The Relationship of Korean Students’ Age and Years of English-as-a-Foreign-Language (EFL) Exposure with English-Reading Ability: A Cross-Age Study

Jill Fitzgerald

MetaMetrics and The University of North Carolina at Chapel Hill Emerita
and Adjunct Research Professor, USA

A. Jackson Stenner

MetaMetrics and The University of North Carolina at Chapel Hill
Adjunct Research Professor, USA

Eleanor E. Sanford-Moore, MetaMetrics, USA

Heather Koons

MetaMetrics and The University of North Carolina at Chapel Hill
Clinical Assistant Professor, USA

Kimberly Bowen

MetaMetrics, USA

Kee Hyung Kim

Natmal Co., Ltd, Lexical Data-Processing Institute, Korea

July 9, 2013

This is an accepted (not copy-edited) manuscript of an article published in Reading Psychology as:


The copy-edited article is also available online:


Running Head: Korean English Reading
Abstract

The purpose of the present cross-age study with South Korean students was to investigate the relationship of age and years of English-as-a-foreign-language (EFL) exposure with English-reading ability. South Korean third-grade through university students who were studying EFL reported English exposure and completed a standardized English reading test which resulted in a reading ability score locating each individual’s EFL-reading ability on a continuous, developmental scale. A regression analysis revealed that age and English exposure interacted in relation to students’ reading ability. Description of the cross-age English reading levels was among the first to portray EFL-reading ability at multiple ages.
The Relationship of Korean Students’ Age and Years of English-as-a-Foreign-Language (EFL) Exposure with English-Reading Ability: A Cross-Age Study

The purpose of the present cross-age study with South Korean students was to investigate the relationship of age and years of English-as-a-foreign-language (EFL) exposure with English-reading ability. The main research question was, “Do individuals’ age and number of years of English exposure interact in relation to English-reading ability?” The study also revealed the cross-age comparative status of South Korean students’ EFL reading ability.

Rationale

In today’s world where English is the global language for communication in commerce, science, publishing, and education, EFL reading proficiency is equally as important as oral proficiency (cf. Crystal, 2003). English-as-a-foreign-language reading exams are used in some countries to screen applicants for jobs, and high scores on English-reading tests constitute common entrance requirements around the world, not just in countries where English is the dominant language (cf. Crystal, 2003). Rich description of the interplay of age, EFL exposure, and EFL proficiencies could contribute to the advancement of students’ EFL proficiencies, including EFL reading proficiencies because such description could inform classroom instruction and policy (Hakuta, Bialystok, & Wiley, 2003; Muñoz, 2007). For instance, decisions about whether to introduce formal EFL instruction, including formal English-reading instruction, in school at early ages rest on the assumption that “earlier is better” for long-range EFL proficiency. Similarly, such research could assist decisions about whether different instructional methods might be suggested at different ages (Cameron, 2001; Muñoz, 2007).

While a modicum of evidence on the subject of age, amount of new-language exposure, and EFL second-language proficiency has accumulated for oral language, far less is known
about the relationships with regard to EFL reading ability. To our knowledge, there are no studies investigating either rate of learning to read specifically in EFL or the relationship between the onset of learning EFL and EFL reading ability attainment. Nor are there cross-sectional studies that aim to examine the interplay of age and language exposure in relation to EFL reading ability, or even studies that simply aim to describe the second-language reading abilities of students at different ages with differing degrees of exposure to the second language.

With an eye toward the possibility that the literature on EFL oral language might assist hypothesis development for our study of EFL reading, in the following sections results of relevant research are provided about the relationship between age, length of exposure to the new language, and EFL oral proficiency. As well, a theoretical stance on the relationships for EFL reading is formulated.

**Age, Exposure, and EFL Oral Proficiency**

The impact of age and amount of new-language exposure on oral language proficiencies has been a significant topic in the field of Second Language Acquisition (SLA) for decades (Muñoz, 2008). In general, the focus of investigations has been divided between an emphasis on the rate of second-language acquisition on the one hand and the impact of initial age of second-language acquisition on eventual second-language oral proficiency on the other.

With regard to rate of acquisition, while it is commonly assumed that children learn a second language faster than adults, there is virtually no evidence to support the supposition (DeKeyser, 2013). In fact, in most studies, including studies accomplished in foreign-language-learning situations, younger children have been shown to be slower in attaining second-language oral proficiencies than older children and adolescents (García-Mayo & García Lecumberri, 2003; Krashen, Long, & Scarcella, 1979; Muñoz, 2006, 2008).
At the same time, research accomplished almost exclusively in immersion settings that are typical for immigrants has consistently revealed a negative correlation between initial age of learning the second language and ultimate oral proficiency attainment (Abrahamsson, 2012; Krashen et al. 1979). More than two dozen studies now support such a conclusion, with most studies employing oral measures of morphosyntax and pronunciation (Birdsong, 2006). The negative correlation between age of acquisition and ultimate attainment of oral proficiency has generally been found to be linear, with correlations ranging from -.45 to -.77, with a median of about -.64 (Birdsong, 2006).

However, it is not clear that the results on the relationship between age of acquisition and ultimate oral proficiency extend to the foreign-language learning setting. Foreign-language learning studies since the 1970’s have yielded conflicting results and have not demonstrated the advantages of an early start (Muñoz, 2008). In particular in EFL situations, when younger starters have been shown to have a long-term oral proficiency advantage over later starters, it generally occurs only when the younger students have also had a greater amount of exposure as compared to later starters (Muñoz, 2008). Moreover, even in the latter case, the advantage is not impressive. The failure of extension to the EFL situation is not entirely surprising because learning a second language as an immigrant when immersed in the dominant language is plausibly different in several ways from learning to read in designated foreign-language school classes where one’s own native-language is the dominant, every-day language (Bernhardt, 2011). Motivations, extent of daily encounter with the foreign language, extent of required and optional reading in the second language that one accomplishes, familial language, and method of instruction, among other related factors, are situational contexts that potentially impact one’s English-reading ability trajectory (Muñoz, 2008).
A Marriage of Theories

Still, some basis for hypothesizing about the relationships when EFL reading is involved would be fruitful for the present investigation and for future ones as well. It is possible that theories that underpin SLA relationships among age, exposure, and oral proficiency might also support hypotheses about the relationships for SLA reading. As well, a marriage of aspects of oral language theory with EFL reading theory might provide additional support.

Theoretical underpinnings, oral SLA/EFL. In the SLA and EFL-specific literature on the relationships among age, exposure, and oral language proficiency, two dominant theoretical themes appear. The first is proposed as support for the ubiquitous finding that older students tend to achieve higher levels of proficiency faster than young students. The effect is generally interpreted as a consequence of older students’ superior cognitive skills (Muñoz, 2008). With age, typically-developing students acquire a vast array of declarative and procedural knowledge about their native language and about learning in general, knowledge which they can call on to scaffold and support their new-language oral development.

Second, the Critical Period Hypothesis (CPH) popularized by Lenneberg (1967) over 30 years ago continues to be frequently evoked and strongly debated as an explanation for eventual varying oral language proficiency by age and exposure to the second language. The basic tenet of the CPH is that there is an ideal period of time (up to approximately 12 to 15 years of age) during which language acquisition is easier, and once the ideal period is passed, language acquisition becomes more effortful, with reduced likelihood of ultimately achieving native-like proficiency (cf. DeKeyser, 2013).

A corollary to the CPH is that different types of cognitive processing happen at different ages (DeKeyser, Alfi-Shabtay, & Ravid, 2010). In brief, one explanation is that younger children...
rely more on using their memories to learn a second language, whereas older children and adults rely more on rule-based learning (Nikolov & Djigunovic, 2006; Paradis, 2004). Younger children are thought to automatically acquire morphosyntactic and phonetic-phonological systems from mere exposure (cf. Lenneberg, 1967; Bley-Vroman, 1990) by using innate domain-specific mechanisms (Bley-Vroman, 1989; DeKeyser, 2000) and procedural memory resources (Paradis, 2009). Declining memory forces later second-language learners to consciously analyze language and be more explicit as they learn, using mostly declarative memory resources (Abrahamsson, 2012; Paradis, 2004). A second account is that the biological plasticity of procedural memory for language begins to decrease after approximately age 5, and therefore, cognitive reliance on conscious declarative memory increases, with a most dramatic shift at approximately age 7 (Paradis, 2004).

However, considerable controversy exists over the validity of CPH for EFL situations. For both immigration and EFL settings, results of investigations using large-scale data sets do not support a critical period of peak sensitivity for advancing individuals’ second-language oracy (Birdsong, 2006; Hakuta et al., 2003; Chiswick, Lee, & Miller, 2004; Nikolov & Djigunovic, 2006). Rather than finding discontinuities before and/or after certain periods, such as before or after ages 12 to 15, several sets of researchers found that the degree of eventual success, typically defined as native-like oral proficiency, steadily declined through the life span (e.g., Bialystok & Hakuta, 1994; Johnson & Newport, 1991).

**Theoretical underpinnings, EFL reading.** A key concept in one of the most widely cited research-based theoretical perspectives on second-language reading is that individuals rely on multiple information sources: a) level of native-language literacy; b) new-language knowledge, generally implying knowledge of phonology, morphology, and grammar; and c) a cluster of
remaining factors including comprehension strategies, level of reader engagement, content and domain knowledge, interest, and motivation (Bernhardt, 2011).

Support for the first information source comes from research in which native-language word-reading ability was highly related to second-language word reading ability, from very young to older adolescent students (e.g., Abu-Rabia, 1997; Durgunoğlu, Nagy, & Hancin-Bhatt, 1993) and across typically-developing students and struggling readers (e.g., Abu-Rabia, 1997; Da Fontoura & Siegel, 1995). However, the phonological and visual mental processes underlying word reading may be strongly influenced by the degree of orthographic similarity between the two languages (e.g., Abu-Rabia, 1997; Da Fontoura & Siegel, 1995). Although to our knowledge no research provides evidence of the degree to which native-Korean-reading ability transfers to English-reading ability, at least some degree of transfer might be anticipated because Korean and English share similar phonological processes in that both are alphabetic languages. However, English is less regular (that is more “opaque”) in symbol-to-sound relationship than Korean (which is more “transparent”) (Koda, 2008; Perfetti & Dunlap, 2008).

The second factor, new-language knowledge, has been supported as critical in that selected measures of new-language oral proficiency, such as phonological awareness, have tended to correlate highly with new-language word-reading ability for second-language learners (e.g., Durgunoğlu, et al., 1993). Second-language listening comprehension and oral vocabulary meaning have also been positively related to second-language reading comprehension (e.g., Verhoeven, 2000). Turning to the third and final cluster of information sources, strategic abilities, such as specific comprehension strategies, developed in native language were related to use of strategic abilities in new-language reading (e.g., Schoonen, Hulstijn, & Bossers, 1998).
Importantly, prior research supports the contention that the three sources of information gradually shift in importance as individuals become more and more proficient SLA readers. Native-language literacy understandings ultimately reach a point where they explain approximately 20% of the variance in readers’ new-language comprehension; new-language knowledge, primarily syntactic knowledge, explains approximately 30%; and 50% of the unexplained variance is attributed to the third cluster of remaining factors (Bernhardt, 2011).

Marrying theories together for EFL reading. Theoretically, taken together, the preceding EFL research and theory suggests older students might acquire EFL reading faster than younger ones in part because they have acquired more native language oracy and literacy (Bernhardt, 2011; Muñoz, 2008). Older students should have more native-language cognitive declarative and procedural knowledge to bring to bear in English. As well, for EFL reading, at a given age, amount of English exposure likely broadens and deepens individuals’ understandings about phonological, syntactic, and structural properties of English (cf. Bernhardt, 2011). Accompanying increasing exposure to English is a gradual shifting of importance of native language literacy, new language knowledge, and other language factors (Bernhardt, 2011) for EFL reading. That gradual shifting in conjunction with conflicting findings for the CPH in EFL settings, some of which emphasize amount of exposure over peak age periods, suggests that the CPH may not be pertinent for EFL reading. Instead, a cross-sectional linearized pattern of EFL reading growth by age and EFL exposure might be theorized—versus a discontinuous pattern.

The Present Study Hypotheses and Design

To date, little is known about individuals’ actual EFL reading levels or whether their reading levels are associated with their age and/or exposure to EFL. The present study was a initial step accomplished in one country toward exploring and describing the relationship
between age, exposure, and EFL reading ability status across grades. Many variables such as student motivation, native-language reading, and native- and English-oral ability need to be investigated. As well, to address issues such as whether there is a CPH for EFL reading, studies in which adults with varying degrees of English reading exposure are tested for English reading proficiency need to be accomplished. We did not attempt to address a broad array of factors. Instead, our study is a modest effort toward an initial exploration and description of the EFL reading ability cross-age pattern in relation to age and EFL exposure in one country, employing a limited set of variables.

Given our main research question, “Do individuals’ age and number of years of English exposure interact in relation to English-reading ability,” the prior research and theory on oral second-language proficiency and prior second-language reading theory led to selected hypotheses for the present study. The main hypothesis was that there would be a significant interaction, expecting that amount of English exposure would moderate the effect of student age and vice versa. First, there might be a trend toward older students with lower levels of English exposure performing similarly to younger students with higher levels of English exposure. We hypothesized such an interaction because prior research supported the contention that older students tend to learn oral EFL faster than younger ones. As well, from second-language reading theory, we thought that younger individuals might not have sufficiently developed native-language reading to transfer to EFL reading, whereas older students would have a richer and higher-level native-reading ability and therefore more opportunity for knowledge transfer from native-language to the English-reading task, which could lead to more accelerated initial learning. Second, within an age group, individuals who had been exposed to EFL for a longer period of time might have gained greater oral English proficiency in addition to having
accumulated greater amounts of English-reading practice, both of which could lead to higher-level reading ability, as compared to their same-age peers with less exposure to English. Third, from the failure of oral proficiency research on CPH in EFL situations to produce a peak sensitive period, and from related research results which have pointed to primary importance of exposure over a sensitive peak age period, a generalized disordinal linear cross-sectional trend might be predicted. That is, when examining students within age group, incremental EFL reading ability increases by exposure would vary across age groups, but with no distinctive differential differences in pattern between older versus younger students.

The study was a cross-sectional one in which 3,529 South Korean third-grade through university individuals who were studying EFL took an English-reading test that was specifically designed for South Korean students studying EFL, and students also reported number of years of English exposure. A regression analysis was conducted to explore the possible interaction of age and number of years of English exposure as related to EFL-reading ability.

**Methods**

**Participants**

Participants were 3,529 South Korean individuals between the ages of 10 and 49 in 3 elementary schools, 1 middle school, 1 high school, and 2 universities. There were 653 students in fifth and sixth grades (342 males, 311 females; ages 10 to 13), 898 (446 males, 452 females; ages 12 to 15) in seventh and eighth grade, 932 (248 male, 684 female; ages 14 to 17) in ninth and tenth grades, 800 in one school that included eleventh and twelfth grade (all female; ages 16 to 19), and 246 (125 male, 121 female; ages 18 to 49) from two universities. The schools and universities were located in three large cities in South Korea with populations of over 1,000,000 people.
At the time of testing, students self-reported the amount of time they spent reading in English outside of school. Using an intraclass correlation method (Ginns & Barrie, 2004), the reliability of the estimates of age group mean reported time spent in English reading was .94. Interestingly, as shown in Table 1, as age increased, there was a tendency for students, on average, to read in English for fewer hours, with the youngest students reporting that they read, on average, 2.35 hours per week, declining to .81 hours per week for 18 to 19 year olds and 1.91 hours per week for those over 19. There was considerable variability in time spent reading English outside of school at each age group, with the median of one hour at each age group except the 18-19-year-old group, where the median was zero hours per week.

**The South Korean Situation**

Information about how EFL and EFL reading are taught in South Korea may be useful context for interpreting results of the present study. Learning EFL is highly valued in South Korea and has been included in the national curriculum at the higher grade levels at least since 1955. More recently, the national curriculum has incorporated EFL beginning at the third grade level. However, instantiation of EFL at the third grade and other elementary grade levels has been slowly progressing, and many schools do not yet provide EFL at the elementary level. Earlier guidelines suggested that by 2010 one native-English-speaking teacher was placed in every middle school, but data are not available to indicate the degree to which the guideline was implemented. In the national curriculum, instructional EFL emphasis in the early grades is on conversation, while in the middle and high school grades, there is more emphasis on learning grammar (Chang, 2008). However, some empirical reports indicate that comprehension and/or conversational ability remain the focus even through the university level (Busbee, 2001; Joh & Choi, 2011).
In the mid-90’s a communicative language teaching method was introduced into the curriculum, a method which included a focus on teaching EFL reading by provision of pre-reading activities, encouragement to skim and scan, and inclusion of discussion or summarization of what was read (Joh & Choi, 2001). However, a survey of 97 middle-school English teachers in South Korea revealed that over 75% of the teachers were not following the new curriculum. Instead, they relied heavily on teaching reading in a “traditional” manner, by introducing new words in the text to be read, expecting line-by-line translation, sometimes having students listen to recorded reading materials, explaining syntactic structures, and asking comprehension questions at the end of reading (Joh & Choi, 2001).

Materials

Number of years of English exposure. At the time of testing, students self-reported, and, in non-university grades, teachers affirmed, the number of prior years a student had studied English. Because some students indicated years of study before formal instruction in schooling likely began, we have interpreted the variable as “years having been exposed to English.” It is well known in Korea that English tutoring occurs outside of the educational system, and as well, some students in our sample may have had parents who spoke English at home and encouraged or taught their children English. Using an intraclass correlation method (Ginns & Barrie, 2004), the reliability of the estimate of age group mean reported years of English exposure was .99.

The E-LQ English Reading Comprehension Assessment. The E-LQ English Reading Comprehension Assessment (Natmal Corporation, 2011) measures an individual’s English-reading comprehension level. The test is specifically for Korean students studying EFL. The test is intended to: measure a student’s English-reading ability beginning at the fifth/sixth-grade level as well as growth across time.
The test contains 36 to 40 items (increasing by grade level) with either one or three task/item types, depending on student grade level: At the fifth/sixth-grade level, three item types are included: a) Item-type 1: Students read one-sentence with a blank substituted for one word and four word-completion choices. An example is: “Can Sal ______ the cat?” with the four choices, pet, ban, pit, dot. b) Item-type 2: Students read two sentences with a blank substituted for one word in the second sentence and four word-completion choices. An example is: “The dog could not catch the cat. The cat was very ______” with the four choices, fast, big, long, funny. c) Item-type 3: Students read a short paragraph 50 to 125 words long and read one sentence calling on an intersentential relationship with a blank substituted for one word with four word-completion choices. At the rest of the grade levels, only item-type 3 was included. Fiction (25% of a test) and non-fiction (75% of a test) content areas include the natural sciences, arts and humanities, social science, and literature. The test is untimed, but designed to be completed within 40 to 50 minutes.

The resulting score is the Lexile, a standardized equal-interval unit of measure set on a continuous, developmental scale that may range from below 0 to over 2000. A Lexile is 1/1000th of the difference between the mean difficulty of mid-first-grade material and the mean difficulty of college and workplace passages (Burdick, Swartz, Fitzgerald, Stenner, Burdick, & Hanlon, 2011).

Test reliability (Chronbach’s Alpha) ranged from .75 to .88 across forms. Content validity was addressed by inclusion of fiction and non-fiction content areas. Also, passage difficulties were aligned to the difficulty of texts that students enrolled in the various English language course levels are likely to encounter. One text per grade commonly used in third-through twelfth-grade South Korean English-language courses were obtained from the Natmal
Corpoartion in Seoul, Korea. The texts were used to guide the target levels and ranges for the *E-LQ English Reading Comprehension Assessment* test passages. Lexile measures were obtained for each of the texts, and item vocabulary difficulty (measured in Lexiles) for the *E-LQ* was matched to Lexile difficulty of passage. Construct validity evidence was provided through a set of linking studies for the Lexile’s predictive ability for 14 other standardized assessments of English reading, including the *Stanford Diagnostic Reading Test* (Karlson & Gardner, 1995), the *Gates-MacGinitie Reading Test* (MacGinitie, MacGinitie, Maria, Dreyer, & Hughes, 2006) and the *Iowa Tests of Basic Skills* (Hoover, Dunbar, & Frisbie, 2005), with sample sizes ranging from 956 to 16,363 students across second through post-secondary grade levels. Correlations ranged from .60 to .93, with a median of .88 (Stenner, Burdick, Sanford, & Burdick, 2007; Stenner, Horabin, Smith, & Smith, 1988).

**Procedures**

All students who were present on the day of the testing took the *E-LQ English Reading Comprehension Assessment*. Testing was completed across the school sites within a two-month span. Reading directions from a script, teachers administered the untimed test to whole-class groups of students. Students were told that the test is a reading comprehension test, three practice items were completed, and directions were repeated if requested by the students. Students completed the test within 50 minutes.

**Results**

**Preliminary Analyses**

As can be seen in Table 1, and as would be expected, the average number of years of English exposure increased with grades, from 3.09 at the 10- to 11-year-old age level to 10.50 for 20 years and older. Similarly, the variance for number of years increased with grade, from a
standard deviation of 1.63 years at the lowest level to 3.86 at the highest age level. It does appear that a few students at each age group had been learning English at home, and/or had been tutored in English, before entering school. For example, 2 students in the 10-11 group had been learning English for 9 years, 1 in the 12-13 group had been learning English for 13 years, 2 in the 14-15 group had been learning English for 15 years, 2 in the 16-17 age group had been learning English for 17 years, 1 in the 18-19 group had been learning English for 19 years, and 1 in the oldest group had been learning English for 35 years.

As can be seen in Table 2, within each age group there was a wide range of number of years of exposure to English. The range of English exposure witnessed within age groups in Table 2 is possible for the following reasons: a) The national curriculum only fairly recently included an EFL course in the elementary grades (starting at grade three), and inclusion is not widespread nationally. At the time of data collection, some, or even many, older students in our study may not have been exposed to EFL course work beginning in elementary school. b) Some schools may have offered EFL in the elementary grades at any point in time, even before the national curriculum included it beginning at grade three. c) Some children may have immigrated to Korea in later school years and therefore not have had the EFL courses at earlier grades. d) Some children may have had English tutoring prior to experiencing formal EFL courses in school. e) Some may have come from families where English was spoken (and/or read) in the home.

Within each age group, on the whole, as number of years of exposure to English increased, the mean and median Lexile increased. Also, holding number of years of English exposure constant, and moving across age groups, again on the whole, the mean and median Lexile increased. Notably, within years of English exposure by age group, there was a wide
range of reading ability. For instance, for the youngest group who had been exposed to English for one to two years, the standard deviation was 187.13 and the range was -468 to 378 Lexiles. For the oldest group, for those who had been exposed to English for 9 to 12 years, the standard deviation was 163.04 and the range was 704 to 1258 Lexiles.

The zero-order correlation of years of English exposure with Lexile measure was moderately strong and significant, $r_{xy} = .54, N = 3,193, p < .01$, affirming the cross-age pattern shown in Table 1 for a generally positive relationship between years of English exposure and demonstrated English reading ability. Normally, we might think that the correlation between age and English exposure would be extremely high—because older students could have more/longer opportunity for English exposure than younger ones. However, the correlation of .54 left 69.75% of the variance in the relationship unexplained. Surprisingly, the zero-order correlation of number of hours spent reading English weekly outside of school with Lexile score revealed there was no relationship, $r_{xy} = -.02, N = 3,188, p < .39$. There was only a very weak tendency for individuals who had been exposed to English longer to spend more time reading English outside of school, $r_{xy} = .09, N = 3,138, p < .01$.

**Regression Analysis Procedures**

Analysis was accomplished through a linear regression model with reading ability measured by the Lexile as the dependent variable; test-level taken by the participant was a control variable for test difficulty; and predictor variables were participant’s age, number of years of English exposure, and the interaction of age and number of years of English exposure. Prior to reporting the results of the regression, treatment of missing data and outliers is presented, discussion of tests of assumptions is provided, and assessment of collinearity is discussed.
**Missing data.** Regarding treatment of missing data, 336 students had missing data for number of years of English exposure. A comparison of age distributions for students who did and did not have data for years of English exposure revealed the two groups were similar in age ($M = 15.13$ and $14.51$, $SD = 3.22$ and $3.05$, respectively). The Lexile measure means, medians, and SDs for the two groups, respectively, were similar: a) for the former group—444.99, 449.00, and 316.72; b) for the latter group—349.60, 330.00, and 322.98. For a scale with over 2000 units, a mean/median difference of approximately 100 is relatively small. In the regression analysis, missing data were deleted in list-wise fashion.

**Assumptions.** Statisticians differ on the extent to which tests of assumptions for linear regression are required. One group believes that some assumptions for linear regression should be tested. Among the most commonly stated assumptions are linearity between predictors and outcomes, homoscedasticity of the relationship of errors with the outcome variable, and normality of the distribution of errors in the population (e.g., Draper & Smith, 1998). Other statisticians provide evidence that linear regression is robust to violations of the assumptions. For instance, Cohen and Cohen (1983) state that when regression is used for descriptive purposes no tests of assumptions are necessary. The latter group contends that even in the case of inferential statistics when one intends to refer to the population, only one assumption—homoscedasticity—is applicable (Cohen & Cohen, 1983). Our stance is consistent with the latter of the two groups. However, to err on the side of caution, and for readers who adhere to the importance for tests of assumptions, the three most commonly cited assumptions were tested.

We tested for linearity by examining the scatterplots for each independent variable as a function of Lexile. The relationship between Lexile and participant age was predominately linear with very slight flattening in the extreme upper range. The correlation of age with Lexile was .73.
(\(p < .01\)). The relationship between Lexile and number of years of English exposure was also
linea, with a correlation of .54 (\(p < .01\)). Homoscedasticity was checked by examining the
standardized residuals as a function of standardized predicted value. Except for minor lower and
upper outliers, variance was similar at points along the x-axis. We assessed normality of the
residuals in the population by examining Q-Q plots of the observed values with expected normal
values. Excepting very slight non-normality in the extreme lowest ranges for the predictor
variables, the residuals conformed to the assumption of normal distribution in the population.

**Outliers.** Inspection for outliers was accomplished by examining frequencies for each of
the variables. There were no clear outliers for any variables.

**Collinearity.** As would be expected, the correlation of age with years of English
exposure was moderate (.54). Older students have more opportunity for English exposure. The
VIF for age was 8.12, and for years of English exposure it was 1.60, both under the commonly
used VIF criteria for collinearity of 10. However, given the correlation between the two
predictors, the regression coefficients may be slightly diminished, but as will be reported in the
next section, the coefficients were in the expected positive direction and were reasonably strong.

**Was there an Interaction of Age and Number of Years of English Exposure?**

After entering the test-level as control, using the enter method for remaining variables, a
significant model emerged (\(F_{7,3185} = 918.43, < .01\)). The adjusted \(R^2\) was .67. That is, 67% of the
variance in student Lexile measure was accounted for by the predictors.

EFL reading ability was impacted by an interaction of participant’s age in conjunction
with number of years of English exposure, \(t_{1,3185} = -5.44, p < .01\). (See complete source of
variance information in Table 3.) This was a slightly disordinal interaction that indicated that
neither age nor years of English exposure separately would sufficiently explain reading ability.
level. For every year increase in age there was a small decrement in English reading ability increase (-.91). Figure 1 shows the Lexile medians for the slightly disordinal interaction. Notable is the pattern of gradually increasing reading ability as age increased, but also clear is the impact of years of English exposure. That is, performance was impacted by the combination of age and years of English exposure taken together. As can be seen in the figure, the slightly disordinal nature of the interaction was largely attributable to those exposed to English for 6 to 8 years. Although on the whole, Lexile performance was higher as age rose, and within each age group, students exposed to English for a longer period tended to perform at a higher reading ability level as compared to those exposed to English for shorter periods. For instance, for 12 to 13 year olds, students with 6 to 8 years of exposure (Lexile median = 257; n = 160) outperformed those with 1 to 2 years of exposure (Lexile median = 131; n = 153) by, on average, 126 Lexiles. Similarly, at ages 18 to 19, students with 9 years or more of exposure (Lexile median = 813; n = 160) tended to outperform those with 3 to 5 years of exposure (Lexile median = 727; n = 46) by 86 Lexiles. We provide these exemplative comparisons, as well as ones that follow, solely for the purpose of exploring and explaining the cross-age patterns and the interaction effect.

However, the advantage for longer English exposure tended to peak at about 9 to 12 years. That is, as can been seen in Figure 1, for example, for those who were 16 years old or older, there was little difference in performance when compared to those exposed 9 to 12 years and 13 years or more (the purple and blue lines are practically synonymous).

The interaction effect is also seen in Figure 2 that shows the interquartile Lexile ranges by student age and number of years of English exposure. The top of the top whisker is the highest score any student achieved, the top of the box is the 75th percentile, the middle bar is the 50th percentile, the bottom of the box is the 25th percentile, and the bottom of the lower whisker
is the minimum score any student achieved. Several points are noteworthy. First, it is clear that a student’s relative standing can not be judged simply by locating a student’s score either according to age or to number of years of English exposure separately. As an example, some older students did not outperform younger students by much—if they had fewer years of English exposure than did younger students. For instance, the median Lexile for 14 to 15 year olds who were exposed to English 1 or 2 years (282L; n = 126) was only slightly higher than the median Lexile for 12 to 13 year olds who had considerably more exposure—6 to 8 years of English (257L; n = 160). Similarly, 18 to 19 year olds with only 3 to 5 years of English exposure performed on the whole (Mdn = 727L; n = 46) only slightly higher than somewhat younger 16 to 17 year olds who considerably more exposure—9 to 12 years (Mdn = 659L; n = 258).

Second, an inference from Table 2 and from the set of examples in the previous paragraph is that, on the whole, older students may tend to catch onto to EFL reading initially more quickly than do younger students. To illustrate, for one to two years of English exposure, 10 to 11 year olds’ median reading ability Lexile was -29 (n = 109), for 12 to 13 year olds it was 131 (n = 153), and for 14 to 15 it was 282 (n = 126). (The comparable figures continue the pattern for 16 to 17 year olds [Mdn Lexile = 498] and 18 to 19 year olds [Mdn Lexile = 678], though due to small cell sizes [19 and 12, respectively] we suggest caution in interpretation at the highest age levels.)

Third, for patterns for longer English exposure within age group, Figure 2 shows a pronounced different set of patterns for younger versus older students. For children 10 to 11 years old, there was a pronounced upper shift in English reading ability, on the whole, after English exposure for more than five years. On the other hand, for students in the other age groups, there tended to be a more gradually inclining advantage for longer English exposure.
We note that both age and years of English exposure main effects were also statistically significant, \( t_{1,3185} = 7.79, p < .01 \) and \( t_{1,3185} = 13.94, p < .01 \), respectively. However, the significant disordinal interaction effect overrode simple explanation of the main effects.

**Conclusions and Discussion**

The main conclusion was that South Korean student English-reading ability was impacted by an interaction of age and number of years exposed to English. That is, our hypothesis was supported—amount of English exposure moderated the effect of student age and vice versa. Specifically, we hypothesized that there could be a tendency for older students with lower levels of English exposure to perform similarly to younger students with higher levels of English exposure, and there was evidence to support that. The cross-age pattern demonstrated a general upward shift in EFL reader ability with increasing age, but in some instances, older students with lower levels of English exposure performed similarly to younger students with more English exposure. As well, within age group, there was a tendency for students who had been exposed to English for a longer period of time to have higher-level reading ability, but that tendency was moderated by a possible ceiling effect for years of English exposure on median EFL-reader ability level. For older students, more than 9 to 12 years of English exposure did not appear to raise median EFL reader ability levels.

**Discussion**

While it may be intuitive to some to think that age and number of years of English exposure interact to impact EFL reading, to our knowledge, there is no prior evidence to support the contention for EFL reading, nor is there a description EFL reading levels across ages in any country such as the one we present. The cross-age description of South Korean students’ EFL
reading ability is among the first, if not the first, description of its kind. It is noteworthy for several additional reasons.

First, it appears that a typically-developing native-English-speaker developmental English-reading trajectory is not immediately applicable to the EFL learner. On average, young native-English-speaking beginning readers take some significant period of time, up to the end of second or third grade, to fully develop initial knowledge about phonemic awareness, sound-symbol relationships, and word recognition strategies, after which vocabulary knowledge and comprehension strategies begin to explode (Chall, 1996; Clay, 1993). Then acceleration rate tends to remain steady through late middle grades with some decline in acceleration during high school years (cf. Chall, 1996; Fitzgerald & Shanahan, 2000). Clearly, while what is known about native-English reading development may have some import for describing the EFL reading developmental pattern, a more complicated set of patterns emerges for the latter situation where the cross-age results in the present study imply reading ability status is dependent on the combination of age and number of years of English exposure. To address the question, “How is a particular EFL student doing in English reading,” comparing a student’s English-reading status to reading growth curves that are documented for native-English speakers is less meaningful than comparing the student’s status to other students similar in age and with similar amounts of EFL exposure.

Second, in a few recent U. S. studies, English-language learners’ English-reading growth patterns suggest that young English-language learners catch up to monolingual peers rapidly, after which they maintain the same reading levels and the same growth patterns as their peers (e.g., Kieffer, 2011). The present cross-sectional study does not directly reveal a developmental growth pattern, but the interactional nature of age and years of exposure to EFL in the present
A cross-sectional study does suggest that if future longitudinal studies of EFL-reading growth in Korea were accomplished, the developmental pattern may differ from the ELL-reading growth pattern.

Third, on the whole, holding amount of English exposure constant, as was the case in prior research for oral EFL proficiency, older students tended to perform at higher EFL reading levels than their younger counterparts. One explanation for the result is that younger students did not yet have sufficiently developed native-language reading understandings to make as rapid initial progress as older students who could bring to bear, through transfer mechanisms, a more developed, richer and higher-level native-reading ability.

Fourth, the possibility of a sensitive or peak period of 9 to 12 years of English exposure for EFL reading is interesting, but certainly requires more study. The present study cannot sufficiently address the issue because the sample does not include a large number of adults who have reached “ultimate” attainment for EFL reading. However, the finding of the possibility of a peak period of 9 to 12 years of English exposure, raises the question of whether there might be a critical period for English exposure for EFL reading ability, rather than a critical age of onset for ultimate EFL attainment.

Fifth, for South Korean individuals who are college- or university-bound in native-English-speaking countries, such as Great Britain, Australia, or the United States, on the whole, there may be some distance between their peak English-reading ability levels and the English-reading levels of textbooks used in the latter countries. At least in the United States, there is some indication that the typical text-complexity level used in universities ranges from about 1137 Lexiles (5th percentile of university texts measures) to 1583 Lexiles (95th percentile) with a median of 1395 Lexiles (Williamson, 2008; cf. Koons, 2011 for a similar study with texts that
overlapped with ones in the Williamson study). The new Common Core State Standards in the United States set 1355 Lexiles as the target median text complexity level for United States students to be able to read (Common Core State Standards, 2010). Even many eighteen- to nineteen-year olds in the present study who had been exposed to English for 13 or more years fell short of matching such a standard (their median Lexile was 834, and the range was 464 to 1255). One researcher found a similar discrepancy between South Korean college students’ reading ability levels and the English college texts they needed to read, reporting that most students had difficulty with word recognition and translating parts of the texts (Song, 2001).

**Limitations**

Results of the present study should be considered in light of study limitations. First, only South Korean students were studied. Findings from South Korean may not generalize to other countries where English is learned as a foreign language. Instructional factors, which vary from country to country and one national curriculum to another, likely impact students’ English-reading ability growth. Second, although the number of years the students had been exposed to English was known in the present study, the nature of the English exposure was not known. Some students may have had private tutoring, some may have had dual-language parents who began to teach or practice English in the home, and/or some may have had parents who motivated students to learn English more than other parents did. The nature of English learning outside of school, including the extent to which it was spoken at home could have influenced students’ rate of learning to read in English. Third, some might question the self-report for number of years of English exposure. However, at least by the intraclass reliability method assessed in the present study, the measure was quite reliable. Additionally, others have found high levels of reliability for similar self-report measures using other indicators of reliability. For
instance, Guglielmi (2008) estimated reliability for self-reports of second-language learners’
native-language oral proficiency through a multi-trait, multi-method analysis and found strong
evidence for the reliability, stability, and construct validity for the self-reports. Fourth, the extent
to which teachers actually taught the national curriculum was not known in the present study.
Finally, inclusion of native-language reading ability, native-language oral proficiency, and/or
English-oral proficiency in the analyses would have permitted broader exploration of effects,
including a wider array of interactions among variables.

Educational Implications

One significant educational implication is that figures such as shown in Figure 2 could be
used for comparative purposes, that is to compare EFL students’ status and progress in relation to
other EFL same-aged and similarly English-exposed students’ status and progress.

The difference between the median South Korean 18-year-old or older students and the
typical United States university or college-text levels suggested that many college/university-
bound South Korean students might expect to spend extra time reading the English texts and/or
expect to seek additional English instructional support.

Future Research

The present study represented only a basic first step toward better understanding of how
EFL reading ability develops over time. If future researchers could conduct similar explorations
but include native-reading ability and/or second-language oral ability measures, much more
could be learned about which factors moderate EFL reading ability growth. As well, examining
trajectories in relation to classroom English-reading instruction would better inform the
conditions under which trajectories occur. It would be important to learn whether varying
emphases in EFL instruction in general might result in different patterns of EFL reading
progress. For instance, would a curriculum that involved considerable emphasis on conversational English and English composing result in different growth patterns as compared to one heavily focused on grammar learning? Also, studying individuals’ trajectories over long periods of time, for instance, as was done in the United States with the National Educational Longitudinal Studies (Curtin, Ingels, Wu, & Huer, 2002), could provide much more information about situational contexts that impact English-reading growth as compared to cross-sectional studies such as the present one. Cross-country comparisons could help us to better understand the degree to which EFL reading growth is bound to cultural factors. For all such studies, use of a developmental measure of reading, one set on a continuous developmental scale, would be important for measuring reading-ability growth. Finally, because the CPH is so frequently debated in the SLA literature with regard to oral language proficiency, it would be interesting to explore the CPH with regard to EFL reading. With a sample of adults who learned EFL, measures of EFL reading proficiency could be taken as indicators of ultimate attainment. Data about age of onset of learning or length of exposure to English could be collected. The adults’ age of onset of learning English could be split at the commonly assumed CPH age of about 12 or 15 years old, length of English exposure could be held constant, and the researcher could examine whether adults who began EFL reading before the CPH of age 12 or 15 had a proficiency advantage over those who began later.

A major policy implication arises from the need for such future studies—is there political will to provide sufficient funding for such large-scale studies? For instance, could PISA begin to include English-reading tests in countries where English is taught as a foreign language? Do national governments find sufficient reason to explore English-reading trajectories? Large-scale
assessments are complicated, time-consuming, and costly, and the proclivity to conduct them will likely be informed by the extent of worldwide English usage.
References


### Years English Exposure and Hours Reading in English Outside of School by Age Group

<table>
<thead>
<tr>
<th>Age Group</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>Mode</th>
<th>SD</th>
<th>Range</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>Mode</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-11</td>
<td>282</td>
<td>3.09</td>
<td>3</td>
<td>2</td>
<td>1.63</td>
<td>0-9</td>
<td>278</td>
<td>2.35</td>
<td>1</td>
<td>1</td>
<td>2.97</td>
</tr>
<tr>
<td>12-13</td>
<td>715</td>
<td>4.07</td>
<td>4</td>
<td>3</td>
<td>2.03</td>
<td>0-13</td>
<td>699</td>
<td>1.93</td>
<td>1</td>
<td>1</td>
<td>2.23</td>
</tr>
<tr>
<td>14-15</td>
<td>829</td>
<td>4.73</td>
<td>5</td>
<td>3</td>
<td>2.75</td>
<td>0-15</td>
<td>851</td>
<td>1.44</td>
<td>1</td>
<td>0</td>
<td>1.87</td>
</tr>
<tr>
<td>16-17</td>
<td>849</td>
<td>7.08</td>
<td>7</td>
<td>9</td>
<td>2.64</td>
<td>1-17</td>
<td>841</td>
<td>1.39</td>
<td>1</td>
<td>0</td>
<td>1.81</td>
</tr>
<tr>
<td>18-19</td>
<td>416</td>
<td>7.19</td>
<td>8</td>
<td>10</td>
<td>3.86</td>
<td>0-19</td>
<td>419</td>
<td>0.81</td>
<td>0</td>
<td>0</td>
<td>1.86</td>
</tr>
<tr>
<td>20+</td>
<td>102</td>
<td>10.50</td>
<td>10</td>
<td>10</td>
<td>6.21</td>
<td>0.5-35</td>
<td>100</td>
<td>1.91</td>
<td>1</td>
<td>0</td>
<td>2.93</td>
</tr>
<tr>
<td>Overall</td>
<td>3193</td>
<td>5.57</td>
<td>5</td>
<td>3</td>
<td>3.33</td>
<td>0-35</td>
<td>3,188</td>
<td>1.55</td>
<td>1</td>
<td>0</td>
<td>2.13</td>
</tr>
</tbody>
</table>

Note. 336 participants, some from each age group, had missing data for years of English exposure, and 341 participants, some from each age group, had missing data for time spent reading in English outside of school.
Table 2

*Lexile Mean, Median, Mode by Age and Years English Exposure*

<table>
<thead>
<tr>
<th>Yr. Eng.</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>Mode</th>
<th>SD</th>
<th>Range</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>Mode</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-5</td>
<td>152</td>
<td>109.70</td>
<td>102</td>
<td>49</td>
<td>222.51</td>
<td>-468/579</td>
<td>375</td>
<td>172.09</td>
<td>174</td>
<td>124</td>
<td>216.78</td>
<td>-449/827</td>
</tr>
<tr>
<td>6-8</td>
<td>14</td>
<td>386.50</td>
<td>349</td>
<td>715</td>
<td>270.32</td>
<td>23/827</td>
<td>160</td>
<td>297.53</td>
<td>257</td>
<td>98</td>
<td>254.60</td>
<td>-409/1184</td>
</tr>
<tr>
<td>9-12</td>
<td>2</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>11</td>
<td>244.09</td>
<td>246</td>
<td>98</td>
<td>321.44</td>
<td>-561/566</td>
</tr>
<tr>
<td>13+</td>
<td>0</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>2</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

| 14-15    |    |       |        |      |       |           | 16-17 |       |        |      |       |           |
| 1-2      | 126| 261.47| 282    | 98   | 164.89| -106/659 | 19  | 543.16| 498    | 495  | 153.10| 272/838  |
| 3-5      | 361| 375.62| 369    | 70   | 192.19| -154/1055| 199 | 582.04| 568    | 471  | 169.56| 234/1255|
| 6-8      | 234| 465.65| 471    | 447  | 206.78| -264/1170| 332 | 589.16| 580    | 422  | 173.84| 191/1199|
| 9-12     | 62 | 525.45| 484    | 422  | 246.22| 9/1175   | 258 | 670.01| 659    | 639  | 162.84| 272/1255|
| 13+      | 4  | --    | --     | --   | --    | --        | 24  | 717.50| 670    | 494  | 216.75| 308/1224|

Table 2 cont’d
Table 2 cont’d

<table>
<thead>
<tr>
<th>Yr.Eng</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>Mode</th>
<th>SD</th>
<th>Range</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>Mode</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-2</td>
<td>12</td>
<td>680.83</td>
<td>657</td>
<td>617</td>
<td>173.54</td>
<td>396/949</td>
<td>5</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>3-5</td>
<td>46</td>
<td>744.96</td>
<td>727</td>
<td>748</td>
<td>179.50</td>
<td>396/1105</td>
<td>17</td>
<td>878.47</td>
<td>849</td>
<td>682</td>
<td>191.81</td>
<td>549/1299</td>
</tr>
<tr>
<td>6-8</td>
<td>105</td>
<td>751.00</td>
<td>733</td>
<td>547</td>
<td>178.51</td>
<td>355/1221</td>
<td>15</td>
<td>944.00</td>
<td>969</td>
<td>809</td>
<td>118.69</td>
<td>789/1189</td>
</tr>
<tr>
<td>9-12</td>
<td>160</td>
<td>789.89</td>
<td>813</td>
<td>918</td>
<td>156.23</td>
<td>396-1206</td>
<td>30</td>
<td>990.23</td>
<td>990</td>
<td>789</td>
<td>163.04</td>
<td>704/1258</td>
</tr>
<tr>
<td>13+</td>
<td>41</td>
<td>818.85</td>
<td>834</td>
<td>639</td>
<td>181.74</td>
<td>464/1255</td>
<td>11</td>
<td>956.45</td>
<td>969</td>
<td>869</td>
<td>122.22</td>
<td>704/1131</td>
</tr>
</tbody>
</table>

Note. 490 participants had missing data for either years exposed to English and/or Lexile score. Also, descriptive statistics are not shown in the table for cells with fewer than 11 participants.
Table 3

*Regression Sources of Variance Results*

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>B Unstandardized</th>
<th>Est.SE</th>
<th>F, p&lt; (df)</th>
<th>R, Adjusted R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model: Control, Age, Years English Exposure, Age by English Exposure</td>
<td>182.50</td>
<td>918.43, .01 (7, 3185)</td>
<td>0.82,0.67</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>22.21</td>
<td>2.86</td>
<td>7.78, .01 (1,3185)</td>
<td></td>
</tr>
<tr>
<td>English Exposure</td>
<td>17.63</td>
<td>1.23</td>
<td>14.37, .01 (1, 3185)</td>
<td></td>
</tr>
<tr>
<td>Age by English Exposure</td>
<td>-0.91</td>
<td>0.17</td>
<td>-5.44, .01(1, 3185)</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1

*Medians for Age by English-Exposure Interaction*

Note. Cells with fewer than 11 participants were omitted.
Figure 2

Interquartile Lexile Ranges by Age and Number of Years English Exposure

Age Groups (e.g., 10 to 11) with Number Years English Exposure (e.g., 1 to 2)

Figure 2 cont’d.
Figure 2 Cont’d.

Age Groups (e.g., 10 to 11) with Number Years English Exposure (e.g., 1 to 2)