passages in the IRI. Accordingly, at ORF-IL of the LF readers, the ORR data of the 17 participants were used for the data analysis. Since the reading rate was
calculated when the participants read the texts for comprehension, ORR at ORF-IL means the reading rate at which the participants were able to read with more than 90% of reading comprehension. Likewise, ORR at ORF-FL was the reading rate when the participants read with less than 50% of reading comprehension for the texts. The data of the reading rate are presented in Table 4.7.

In general, the participants’ ORF and ORR showed the positive relation. As ORF was higher, the reading rate was faster. At ORF-IL, the ORR score of the HF readers was the highest (130.36 WPM), followed by the IF readers (115.69 WPM). The ORR score of the LF readers was the lowest (104.36 WPM). At ORF-IL, the intergroup differences of the mean ORR were approximately 11 to 15 WPM. The similar pattern was observed at ORF-FL. Thus, the results seemed to indicate that the reading rate is likely to be related to reading comprehension ability. In addition, the participants in each group tended to read slower at ORF-FL than ORF-IL. The difference of ORR between ORF-IL and ORF-FL was approximately 11 WPM, 10 WPM and 12WPM in sequence from the LF readers to the HF readers.
The differences of the reading rate among three reader groups need to be considered in detail. The finding indicates the intergroup difference of ORR at ORF-IL and ORF-FL ranged from 11 WPM to 15 WPM. Concerning the meaning of the reading rate, the research conducted in the L1 setting provides an insight into the interpretation of the data (Carver, 1992; Hasbrouck & Tindal, 2006; Morris et al, 2013, 2017). Carver (1992) who investigated the reading rate of L1 readers found that the reading rate of young L1 readers tends to increase at average 14 WPM per year. As there are discrepancies of the reading acquisition process between L1 readers and L2 readers, it could be inappropriate to compare them one to one. Nevertheless, considering how much print L1 readers are exposed to for a year, the intergroup difference of reading rate in L2 reading

Table 4.7

<table>
<thead>
<tr>
<th>ORF-IL</th>
<th>ORF-FL</th>
</tr>
</thead>
<tbody>
<tr>
<td>LF</td>
<td>IF</td>
</tr>
<tr>
<td>N</td>
<td>17</td>
</tr>
<tr>
<td>Mean</td>
<td>104.36</td>
</tr>
<tr>
<td>SD</td>
<td>16.22</td>
</tr>
<tr>
<td>Min</td>
<td>63.1</td>
</tr>
<tr>
<td>Max</td>
<td>131.16</td>
</tr>
<tr>
<td>Range</td>
<td>68.06</td>
</tr>
</tbody>
</table>
seemed to be significant and it means that L2 readers of the different reading rate appear to have the substantial differences in reading comprehension ability.

However, the difference of reading rate should be interpreted with caution. Notably, there were significant intra-group variances at each ORF-level. Even in the same group, the range of ORR was more than 70 WPM. That is, the reading rate seemed to vary considerably among the participants even within the same group. This means that even though they understood the meaning of the same texts, some participants processed text in a very slow manner while others did at a rather faster rate. To be sure, it is not right to say that a reader whose reading rate is 50 WPM has the same reading ability as the reader of the oral reading rate of 120 WPM even though they can understand the same text. This means that there is a disparity of the reading ability between students within a reading fluency group and thus caution needs to be applied when the data of reading rate is interpreted.

Thus far, the above findings seemed to reveal that the oral reading rate can be a proxy for the reading ability. However, this finding concerning the reading rate does not fully support the conclusion drawn by Jeon (2012) and Lems (2006). These prior studies examined how much variances of L2 reading comprehension is collectively and individually explained by fluency variables. In each study, passage reading fluency explained 26% and 21.1% of silent reading variance, respectively. From these statistical results, they concluded that passage reading rate (i.e., oral reading fluency) cannot be a proxy for silent reading comprehension. The results of the current study do not support this conclusion.
A possible explanation is that Jeon (2012) and Jiang et al. (2012) assessed reading rate and reading comprehension in different stages by employing the independent measurement methods, whereas the current study examined both components simultaneously. This would not make the findings of the present study comparable with those of Jeon (2012) and Jiang et al. (2012), since ORF in these studies was likely to be the rate of ‘word calling’, which is the prevalent phenomenon among L2 readers (Quirk & Beem, 2012). Thus, as Morris et al. (2013) underscores, reading rate can be an efficient and valid measure of fluency “if recorded in a reading-for-meaning context” (p.53). There are also arguments that readers’ performance is impeded not only by the decoding or comprehension factors, but also by their inability to process text in an efficient and fluent manner (LaBerge & Samuel, 1974; Rasinski & Padak, 1998). Therefore, reading rate can be an indicator of L2 reading ability only when reading rate is measured on condition that readers are reading for meaning and understanding.

By the same token, reading rate at ORF-IL and ORF-FL needed to be interpreted in different ways. As mentioned above, the reading rate can be a valid indicator of reading ability in the reading-for-meaning context (Morris et al., 2013, 2017). In the current study, as the reading comprehension score was higher than 90% at ORF-IL, it is legitimate to say that reading rate is a predictor of the different reading abilities at this level. However, the reading rate at ORF-FL needs to be seen in a different way. As shown in Section 4.1.2, a substantial portion of the participants decoded words without accessing to the meaning of the words. Therefore, it is questionable whether their reading rates at ORF-FL
are their true reading rates. The pattern found in this study is consistent with the previous research (Nakamoto et al., 2007; Quirk & Beem, 2012). For example, in the research on word calling, Quirk and Beem (2012) found that the gap exists between reading comprehension and reading fluency (i.e., reading rate) in L2 reading. In their study, approximately 55% of the total participants had significantly higher reading fluency scores than their reading comprehension scores, which suggested that a significant number of word callers exist among L2 readers. From these findings, they concluded that decoding skills develop more rapidly for L2 readers than comprehension skills. Taken together, the L2 readers at ORF-FL would presumably read more slowly than the findings revealed in the current study if they focused more on comprehending the text.

Lastly, these findings need to be considered in the context of reading assessment and the synthesis of the findings is worth being noted. First, it was identified that approximately 98% of the participants’ ORF-FL was within the readability level 8. Meanwhile, the average readability level of the passage of MCSAT was 8.5 and 80% of the passages has over readability level 8. Second, about 50% of the entire participants, whose ORF-FL ranged approximately from level 4 to level 6, tended to read 109 WPM at their ORF-FL, and it means that they read the passage at 109 WPM when they could not properly understand the reading passage. It was also highly possible that they would read more slowly if they focused more on the meaning of the text. The similar result was derived from Jeon (2012) who reported that the Korean high school students read the passages of readability level 9 at 62.5 WPM. Additionally, since the participants
in Jeon’s (2012) were asked to read without understanding meaning of the passages, it is likely that the reading rate would become much slower if they read in the reading-for-meaning setting. Third, the total number of words in MCSAT was reported to be 3707 words. It means that if the time constraint² is considered, the test-takers are required to read the passages at about 82 WPM. Taken all together, it can be concluded that the passages of MCSAT do not seem to be appropriate for assessing Korean students’ English reading comprehension abilities when two factors such as the students’ reading rate and the readability level of the reading passage were considered.

4.2 L2 Reading Experience across the Reading Fluency Groups

In respect to the second research question-what kind of L2 reading experiences the participants in the different fluency groups have, a semi-structured oral interview was conducted. The major differences of L2 reading experiences among the three fluency groups (i.e., LF, IF, and HF) were found in terms of how steadily they were engaged in L2 learning. Specifically, how much L2 reading practice the participants have had is attributable to the L2 reading

² According to KICE, students are given 45 minutes for the reading section of the CSAT
fluency development. The excerpts from the participants’ interviews provide some supporting individual experiences or episodes they told. First of all, the L2 reading experiences of the HF readers are presented.

4.2.1 High fluency group

The HF readers differed substantially to the other two groups in terms of how steadily they had been engaged in learning English mostly focusing on reading. They reported that they have learned English on a continual basis.

Interview Extract 1 (S2)
I started to learn English when I was a first grader. And I have studied English since then. I attended English Hakwon and read quite many English books at that time…Since I became a middle school student, I have studied alone. I like English.

Interview Extract 2 (S8)
I started to learn English when I was seven years old. I learned something like phonics at a place like the community service center. I studied English hard, memorizing words and studying grammar…These days, I don’t study English every day because I have to study math, but I study regularly like three days a week… I think I know enough vocabulary to the extent which it makes comprehension possible.

Interview Extract 3 (S24)
I started to learn English from my mother when I was a kindergarten student. And I attended several English Hakwon. I have studied English steadily since then. I prefer learning English to Korean. Especially, I like learning (English) grammar.

Data from the interviews revealed that the participants had been
predominantly learning English steadily for academic purposes or for their preference. Concerning the first time they were exposed to L2, it seemed that some students started to learn English a little earlier than others. A few students started to learn English when they were kindergarten students and others reported that they were first exposed to L2 learning through receiving the formal education in the elementary school. Even though it was found that the HF readers seemed to be engaged in learning constantly, the three participants who have the highest level of ORF had the distinctive reading experiences. What follows is their interview excerpts and the ORR scores.

**Interview Extract 4 (S2): ORR-142 WPM**

I started to learn English when I was a first grader. And I have studied English since then. I attended English *Hakwon* and read quite many English books at that time. The books were thin and easy ones like *Charlie and the Chocolate Factory*. With these books, I naturally started reading English book. I attended the institute for four years and read quite many books. Even after quitting the institute, I continued to read English books by myself…These days, I don’t have enough time to read, but when I feel bored of reading test practice texts of the CSAT, I read English novels.

**Interview Extract 5 (S7): ORR-153 WPM**

I have experience of studying abroad. I stayed in the U.S. for four years from 4th grade to 7th grade. I received the regular education in the school. For the first few months, I was placed in a kind of ESL class, focusing on English learning. After that, I took the regular classes.

**Interview Extract 6 (S29): ORR-171 WPM**

I attended the English kindergarten and the English elementary school. I think
it helped a lot. In the elementary school, I learned every subject in English… Yes, it was for about 8 years. After that, I have gotten English private tutoring so far.

The reading performances of these three interviewees differed markedly from other participants in terms of the reading rate and comprehension. The distinctive feature about their L2 reading was the relatively abundant amount of L2 reading experience and the kind of reading material. The first interviewee who had a very good reading habit reported that she had read books on a regular basis. The second interviewee was exposed to L2 reading as he received the formal education for four years in the L1 country. While the third interviewee started to study English in Korea, she attended the kindergarten and elementary school, providing all the courses in English. Even though the second and the third interviewees did not report that they enjoy reading books written in English, they seemed to share the experiences of ‘reading to learn’ in common.

Aside from these participants, the vast majority of the participants in the HF group noted that they have read two or three English books so far in the regular classes or to complete the tasks required for the performance tests. The most frequently mentioned reading material they use was English textbooks or test practice workbooks for the CSAT. When asked how they had processed the text and practiced reading with this material, they remarked that they focused on analyzing sentence structures and figuring out the main point of the text. They also reported that they can understand to some extent only if they knew most of the words in the text. The following extract from an interviewee is illustrative of the typical reading experiences the HF readers have.
Interview Extract 7 (S8)

To be honest, I don’t enjoy reading. When I can read smoothly, it feels good. But usually, there are a lot of difficulties… many difficult words in the reading passages. To get a good score in the CSAT, I have to memorize vocabulary, which I do not want to do. So, reading is not interesting. But I am doing it because I have to.

4.2.2 Intermediate fluency group

The responses from the IF group were at variance with the HF group. The amount of L2 reading was relatively small compared to that of the HF group even though the significant differences were not detected in terms of the beginning years of learning English. The comments from the interviews suggested that they had not been constantly engaged in learning English since they started to learn it.

Interview Extract 8 (S1)

I started to learn English very early like when I was three or four years old. My mother showed me some English videos… but I haven’t studied English that much. Actually, I don’t hate to study (English), but I am doing it because I have to. Now, it’s time to study hard.

Interview Extract 9 (S14)

I started to learn English in the elementary school, but I have not been concentrated on learning itself. Since I was a high school student, I have been doing some more… These days, I receive the private English tutoring.
**Interview Extract 10 (S31)**

I like learning English, but I am not good at it. I received the level five at the last mock-CSAT… I like speaking and listening in English. I think I am good at them to some extent… because I attended the English kindergarten. But I can’t read well. I haven’t read that much.

**Interview Extract 11 (S18)**

Actually, I tend to study English as far as it doesn’t make a problem in my grade. When I was young, my father made me memorize the English dictionary… Yes, the dictionary for the young learner. I can’t remember whether I actually like it or not. But anyway, it helped a lot. I didn’t study English hard at the middle school because my English score was quite good enough even though I didn’t study. But, I got the level three at the first mock-CSAT last year. So, I have had the private tutoring since then.

Explicit in the participants’ responses was the fact that the amount of L2 learning differed significantly between the HF readers and the IF readers, which seemed to greatly contribute to the pace of reading fluency development. When it comes to the reasons for the intermittent L2 learning, these interviewees answered that they were not fully engaged in learning L2 or did not enjoy learning it. Concerning their L2 reading experience, how they have done L2 reading practice needed to be noted and is described as follows.

**Interview Extract 12 (S36)**

I have read two English books so far because it was required for the performance assessment… I spend about 10 hours a week studying English at school and the English *Hakwon*. In the class, we learn to read, using the English textbook and the workbook… Our English teacher interprets the text for us and I listen to it.
Interview Extract 13 (S37)

I usually practice English reading with workbooks. In 45 minutes, I try to read all the test text, and find the answers. … It is impossible to read all the text intensively within the time limit, and the text levels are so high. So, I read the first sentence and the last sentence of the text and guess the answer. It works quite well.

Interview Extract 14 (S3)

I like reading novels written in Korean but I have never read English books. My English learning is usually about reading test passages and finding the answers. My weakness in English reading is lack of vocabulary knowledge… I sometimes find the answer intuitively.

Interview Extract 15 (S27)

I have read four English books in the school… thin one like a fairy tale. When I read the textbook, I found it difficult because of vocabulary and grammar.

Obviously, it was revealed that the participants in the IF group had a limited amount of L2 reading experience. The majority of the participants seemed to have problems with lower-level processing due to lack of the vocabulary or syntactic knowledge. Notably, it seemed that they have formed an inappropriate habit of reading text. Some students reported that they do not read the whole text to comprehend it, but guess the content based on few sentences or words. Also, shockingly, some participants appeared to regard the English reading activity as reading the test passages. In addition, it was delivered that some of the IF readers had difficulty in reading and comprehending at the same
time when performing the IRI tasks. The following excerpt illustrates the fact.

*Interview Extract 16 (S16)*

I have never read English books, and reading like this (the IRI administration) is different from what I have usually done. I have never read like this. Reading and comprehending the text at the same time is so daunting. When I read the text, I first read the sentence and then think to find out what it means.

### 4.2.3 Low fluency group

The result of interviews of the LF group revealed that they virtually had no L2 reading practices. In fact, even though they listened to the lecture provided by teachers, they have not been actively engaged in learning L2 itself. Concerning how long they have been engaged in L2 learning, the majority of the LF readers responded as follows.

*Interview Extract 17 (S10)*

I have not studied English until last year. I am trying hard nowadays. I don’t have enough vocabulary knowledge to comprehend the text and I have problem in English grammar.

*Interview Extract 18 (S12)*

I have just started to study English. Actually, I haven’t touched it so far. I used to score 20 points on the English exam. Still it is hard to read. I barely read. I learned to pronounce recently.
Interview Extract 19 (S34)

I don’t like to study English, so I haven’t studied hard so far. I know only a few English words. I usually score 50% in the English tests… I don’t attend English 
Hakwon. I am studying only in the school.

What is notable in the interview extracts is the duration of their L2 learning. They all reported that they have not started to be engaged in learning English until recently even though they have been exposed to it in formal schooling for approximately 9 years. As S22 described his reading ability, the majority of the LF readers seemed to have problems in the decoding skills.

Concerning how they read the L2 texts, the LF readers replied as follows:

Interview Extract 20 (S22)

In the class, we learn English using the English textbook and the workbook. Usually, the English teacher reads the text to us and translates it into Korean. I listen and write it down….I can’t understand the text alone.

Interview Extract 21 (S20)

I am not good at reading in English…When I read the text, I read a sentence and take time thinking about what it means. It takes 10-15 minutes to read a whole passage (test practice passages of the CSAT). In the test situation, I don’t read the whole text and try to find answers.

As illustrated in the interview, the LF readers seemed to have little experience of processing the sentence units as well as struggle with recognizing and pronouncing words. Therefore, they appeared to have difficulty understanding the English texts in their textbooks or workbooks for themselves.
4.2.4 Discussion

The results from the interviews revealed the different reading experiences L2 readers have. Overall, most participants had the limited amount experience of L2 reading, but the quantity of L2 learning appeared to affect the L2 reading fluency development. In particular, as mentioned in the interviews of the HF reader group, the accumulated linguistic knowledge such as grammar and vocabulary seemed to have the significant influence on L2 reading. This is likely to back up the findings that syntactic knowledge (Brimo, Apel, & Fountain, 2017) and vocabulary knowledge (Pulido & Hambrick, 2008) can explain variances in L2 reading comprehension ability and related literacy skills.

While the L2 linguistic knowledge is obviously necessary for reading success, it is not likely to be sufficient itself for the fluent L2 reading skill development. As revealed from the finding of the present study, there were differences among the HF readers in terms of the reading rate and comprehension ability. In fact, the readers who have done ‘the act of reading’ for an extended period of time showed the higher achievement than those who have studied English by translating and analyzing English sentences. This result would support the theories and the previous research regarding the fluency development (Favreau & Segalowitz, 1982; Grabe, 2010; Haynes & Carr, 1990; Koda, 1996; Logan, 1997; Segalowitz & Segalowitz, 1993) What they posit in common is that the fluency development comes as a result of a set of skills (i.e., rapid word recognition, rapid reading rate, accuracy in comprehension) that require the development of automaticity, a large sight vocabulary, and an
extensive period of implicit learning. In the reflection of the fact that fluent reading skills emerge as a result of implicit learning rather than explicit learning of linguistic knowledge (Grabe, 2010), it can be concluded that the text processing experience (Koda, 1996, 2005), namely, the amount of the implicit learning contributed to the different developmental pace of reading ability among L2 readers.

Compared to the HF readers, the IF readers appeared to have a relatively small amount of linguistic knowledge due to discontinuous L2 learning, which might cause the underdevelopment of their L2 reading fluency. In fact, they had difficulty reading and understanding the reading material. In addition, since their reading experience was limited in reading and translating the textbooks or test workbooks, the majority of the IF readers appeared to confuse English reading with text translation, which is backed up by the previous studies (Kim, 2008; Song, 1998; Yang et al, 2009). Furthermore, as illustrated in the results of Section 4.1.1, the levels of the reading material which was revealed to be too high for this reader group seemed to aggravate the situation, and thus, made them rely on the translation rather than simultaneous text processing. More strikingly, some readers addressed that they tended to read only part of the text or a few sentences to comprehend the passages of the workbooks. However, as the automaticity theory (LaBerge & Samuel, 1974) postulates, reading comprehension is made possible with the execution of lower-level processing, and reading fluency comes as a result of efficient and automatic word recognition. Therefore, it is legitimate to suggest that to cultivate fluent reading
skills, the IF readers need to attend to every word and practice reading with extended texts that have the appropriate readability level (Grabe, 2009; Rasinski et al., 2012).

Finally, a number of the participants in the LF reader group were reported to have the scant amount of L2 reading experience. They had difficulties in word identification skills, in particular, due to lack of the phonological knowledge. The participants’ responses to the interviews confirmed the results of Section 4.1.2 and 4.1.3 and are consistent with the previous studies which emphasize the importance of phonological processing skills, especially in the early stage of literacy development (Grabe, 2009; Hulme, Snowling, Caravolas, & Carroll, 2005). From the result of the interviews, it can be inferred that their deficient reading performances resulted from the underdeveloped phonological skills, the key aspect of the word recognition skill.

The description of the differences found in L2 reading experiences between more fluent readers and less fluent readers clarified the fact that the different level of reading fluency might be the outcome of the incremental L2 reading experiences, combined with the amount of linguistic knowledge as a critical factor.
CHAPTER 5
CONCLUSION

This chapter is composed of two sections. Section 5.1 summarizes and presents the major findings of the current study and the implication on English education in the Korean context. In Section 5.2, some limitations of this study are discussed and suggestions are made for further studies.

5.1 Major Findings and Pedagogical Implication

The present study was designed to explore and diagnose English reading fluency among Korean high school students, combined with their L2 reading experiences. This study dealt with two key points in the research questions: 1) the diagnosis of Korean EFL high school students’ ORF using the IRI, 2) the experiential difference involved in the development of L2 reading fluency. The key findings of the current study are summarized as follows.

First, the high school students at eleventh grade in the study showed wide dispersion ranging from 10 to 12 grade levels at ORF-IL and ORF-FL. Notably, about 40 percent of the students were able to read text of the levels 2 and 3 satisfying the performance criteria of the independent level whereas 60 percent of the students showed frustration in reading the passages at the levels 2 through 5. The significant relationship was found in reading performances between ORF-IL and the mock-CSAT level even though the coefficient were
expected to be higher as both tests measured reading comprehension.

Second, there was less variability in word recognition accuracy across the ORF levels at both independent level and frustration level. The students can read texts at both independent level and frustration level with high accuracy. Across all independent levels, reading accuracy means was over 99%, while at the frustration level, reading accuracy means ranged from 99.6 % to 100% except for level P and level PP. The word recognition pattern at the frustration level was consistent with the previous studies and the result implied the ‘word calling’ phenomenon is prevalent among L2 readers.

Third, there were significant differences in the number of the decoding errors across the reading fluency groups and it implied the crucial role of word decoding in the reading fluency development. Mispronunciation accounted for the substantial portion of the decoding errors found in the LF group while the ratio of the substitution errors was on the increase as the students’ reading fluency levels become higher.

Fourth, in respect of reading rate, an increasing pattern was detected according to the students’ fluency levels at both independent level and frustration level. The students in the HF group could read texts at the rate of 130.36 wpm, the IF group at 115.69 wpm, and the LF group at 104.36 wpm. Overall, the students tend to read at the slower rate at the frustration levels than at the independent levels.

Lastly, regarding the L2 reading experiences involved in the development of English reading fluency, the results revealed that there were
significant differences in the total amount of L2 exposure and reading experiences across the reading fluency groups. The readers with higher fluency tended to be engaged in a relatively large amount of reading practices, while readers with lower fluency had a smaller amount of text processing experience.

Based on the major findings aforementioned, the conclusions drawn from the present study propose the following implications for L2 reading education. First of all, the qualities the fluent reader has definitely need to be considered in the L2 educational context if the ultimate goal of the L2 reading instruction is to foster good readers. Fluent readers can recognize words effortlessly and efficiently with minimal attention, so successful reading comprehension is made possible. Even though successful reading comprehension is not founded solely on the efficient word identification processing and other lower level skills, building automaticity in the lower level processing of reading is essential. Even for L2 readers, these efficient lower level skills are the prerequisite for successful reading comprehension. In fact, lack of automaticity in these skills results in slow, laborious reading (Rasinski et al., 2012). With this regard, Nuttall (1996) posits a “vicious circle” to describe readers who cannot develop good reading skills. Slow readers do not read much, and if they do not read much, they do not understand. If they do not understand, they cannot enjoy reading. Therefore, considering that the fluency development involves automatic lower-level processes, which is made possible when the readers “make the best use of what is already known” (Nation, 2009, p.2), it is pivotal to provide a large amount of L2 reading practice and exposure with the appropriate level of texts.
Especially, teachers should provide opportunities for students to read an extensive amount of books on their own with pleasure. In order to find the appropriate level of book for each student, it is also paramount that teachers are well aware of their students’ reading abilities.

In addition, as the fluency development can be supported by a number of instructional practices, it is necessary to develop the reading curricula that focus on building reading fluency. A number of researchers have suggested various fluency instructions which can be incorporated into the regular reading class to assist students experiencing difficulties in reading (Rasinski, 2006, 2010; Rasinki & Padak, 1998). These practices include model fluency reading, assisted reading, repeated reading, and word-recognition exercises. These instructional practices hold promise for gaining comfort with print, thereby enabling the transition from ‘learning to read’ to ‘read to learn’ (Chall, Jacobs, & Baldwin, 1990).

Subsequently, the present study offers a valuable insight into the reading assessment in Korea. Explicit in the results of the current study is that the reading passages of the mock-CSAT are predicted to be too difficult for students to read with the moderate degree of comprehension. As the text difficulty is one of the critical factors contributing to the difficulty level of the reading comprehension test (Freedle & Kostin, 1993, 1999), the test validity and reliability will be critically compromised as the test cannot measure reading comprehension ability properly which it ultimately intends to measure. Furthermore, with the introduction of the criterion-referenced English test into the CSAT in 2018, more empirical research is called for to investigate actual
reading abilities of L2 readers. Regarding the criterion-referenced test, Glaser (1963) stipulated that the achievement measurement should involve “the determination of the characteristics of student performance with respect to specified standard” (p. 519). That is to say, to ensure whether students attain the criterion-reading ability, the target performance of reading and stage along the route to that performance should be described. Consequently, only when it is possible, the validity of the newly-introduced English test of the CSAT can be argued and justified. In this sense, IRI is worthy of notice as a reading assessment tool. As it is individually administered and responsive to students’ performance, it can produce valid results and provide more reliable information about L2 reading growth. Therefore, IRI deserves to be accentuated in a Korean educational context as for it can provide valuable information on multiple aspects of students’ reading abilities.

5.2 Limitations and Suggestions for Future Research

The present study has some limitations that need to be considered concerning the methodology in spite of the fact that each experiment was meticulously designed. First of all, the sample size was relatively small due to the test format requiring the individualized test administration; the findings of the study cannot be generalized to other groups of students. In order to ascertain
better reliability of the results from the statistical analyses, it might have been better to have more participants take the IRI Test. Since the participants were limited to eleventh-grade high school students, the findings of the current study cannot be generalized to different grades of students. If research using the IRI similar to the present study is conducted on learners at different grades, a more integrated picture about the reading fluency development of Korean EFL students will be obtainable.

In addition, even though this study provides valuable information concerning the reading rate of the L2 reader, caution must be applied. This study had limitations in measuring the reading rate at their frustration level. If other measurements are employed such as eye-trackers, it is possible to reach a more precise understanding of the patterns of L2 learners’ reading skill development. In addition, with a small sample size, the findings might not be generalizable. The results regarding reading rate represent only a start and need to be confirmed or challenged in future studies of this type.

Lastly, the present study employed the text from the published informal reading inventory, which examined the readability of the text based on the readability formula and verified it by the process of field testing. However, the specific levels of the texts which were assumed to be appropriate for L1 readers might not be applied to L2 readers. Besides, as Klare (2000) claimed, the conventional readability formulas are limited in predicting characteristics that contribute to the text difficulty, other than semantic difficulty and syntactic complexity. The factors, such as the organization of the text, topic familiarity, the
coherence of the text and cultural knowledge should be more carefully considered for designing fluency testing measures, especially for the L2 readers.
REFERENCES


Miller, J., & Schwanenflugel, P. J. (2008). A longitudinal study of the development of reading prosody as a dimension of oral reading fluency in


- 112 -


Rasinski, T. V., & Samuels, S. J. (2011). Reading fluency: What it is and what it is not. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say*
about reading instruction (pp. 94-114). Newark, DE: International Reading Association.


APPENDICIES

APPENDIX A. Recruitment document for participants ......................... 121
APPENDIX B. A sample of test Form A for students (Level 3) ............... 122
APPENDIX C. A sample of test Form A for teachers (Level 3) ............. 123
APPENDIX D. Interview Questions .............................................. 124
APPENDIX A. Recruitment Document for Participants

This recruitment document was approved by IRB on June 2017.
APPENDIX B. A Sample of Test Form A for Students (Level 3)

Read this story to find out about what two youngsters decide to do.

It’s summer, it’s hot, and it’s Monday.
“What do you want to do?” asks Slam.
“I do not know,” says Dunk. “It’s too hot for a game of basketball.”
“Do you want to play checkers?” asks Slam.
“No,” says Dunk. “We did that for the last six days. Why not go swimming?”
That’s when Slam and Dunk go down to the kitchen. Grandpa is there. He has a big grin. He’s got something in his hands, but Slam and Dunk cannot see what it is.
“Do not look,” says Grandpa. “See if you can think of what I have in my hands.”
“A sandwich!” says Slam.
“A cactus!” says Dunk.
“No,” says Grandpa. “Look! This is what we will do for fun today!”
“Go fishing!” yell Slam and Dunk. “What a good idea!”

---

4 The actual test set of L2 reading fluency tasks in the administration of the current study was produced by combining reading passages from Informal Reading Inventory (Burn & Roe, 2011) which consisted of 14 passages in total. The directions for the test passages were provided in Korean.
APPENDIX C. A Sample of Test Form A for Teachers\(^5\) (Level 3)

INTRODUCTORY STATEMENT: Read this story to find out about what two young men decide to do.

It's summer, it's hot, and it's Monday.

"What do you want to do?" asks Slam.

"I do not know," says Dunk. "It's too hot for a game of basketball."

"Do you want to play checkers?" asks Slam.

"No," says Dunk. "We did that for the last six days. Why not go swimming?"

That's when Slam and Dunk go down to the kitchen. Grandpa is there. He has a big grin. He's got something in his hands, but Slam and Dunk cannot see what it is.

"Do not look," says Grandpa. "See if you can think of what I have in my hands."

"A sandwich!" says Slam.

"A crutch!" says Dunk.

"No," says Grandpa. "Look! This is what we will do for fun today!"

"Go fishing!" yells Slam and Dunk. "What a good idea!"

Source: Go to Hawaii, by Chris Sawyer (Gateway Learning Corporation, 2001), pp. 9-13.

---

**COMPREHENSION QUESTIONS**

<table>
<thead>
<tr>
<th>Type</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>main idea</td>
<td>1. What is this story about? (Slam and Dunk decide what to do.)</td>
</tr>
<tr>
<td>detail</td>
<td>2. In what season does the story take place? (summer)</td>
</tr>
<tr>
<td>cause and effect</td>
<td>3. What caused Dunk not to want to play basketball? (It was too hot.)</td>
</tr>
<tr>
<td>inference</td>
<td>4. What have they done for the last six days? (play checkers)</td>
</tr>
<tr>
<td>inference</td>
<td>5. What part of the house were Slam and Dunk in at the beginning of the story? (upstairs)</td>
</tr>
<tr>
<td>sequence</td>
<td>6. What room do Slam and Dunk go to after Dunk suggests going swimming? (the kitchen)</td>
</tr>
<tr>
<td>inference</td>
<td>7. How does Grandpa feel when Slam and Dunk see hin? (pleased: happy)</td>
</tr>
<tr>
<td>vocabulary</td>
<td>8. What is a cactus? (a plant that usually has stickers, prickles, or spines and grows in dry areas, such as deserts.) [Accept reasonable approximations of this definition]</td>
</tr>
<tr>
<td>inference</td>
<td>9. Which guess about what Grandpa had in his hands was best? Why do you say so? (the sandwich, because a cactus might hurt his hands) [Accept any reasonable rationale for the choice made.]</td>
</tr>
<tr>
<td>inference</td>
<td>10. What did Grandpa have in his hands? [Accept a fishing pole, a fishing rod, or fishing gear of any kind.]</td>
</tr>
</tbody>
</table>

---

\(^5\) The test forms for teachers are from Informal Reading Inventory (Burn & Roe, 2011), which consisted of 14 tasks in total. The directions and the questions for the test were all provided in Korean. The original version of the reading comprehension questions was used for the test administration except for the vocabulary questions. The vocabulary questions were excluded from the test administration according to Taguchi et al. (2004).
APPENDIX D. Interview Questions

Learner Background

- 1학년 때 언어성적과 영어 성적은 어떠하십니까?
- 국어 책은 얼마나(어떤 종류) 읽습니까?
- 국어책을 읽는 것은 좋아합니까?
- 영어 학습을 시작한 나이는 몇 세 입니까? (영어를 배운 기간은 얼마나입니까?)
- 영어는 듣기, 말하기, 읽기, 쓰기 별로 어느 정도 입니까?
- 공인 인증 영어 점수를 보유하고 있습니까? 어떤 시험? 몇 점?
- 영어권 국가에서 여학연수를 받거나 학교에 다닌 적이 있습니까? ( 언제? 몇 년 동안 입니까?)
- 영어에서 가장 쉽거나 필요한 영역은 무엇입니까?
- 영어에서 어려운 영역은 무엇입니까?
- 영어 읽기에서 가장 어려운 점은 무엇입니까?

Reading in School

- 학교 수업에서 주로 배우는 영어 영역은 무엇입니까?
- 학교 영어 수업이 영어 실력 향상에 도움이 된다고 생각합니까?
- 학교 수업은 읽기 실력 향상에 도움이 된다고 생각합니까?
- 학교 수업시간에 영어 교과서는 어느 정도 사용합니까?
- 교과서 외 다른 학습 자료는 사용합니까?

Private Tutoring

- 영어 사교육은 받은 적이 있습니까? (과외/학원/학습지/EBS)
- 사교육이 영어 실력에 향상을 줄니까?
- 사교육에서 배운 것은 주로 무슨 영역입니까?

Reading Outside School

- 영어 TV 프로그램이나 영화를 보는데 사용하는 시간은 1주일에 얼마나 됩니까?
- 영어 라디오나 이야기는 듣는데 사용하는 시간은 1주일에 얼마나 됩니까?
- 영어책, 신문, 잡지, 만화, 인터넷 등을 읽는 시간은 1주일에 얼마나 됩니까? 얼마나 자주?
- 학교 수업 외에 영어를 공부하는 시간은?
- 집에 영어책을 몇 권이나 보유하고 있습니까?
- 부모님은 어려분이 어렸을 때 얼마나 자주 영어책을 읽어주셨습니까?
- 부모님은 어려분이 읽은 영어책에 관심이 있습니까?
- 가정에서 구독하는 영어신문이나 잡지가 있습니까?

- 124 -
국 문 초 록

본 연구는 한국인 고등학교 영어 학습자를 대상으로 제 2 언어로서의 영어 읽기 유창도가 능력이 어느 정도이며, 다양한 리터러시 관련 경험 속에서 어떠한 과정을 거쳐 발달하게 되는 지를 살펴보고자 한다.

읽기의 유창도란, 하위수준의 읽기 기술과 상위 수준의 읽기 기술을 모두 아우르는 전반적인 읽기 능력이라고 정의할 수 있다. 즉, 읽기의 이해는 단어를 인식하는 독자의 유창성에 달려 있으며, 이는 글을 읽으면서 단어를 인지하는 속도와 정확성, 그리고 표현력으로 나타난다. 그러나 그간 외국어로서의 영어 학습자가 어느 정도의 유창도를 가지고 글을 읽을 수 있는 지에 대한 연구가 많이 시행되지 않았기에, 한국어 리터러시 학습 배경을 가진 영어 학습자들의 읽기 유창도 발달 과정을 살펴보고자 한다.

본 연구의 참여자인 58명의 고등학교 학습자는 비공식 읽기 평가도구를 이용한 읽기 유창도 테스트 과업을 수행하였다. 읽기 과업에서 참여자는 영어로 모국어로 사용하는 학습자의 학년수준에 맞게 만들어진 일련의 텍스트를 소리 내어 읽은 후, 읽기 이해 질문에 답하였다. 읽기 과제가 끝난 후, 학습자들은 인터뷰에 응하였다. 읽기 과제의 분석 결과, 학습자들의 읽기 유창도는 같은 학년 내에서 넓게 분포하고 있었다. 제 1 언어 사용자 기준으로, 유치원 이하의 단계에서 7학년 단계까지 이르러 있었으며, 독립적으로 읽을 수 있는 수준과 학생들이 좌절을 느끼는 수준으로 나누어 본 결과, 참여한 학생의 40%는 2학년이나 3학년 수준의 글을 스스로 읽고 이해할 수 있는 것으로 나타났다. 하지만 참여학생의 50% 학생들은 3학년에서 5학년 수준 사이의 글을 50% 이하 정도 밖에 이해할 수
없었다. 이러한 학생의 분포는 학생들의 모의고사 성적과 유의미한 상관관계를 보였으나, 두 시험 모두 읽기 이해를 측정했음에도 불구하고 높은 상관관계를 가지지는 않았다. 읽기 속도는 학생들의 유창도가 높아짐에 따라 증가하는 추세를 보였고 읽기 능력과 읽기 속도가 정적 관계에 있다는 것이 확인되었다. 반면 단어 인식의 정확도는 학생들의 읽기 능력별로 큰 차이를 드러내지 않았다. 한편, 유창도를 기준으로 그룹별로 나누어 본 결과, 학생들의 읽기 오류는 읽기 유창도가 낮을수록 많은 오류가 발견되었다. 또한 읽기 유창도가 낮을수록 잘못된 발음으로 인한 오류가 많은 비율을 차지하는 반면, 유창도가 높아질수록 다른 단어로 대체하여 발음하는 오류가 점점 많은 비율을 차지하였다. 마지막으로, 읽기 유창도 발달에 관련된 제2언어 읽기 경험에 관한 학습자 인터뷰는 제2 언어 읽기 활동의 양 면에서 학습자간 큰 차이가 있었음을 보여주었다. 결과에 근거하여, 본 연구는 학습자의 읽기 유창도 발달을 위한 읽기 수업에 대한 시사점과 이후 연구에 대한 제언을 결론부에 제시한다.

주요어: 제2언어 읽기 유창도 발달, 읽기 속도, 단어인식의 정확성, 단어인식 오류, 제2 언어 읽기 경험, 비공식 읽기 평가 도구

학번: 2016-21792