English Language & Literature Teaching, Vol. 26, No. 3 Fall 2020
https://doi.org/10.35828/etak.2020.26.3.101

An Analysis of the Continuity of the Listening Scripts of Middle School English Textbooks with Coh-Metrix

Jisu Ryu[*]
(Konkuk University)
Moongee Jeon[**]
(Konkuk University)

Ryu, Jisu & Jeon, Moongee (2020). An analysis of the continuity of the listening scripts of middle school English textbooks with Coh-Metrix. *English Language & Literature Teaching, 26(*3), 101-121.

This study aims to analyze the linguistic continuity of the listening passages of middle school English textbooks revised by the 2015 National Curriculum with Coh-Metrix, a computer system designed to analyze various psycholinguistic features of text corpora. Specifically, a total of 310 listening scripts were constructed to examine the continuity among three grades for 23 Coh-Metrix measures. The measures selected for this study included basic counts (the number of words, the number of sentences, word length, sentence length), word frequency, word features (imageability, concreteness, age of acquisition), lexical diversity (type-token ratio), personal pronouns, connectives (causal connectives, additive connectives, temporal connectives), readability scores (Flesch Reading Ease score, Flesch-Kincaid Grade Level), syntactic complexity, referential cohesion, and semantic cohesion. The results showed that the continuity was appropriately established for the number of words and sentences, sentence length, additive connectives, words before main verbs, Flesch-Kincaid Grade Level, and word features. It, however, was not reflected for word length, word frequency, lexical diversity, noun phrase density, semantic and reference cohesion, causal and temporal connectives, pronouns, and Flesch Reading Ease score. The findings provide several pedagogical implications for the educational policy of National Curriculum and the development of English language teaching materials.

[Coh-Metrix/middle school English textbooks/listening scripts/corpus/continuity]

---

[*] The first author
[**] The corresponding author

I. INTRODUCTION

With the advent of internationalization period, the interest in English education as well as its status as an international language has been increasing. As part of efforts to enhance national power in line with this trend, it is a highly significant and timely goal to ensure that English education is systematically designed and implemented. This is particularly the case in EFL settings where the amount of exposure and the use of English in everyday life is limited; in these settings, English textbooks used in public school serve as the primary input in English education for teachers and learners (Jeon, 2011). For this reason, studies on English textbooks have been actively conducted (Bae, 2019; Go & Park, 2007; Jeon, 2011, 2015; Kim, J. & Jeon, 2016; Kim, S. & Jeon, 2013; Son & Lee, 2005)

Previous studies on English textbooks have undergone a major shift in methodology. Early studies were primarily conducted based on the manual work in data collection and analysis, and qualitative research was mainly carried out (Kim, 2001; Son & Lee, 2005; Song & Yoon, 2003). These studies were limited in accuracy, objectivity, and the amount of data to be analyzed, and they focused on several specific measures rather than a wide range of measures (Park, 2007). With the help of advanced technologies made in the fields of computational linguistics (Jurafsky & Martin, 2008), corpus linguistics (Meyer, 2002), and several text analysis tools (e.g., Wordsmith Tools and Antconc), however, many textbook researchers in English education have been able to analyze a considerable amount of data in a more objective and accurate fashion (Lee & Lee, 2019; Yu & Kim, 2015).

Studies using these text analysis tools complemented the limitations of the early manual-based research. Nevertheless, these analysis tools that had been widely used in previous studies were predominantly limited to the word or sentence level which made it difficult to examine various linguistic and psycholinguistic aspects of textbooks. In order to promote a more comprehensive understanding of the materials used in English textbooks, it is necessary to take a look at deeper-level measures (e.g., cohesion-related features) and surface measures (e.g., word or sentence-based features). Noting the significance of the comprehensive text analysis, a group of researchers in the Institute for Intelligent Systems (IIS) at the Memphis University have developed an automatic text analysis tool, Coh-Metrix (Graesser, Jeon, Cai, & McNamara, 2008; Graesser, Jeon, Yan, & Cai, 2007). Coh-Metrix provides more than 100 indices on the Internet to measure text features extensively.

Recently, researchers employed Coh-Metrix to investigate various linguistic and psychological features of English textbooks (Bae, 2019; Im & Jeon, 2013; Jeon, 2011, 2014, 2015; Jeon & Im, 2009, 2010; Kim & Jeon, 2013; Lee & Hyun, 2013; Song, 2013).

They investigated how the text difficulty across grade levels systematically differed in terms of diverse measures (Bae, 2019; Kim & Jeon, 2013) and how linguistic and psycholinguistic features were comparable across textbooks published by different publishers within a particular school year (Im & Jeon, 2013; Jeon & Im, 2010). The general findings did not show a consistent direction, indicating that text complexity gradually did not increase as grade levels went up. There were, also, substantial differences among textbooks published by different publishers within the same grade. The aforementioned studies demonstrate how the Coh-Metrix system is successfully used to investigate whether textbooks are systematically designed to maximize learning gains.

Previous studies looking into text complexity across grade levels were mainly conducted for reading materials; in comparison, studies conducted for listening scripts are lacking (Kim & Jeon, 2013; Song, 2013). In addition, there is an urgent need to analyze a series of English textbooks newly developed by the 2015 National Curriculum to examine whether the continuity among textbooks is well established.

This study attempted to examine whether the text difficulty was systematically improved as the grade level increased by analyzing the listening scripts of middle school English textbooks. To this end, the current study made plans to carry out a comprehensive and systematic analysis using 23 indices provided by Coh-Metrix, including surface and deeper measures. Basically, the current study intended to present implications for future textbook development and pedagogical implications for teachers and English language learners.

II. PREVIOUS STUDIES

With the growing importance of English education in an age of globalization, studies on English textbooks as the primary medium for English education have been actively conducted. Research has been conducted on four skills of English (*studies on listening and speaking*: Im & Jeon, 2013; Jang & Jang, 2018; *studies on reading*: Jeon, 2011, 2014, 2014; Jeon & Im, 2009, 2010; Lee & Hyun, 2013; *studies on writing*: Son & Lee, 2005).

Studies on English textbooks have seen crucial shifts in methodological approaches. Early studies were mostly conducted manually in data collection and analysis (Kim, 2001; Son & Lee, 2005; Song & Yoon, 2003). For instance, Kim (2001) conducted an analytical study on listening and speaking activities of the first grade English textbooks developed by the 7th National Curriculum. To this end, a qualitative analysis was performed on whether the textbooks were properly organized by the 7th National

Curriculum and the creative and communicative use of language based on the data analysis framework put forward by previous researchers. The findings showed that a large portion of the textbooks did not conform to the 7th National Curriculum and the principles of creative and communicative language use.

Several qualitative studies based on manual work have been conducted for analyzing English textbooks, including studies on cultural aspects in textbooks (Choi, 2015; Jang & Kim, 2018; Kang & Hong, 2016), students' perception of grammar items in textbooks (Jung & Kim, 2017), conformity with the achievement standards of textbooks (Kwon & Kang, 2017), and types of activities in the English textbooks (Kim, 2018).

While there are studies that clearly benefit from qualitative approaches such as the above, other studies looking into the linguistic components of textbooks (e.g., the text difficulty of textbook materials) use advanced text analysis tools developed in the fields of computational linguistics (Jurafsky & Martin, 2008) and corpus linguistics (Meyer, 2002). Studies using the text analysis tools are detailed in the next section.

1. Studies on English Textbooks Using the Corpus Linguistic Methodology

Technological advances in the fields of computational linguistics (Jurafsky & Martin, 2008) and corpus linguistics (Meyer, 2002) have enabled analysis of a vast amount of textual materials in a more objective and systematic manner, using automated text analysis tools. They have opened new opportunities for textbook research, thereby enhancing objectivity and reliability of data analysis. A number of studies have employed tools such as Wordsmith Tools (Kim & Hwang, 2018; Kim, Lee, & Kim, 2017), NLP Tools (Kim & Chun, 2008), Antconc (Yu & Kim, 2015; Shin, 2017), Range (Lee & Lee, 2019), and VocabProfile (Ha & Cho, 2016).

A particularly notable study was conducted by Shin (2017), who used the keyword extraction function of the AntConc program to investigate the authenticity of vocabulary used in high school English textbooks. The reading passages in English textbooks did not provide enough examples to demonstrate actual uses of function words. In addition, the frequency of negative keywords was much higher in the textbooks compared to the existing native corpus (i.e., the British National Corpus). Based on the empirical data, this study observed that the authenticity of the reading passages in the English textbooks was low.

Text analysis tools can facilitate the objectivity and reliability of textbook research; however, the majority of the existing analysis tools mainly have operated at the vocabulary or sentence level with the limitation on the number of aspects investigated (Jeon, 2011; Ryu & Jeon, 2020). To gain a more comprehensive and multidimensional view of text difficulty in different textbooks, it is necessary to include measures not only

at the lexical level but also at the discourse level such as cohesion and coherence (Jeon, 2015; Ryu & Jeon, 2020). Several studies have been conducted to achieve this goal, and some Coh-Metrix-based studies of English textbooks are presented in the following section.

2. Studies on English Textbooks using the Coh-Metrix System

Researchers at the Institute for Intelligent Systems at the University of Memphis made a concerted effort to develop the Coh-Metrix system, based on computational linguistic algorithms and corpus linguistic data (Graesser et al., 2007). Specifically, the Coh-Metrix system is predicated on the computational linguistics algorithms, such as parts-of-speech tools (Brill, 1995) and Latent Semantic Analysis (Landauer, 2007), including various types of corpus data used to improve the accuracy of genre analysis; therefore, it is assumed that the system provides an objective and reliable data making it possible to measure text characteristics at the discourse level (e.g., cohesion and coherence) as well as at the lexical level.

The validity of the text analysis can be achieved by the systematic analysis using comprehensive features. In light of this, a number of studies have investigated the multidimensional nature of texts using the Coh-Metrix system (Im & Jeon, 2013; Jeon, 2014, 2015; Jeon & Im, 2010; Kim & Jeon, 2013; Lee, 2013; Song, 2013). Jeon (2014) analyzed whether the text difficulty of the reading passages in English textbooks for the first, second and third year of high schools implementing the 2009 National Curriculum. In his study, 10 measures in Coh-Metrix were chosen. The finding was that while the difficulty of frequently occurred words, Flesh Kincaid readability scores, and noun phrase density scores were shown to increase with grade levels, there was no systematic increase in text complexity in terms of word and sentence count, Flesh-Kincaid Grade Level scores, the number of words occurring before the main verb, type-token ratio, and referential and semantic cohesion.

A similar finding was observed in the study conducted by Lee (2013). He analyzed various linguistic, psycholinguistic measures of the reading passages in high school English textbooks for the second and third graders. The result showed that for the number of words, semantic cohesion, and Flesh-Kincaid Grade Level, the third-grade textbooks were more difficult than the second-grade textbooks, whereas the same pattern was not revealed in other measures. To sum up, these studies suggest that textbook developers may have been mostly concerned with the surface level linguistic features in leveling text difficulty across grades; thus, they may be limited to develop textbooks which reflect the actual reading competence of high school students in a more comprehensive fashion.

Although previous studies on the first-to-third grade continuity focused on the reading materials, few analyzed whether there is a systematic inter-grade link in the listening passages. One of the few exceptions is Kim and Jeon's study (2013) which built corpus data collecting listening passages from middle school English textbooks for the first, second and third graders implementing the 2009 Curriculum. They used a variety of measures to examine whether text difficulty of the English textbooks systematically increase with grade levels. The result showed that some measures (i.e., the number of words and sentences, readability scores, and the number of words occurring before the main verb) demonstrated a consistent increase in text difficulty as the grade level increases while this was not the case for other measures (i.e., noun phrase density, lexical diversity, and referential and semantic cohesion).

Previous studies (Kim & Jeon, 2013; Song, 2013) contributed to the extent that they have created a basis for the future research by shedding a new light on the significance of listening passages in the corpus study. In other words, it is suggested that, as in the case of reading, listening proficiency of L2 learners is likely to increase as the grade increases; therefore, a more careful consideration of text difficulty is required to provide a systematic and efficient teaching and learning of English.

Notably, with the recent introduction of the 2015 National Curriculum, English textbooks have been newly developed. Accordingly, it is a timely issue to examine whether the textbooks have been written systematically enough to reflect inter-grade continuity in text difficulty. The current study aims to analyze whether the text difficulty of listening passages of newly developed middle school English textbooks systematically increases with grade levels, using various linguistic and psycholinguistic measures provided by Coh-Metrix.

III. METHODOLOGY

1. Textbook Corpus for the Current Study

In the present study, the entire listening scripts of middle school English textbooks revised by the 2015 National Curriculum for the first, second and third graders were constructed as a corpus for the current study. Specifically, all 13 English textbooks per grade currently published by 10 different publishers were selected for this study. A total of 310 listening scripts extracted from 39 English textbooks were used for analyzing the continuity among grades with a wide range of measures of Coh-Metrix.

2. Corpus Compilation and Textbook Analysis Tool

In this study, a textbook corpus (txt files) in Unicode format was constructed for analyzing the inter-grade connectivity using Coh-Metrix developed by the Institute for Intelligent Systems at the University of Memphis in the United States. Specifically, the textbook corpus consisted of listening scripts extracted from 13 English textbooks. The list of middle school English 1, 2, and 3 textbooks analyzed for the current study was presented in Table 1.

TABLE 1

List of English Textbooks for the Current Study

List of English Textbooks	Publisher
Middle School English 1, 2, 3	Kumsung Publishing Co. (Choi et al.)
Middle School English 1, 2, 3	Darakwon Inc. (Kang et al.)
Middle School English 1, 2, 3	Dong-A Publishing Co. (Yoon et al.)
Middle School English 1, 2, 3	Dong-A Publishing Co. (Lee et al.)
Middle School English 1, 2, 3	Mirae-N Co. (Choi et al.)
Middle School English 1, 2, 3	Visang Education Co. (Kim et al.)
Middle School English 1, 2, 3	NE Neungyule Inc. (Kim et al.)
Middle School English 1, 2, 3	NE Neungyule Inc. (Yang et al.)
Middle School English 1, 2, 3	YBM Co. (Park et al.)
Middle School English 1, 2, 3	YBM Co. (Song et al.)
Middle School English 1, 2, 3	Jihaksa Publishing Co. (Min et al.)
Middle School English 1, 2, 3	Chunjae Education Inc. (Lee et al.)
Middle School English 1, 2, 3	Chunjae Education Inc. (Jeong et al.)

These listening scripts were separated by each chapter. So, the corpus for the first graders contained 103 files. The corpus for the second graders consisted of 104 files and the corpus for the third grade had 103 files. A total of 310 files of listening scripts were analyzed for examining the inter-grade continuity with the Coh-Metrix system. The Coh-Metrix measures selected for the current study included basic counts (the number of words, the number of sentences, the average word length, the average sentence length), word frequency, word features (imageability, concreteness, age of acquisition), lexical diversity (type-token ratio), personal pronouns (the first-person pronouns, the second person pronouns, the third person pronouns), connectives (causal connectives, additive connectives, temporal connectives), readability scores (Flesch Reading Ease score, Flesch-Kincaid Grade Level), syntactic complexity (noun phrase density, words before the main verb), referential cohesion (argument overlap for adjacent sentences, argument overlap for all sentences), and semantic cohesion (LSA cosine scores for adjacent sentences, LSA cosine score for all sentences). The reason these measures were used in

this study is that these measures have been widely used in textbook studies (Jeon, 2015; Jeon & Im, 2010; Ryu & Jeon, 2020) and prior Coh-Metrix studies also have applied these measures for analyzing English corpora (Graesser, Jeon, Cai, and McNamara, 2008; Graesser, Jeon, Yan, & Cai, 2007).

IV. RESULTS AND DISCUSSION

A one-way analysis of variance (ANOVA) was performed to analyze the continuity of the listening scripts of middle school English textbooks for the first graders, the second graders, and the third graders. Specifically, the independent variable was the grade level and the dependent variable was each Coh-Metrix measure. Each statistical analysis was conducted at the significance level of 5% (p = .05).

1. Basic Counts

The results of the basic counts were summarized in Table 2. First, as shown in Table 2, the inter-grade continuity of the listening scripts of middle school English textbooks was well established for the number of words, the number of sentences, and the average sentence length. Specifically, the scores of these measures significantly increased with grade levels went up. Tukey HSD post hoc analyses indicated the significant differences among grade levels (all ps < .05). For the average word length, however, the differences among grade levels were not statistically significant. Given that the word length is one of the critical linguistic factors affecting the difficulty of text, it is advisable that this particular variable should also be taken into account for future textbook development.

TABLE 2

Results of Basic Counts

Indices	ME 1 (n=103)	ME 2 (n=104)	ME 3 (n=103)	F	p
Number of words	301(105)	415(143)	504(170)	52.92	.000*
Number of sentences	55.8(17.4)	67(18.8)	74.6(20.3)	25.95	.000*
Average sentence length	5.38(.69)	6.14(.68)	6.69(.76)	88.22	.000*
Average word length	1.25(.06)	1.25(.05)	1.27(.05)	2.06	.130

*p < .05

2. Word Frequency

Word frequency is an essential factor influencing the text difficulty (Graesser et al.,

2007; Graesser, McNamara, Louwerse, & Cai, 2004). So, it is pivotal to control the frequency of the words used in textbooks for the gradual manipulation of the difficulty of the textbooks across grades in textbook development (Jeon, 2011, 2015). The Coh-Metrix system uses the CELEX vocabulary database created by the Dutch Centre for Lexical Information to calculate vocabulary frequency measures (Bayen, Pippenbrock, & Glickers, 1995). The current study adopted log-transformed word frequency measures instead of raw word frequency scores; the log-transformed scores are commonly used because the distribution of the log-transformed scores tends to be more approximated to the normal distribution (Graesser et al., 2004).

The analysis presented in Table 3 showed that the continuity of the listening scripts in the middle school English textbooks was not well established for the measure of word frequency (for content words). The current results along with the results of the average word length suggest that these lexical elements should be reflected in future textbook development for the text difficulty manipulation.

TABLE 3

Results of Word Frequency

Indices	ME 1 (n=103)	ME 2 (n=104)	ME 3 (n=103)	F	p
Word frequency (content words)	2.48(.14)	2.48(.10)	2.51(.10)	1.85	.160

*p < .05

3. Word Features

Word features such as imageability, concreteness, and age of acquisition are all vital factors affecting the difficulty and comprehension of text (Graesser et al., 2004; Ryu & Jeon, 2020). The Coh-Metrix system provides estimates of word features using the values of the word feature measurements contained in the MRC Psycholinguistic vocabulary database (Coltheart, 1981). The values of the word features are within the range of 100 to 700.

TABLE 4

Results of Word Features

Indices	ME 1(n=103)	ME 2 (n=104)	ME 3 (n=103)	F	p
Imageability	418(18.9)	413(16.7)	407(18.0)	8.43	.000*
Concreteness	381(24.6)	379(18.7)	373(19.8)	3.75	.025*
Age of acquisition	268(27.0)	275(18.2)	281(21.7)	7.61	.001*

*p < .05

The analyses of word features, as presented in Table 4, revealed the statistically significant differences among grade levels for imageability scores. Tukey HSD post hoc analyses were conducted to verify specific differences between grades. The results showed that only the difference between the textbooks of the first and the third graders was statistically significant ($p < .05$), whereas no statistical differences were observed between the first and the second graders, and between the second and the third graders. The same pattern was true for the analyses of concreteness and age of acquisition; that is, the overall differences across grades were statistically significant, whereas only the differences between the first and the third graders were statistically significant ($ps < .05$; i.e., Tukey HSD post hoc analysis). These findings suggest that more concrete guidelines need to be provided in terms of word features for future textbook development.

4. Lexical Diversity

The measure for the lexical diversity is represented as a type-token ratio value (Graesser et al., 2004; Ryu & Jeon, 2020). A type refers to the frequency of the initial appearance of an individual vocabulary, and a token indicates the frequency of the repeated appearances of that particular vocabulary. Therefore, the type value is always 1. For example, if the word "*continuity*" is used five times, the type value of this word is 1 and the token value is 5. In this case, the type-token ratio value is 0.2 (1/5 = 0.2). A high type-token ratio value tends to increase the likelihood that various words are used in the text because the type value is relatively high and the token value relatively is low. Thus, higher lexical diversity values are inclined to increase the difficulty of text due to the increased cognitive load of dealing with novel words (Jeon, 2015; Ryu & Jeon, 2020).

TABLE 5

Results of Lexical Diversity

Indices	ME 1(n=103)	ME 2(n=104)	ME 3(n=103)	F	p
Type-token ratio (content words)	.61(.09)	.62(.08)	.62(.08)	.94	.391

*$p < .05$

The analyses on the listening scripts in the middle school English textbooks, as presented in Table 5, showed that the inter-grade connectivity was not well manipulated for the lexical diversity. These results suggest that lexical diversity as well as word length, word frequency, and word features are required to be considered for the systematic manipulation of text difficulty across grade levels.

5. Personal Pronouns

As the number of pronouns increases in the text, the text difficulty tends to become higher because the search for the referents for them tend to be more complex (Graesser et al., 2007). The Coh-Metix system provides first person pronouns, second person pronouns, and third person pronouns. The analysis of the personal pronouns, as indicated in Table 6, revealed that the inter-grade continuity of the listening materials for the middle school English textbooks were not well-coordinated for all pronouns (all *p*s > .05). These results suggest that these personal pronouns (i.e., first person pronouns, second person pronouns, third person pronouns) should also be considered in developing English textbooks for the future.

TABLE 6

Results of Personal Pronouns

Indices	ME 1 (*n*=103)	ME 2 (*n*=104)	ME 3 (*n*=103)	*F*	*p*
First person pronouns	37.98(13.1)	38.35(10.0)	38.12(10.4)	.03	.973
Second person pronouns	26.05(11.0)	24.69(8.44)	24.07(9.0)	1.16	.314
Third person pronouns	8.86(11.0)	7.41(6.92)	8.20(6.38)	.79	.455

*$p < .05$

6. Connectives

Connectives in the text generally tend to promote the semantic connections between sentences (Millis & Just, 1994; Murray, 1997). The Coh-Metrix system provides three types of connectives, including causal connectives, additive connectives, and temporal connectives.

TABLE 7

Results of Connectives

Indices	ME 1 (*n*=103)	ME 2 (*n*=104)	ME 3 (*n*=103)	*F*	*p*
Causal connectives	22.3(17.9)	22.6(10.5)	25.8(9.96)	2.22	.111
Additive connectives	3.71(4.46)	5.66(4.62)	6.82(5.15)	11.27	.000*
Temporal connectives	19.6(13.5)	19.7(11.5)	20.5(9.83)	.19	.830

*$p < .05$

As presented in Table 7, listening passages of the middle school English textbooks did not show any statistically significant differences across grades for the causal and temporal connectives. There were, however, significant differences in the number of additive connectives. Tukey HSD post hoc analyses revealed that the differences

between the first and the second graders, and between the first and the third graders were significant (all *p*s < .05) respectively, while the difference between the second and the third graders was not statistically significant. One way to interpret these results is to understand that the additive connectives in the textbook tend to increase as the length of sentence increases across grade levels. The current findings point out the need for a more systematic manipulation of text difficulty across grades in terms of the number of connectives (i.e., causal connectives, additive connectives, temporal connectives).

7. Readability Indices

The Coh-Metrix system provides readability indices such as the Flesh Reading Ease (FRE) score and the Flesch-Kincaid Grade Level (FKGL), as presented in Table 8 (Graesser et al., 2007). FRE scores range in 0 to 100, and FKGL scores range in 1 to 12. The higher the FRE score, the less difficult the text; in contrast, the higher the FKGL score, the more difficult the text.

TABLE 8

Results of Readability Indices

Indices	ME 1 (*n*=103)	ME 2 (*n*=104)	ME 3 (*n*=103)	*F*	*p*
Flesh Reading Ease	95.1(4.27)	94.2(4.19)	92.9(4.46)	6.59	.002*
Flesch-Kincaid Grade Level	1.28(.74)	1.61(.69)	1.95(.75)	21.61	.000*

*$p < .05$

The analysis of readability indices in Table 8 indicated that the listening passages of the English textbooks were systematically written, indicating that the text difficulty gradually increased as the grade level went up. According to Tukey HSD post hoc analyses, however, only the difference between the first and the third graders was statistically significant for FRE ($p < .05$). On the other hand, the differences between all grades for the FKGL scores were statistically significant (all *p*s < .05). Therefore, for the FKGL index, the inter-grader connectivity appears to be carefully designed. In addition, considering that the FRE scores and the FKGL scores are determined by the word length and the sentence length, these results are in line with the sentence length analysis.

8. Syntactic Complexity

The Coh-Metrix presents the measures of syntactic complexity by the noun phrase density and the number of words before main verbs (Graesser et al., 2004). The noun

phrase density score denotes the number of modifiers of headnouns divided by the entire words that comprise the noun phrase. In general, the higher the score of the two measures, the more difficult the text is (Graesser et al., 2004; Ryu & Jeon, 2020).

The analysis of the syntactic complexity, as shown in Table 9, revealed that there was no statistically significant difference between grades for the noun phrase density scores. A significant difference, however, weas observable across grades for the words before main verbs. Tukey HSD post hoc analyses verified that all of the differences between each grade level were statistically significant (all *p*s < .05). These results suggest that the text difficulty was well structured across grades for such measures as the sentence length that is relatively easy to control, whereas microscopic features such as the noun phrases used in sentences were not well reflected. The current findings suggest that measures for the noun phrase density should be well reflected in future curriculum.

TABLE 9

Results of Syntactic Complexity

Indices	ME 1 (*n*=103)	ME 2 (*n*=104)	ME 3 (*n*=103)	*F*	*p*
Noun phrase density	.47(.10)	.48(.09)	.47(.09)	.43	.650
Words before main verbs	1.03(.21)	1.14(.24)	1.25(.28)	21.02	.000*

*$p < .05$

9. Reference Cohesion

Reference cohesion is one of the most crucial linguistic components affecting text comprehension (Cirilo, 1981; Haberlandt & Bingham, 1978; Jeon, 2011). Reference cohesion is formed when arguments are commonly used in two sentences in a text. Arguments refer to nouns, pronouns, and noun phrases (Graesser et al., 2004). If the reference cohesion between sentences in text is high, the text comprehension is purported to be facilitated. As a consequence, the text difficulty is reduced.

TABLE 10

Results of Reference Cohesion

Indices	ME 1 (*n*=103)	ME 2 (*n*=104)	ME 3 (*n*=103)	*F*	*p*
Argument overlap (adjacent sentences)	.22(.09)	.25(.09)	.28(.08)	16.15	.000*
Argument overlap (all sentences)	.20(.07)	.23(.07)	.26(.07)	14.92	.000*

*$p < .05$

The results of reference cohesion, as shown in Table 10, confirmed the statistical differences for the argument overlap between adjacent sentences across grade levels. According to Tukey HSD post hoc analyses, all of the inter-grade differences were statistically significant (all *p*s < .05). These results suggest that as the grade level increases, the reference cohesion also increases, which leads to reduction in text difficulty. It is of significance to connect sentences semantically, taking into account the semantic connection of the sentences included in the textbook. This, however, can increase the reference cohesion which reduces the difficulty of text comprehension. This is inclined to conflict with the general expectation that text difficulty gradually increases as grade level increases. Thus, as suggested in the study by Jeon (2015), these results suggest that reference cohesion needs to be controlled systematically. Likewise, the argument overlap for all sentences was statistically difference across grades. The difference between the second and the third graders, however, was shown not to be statistically different based on the Tukey HSD post hoc analysis ($p > .05$). From these results, it is suggested that the argument overlap should be well controlled in future textbook development in order to increase the difficulty of text as learners' expected proficiency increases.

10. Semantic Cohesion

The Coh-Metrix system applies the LSA mathematical model to measure semantic cohesion (Graesser et al., 2004). This model calculates the semantic similarity between words and sentences based on a vast size of corpus data. If the value is high, the cognitive load for text comprehension tends to decrease.

TABLE 11

Results of Semantic Cohesion

Indices	ME 1 (*n*=103)	ME 2 (*n*=104)	ME 3 (*n*=103)	*F*	*p*
LSA cosine (adjacent sentences)	.12(.05)	.11(.04)	.12(.03)	2.35	.097
LSA cosine (all sentences)	.12(.05)	.11(.05)	.11(.04)	.31	.737

*$p < .05$

As shown in Table 11, the analyses of semantic cohesion displayed no statistical differences between grades for the LSA cosine measures for adjacent sentences and all sentences (all *p*s > .05). In other words, the inter-grade connectivity was not well established for these particular measures. These results suggest that computer tools such as the Coh-Metrix system can provide an aid in textbook development and revision

process by providing concrete and numerical evidence for such linguistic properties as semantic cohesion that is difficult to visually inspect otherwise.

V. CONCLUSION

In EFL settings in Korea, English textbooks used in actual educational settings serve as the main source of linguistic input for learners. For this reason, research on English textbooks has been carried out in various scale. Among other components of textbooks, researchers have been paying attention to the textual input in textbooks, given that the text *per se* plays a key role in promoting linguistic proficiency (Chen, 2016, Plakans & Bilki, 2016). Researchers have emphasized the significance of selecting the appropriate level of text difficulty because it is purported that a careful matching of the level of text difficulty to a learner's proficiency levels, or to the slightly higher level, would maximize the learning effect (Berendes et al., 2018; Chen, 2016; Krashen, 1985; Vygotsky, 1978). There has been a series of studies examining whether the difficulty of text gradually increases as the grade increases, or as chapters progress within a grade (Berendes et al., 2018; Chen, 2016). Previous studies have been mainly concerned with reading passage; however, in order to help learners to develop their listening proficiency, it is also to be considered to adjust the difficulty of listening passages according to the increase of grade level. The current study aims to fill this research gap by analyzing the difficulty of listening passages in the newly developed English textbooks for the first, second and third graders developed under the 2015 National Curriculum. Specifically, the difficulty level of English textbooks was examined using the Coh-Metrix system to see if it is systematically increased with grade levels.

In the study, a total of 23 measures were investigated including basic counts, word frequency, word features, lexical diversity, personal pronouns, connectives, readability indices, syntactic complexity, reference cohesion and semantic cohesion. The analysis of basic counts reveals that except for the average length of words, the other measures for the average number of words and sentences and the average length of sentences displayed a consistent pattern in which the difficulty rises as the grade level increases. As the length of sentences increases with rising grade levels, the number of additive connectives increases accordingly. Similarly, it appears that a systematic arrangement of text difficulty is observed among readability indices (i.e., FKGL, FRE scores). The same pattern applies to syntactic complexity: the number of words occurring before main verbs increased with grade levels, which would systematically facilitate learners' acquisition of more complex structures as their grades go up. These results suggest that the macroscopic measures that are easily identifiable seems to be reflected in the process

of textbook development. In addition, word features and reference cohesion which are deeper-level linguistic measures turned out to have been controlled appropriately.

On the contrary, this systematic tendency is not observed in other measures. First, there were no significant differences in word frequency on content words and lexical diversity across grade levels; this was counter to the general expectations that more of less frequent words and more diverse words would be distributed as grade levels increase. In addition, the inter-grade connectivity was not established for the personal pronouns, which implies that these features may not have been considered in the process of textbook development. As for the connectives, the differences in the causal and temporal connectives across grade levels were not statistically significant. For the syntactic complexity, noun phrase density scores were not significantly different across grades. However, to promote the ability to process more complex structures, it may be desirable to consider the noun phrase density within a sentence rather than simply increasing the average sentence length or the number of words. Lastly, for semantic similarities, LSA cosine between adjacent sentences and all sentences were not statistically different across grades.

The current findings are comparable to the previous studies in which they analyzed listening scripts of English textbooks for the first, second and third graders developed under the 2009 National Curriculum (Kim & Jeon, 2013 for middle school; Song, 2013 for high school). They reported a well-established inter-grade continuity for such measures, which are easily identifiable, as the number of words and sentences, readability indices (FRE, FKGL), and the words occurring before main verbs. It is also noted that lexical diversity and syntactic complexity were not appropriately controlled. In particular, semantic cohesion measures, which are not easily identifiable, were not significantly different across grades. The fact that the current results are quite similar to those reported by studies based on the previous curriculum suggests that textbook developers seemed to have followed the similar guidelines in designing the new textbooks. The current study as well as previous studies (Kim & Jeon, 2013; Song, 2013) suggests that more diverse measures should be considered for future textbook development and that automatic text analysis tools such as Coh-Metrix may facilitate the process along with enhanced objectivity and reliability.

Suggestions for further study are presented, based on the limitations of the current study. First, the current study adopted a total of 23 measures for the analysis of listening scripts; this is a significant increase compared to the 10 measures used in the previous studies (Kim & Jeon, 2013; Song, 2013). Nevertheless, more measures need to be included in further studies, which is required to advance textbook development and evaluation and to help classroom teachers to select textbooks to meet the needs of their students by considering more diverse features. In addition, the current study addressed

listening scripts only; further studies need to be conducted to make a systematic, linguistic comparison between reading and listening passages, which may enrich our understanding of the overall characteristics of spoken and written language.

To sum up, previous studies have primarily addressed the issue of the systematic arrangement of text difficulty for reading passages; however, the present study suggests that listening passages can be analyzed to maximize potential learning gains. This study has demonstrated that a variety of measures, provided by Coh-Metrix, can ensure that the analysis is conducted with greater objectivity and validity. In addition, it is expected in further studies that Coh-Metrix serves as an effective tool to analyze various surface- and deep-level linguistic measures for diverse text materials as well as textbooks so that it can enhance the efficiency of the process of developing textbook materials, text evaluation and selection.

REFERENCES

Bae, J. (2019). A continuity analysis of the reading passages in elementary and middle school English textbooks with 2015 Revised National Curriculum. *Modern Studies in English Language and Literature, 63*(2), 81-107.

Berendes, K., Vajjala, S., Meurers, D., Bryant, D., Wagner, W., Chinkina, M., & Trautwein, U. (2018). Reading demands in secondary school: Does the linguistic complexity of textbooks increase with grade level and the academic orientation of the school track? *Journal of Educational Psychology, 110*(4), 518-543.

Brill, E. (1995). *The CELEX Lexical Database (CD-ROM)*. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Chen, A. C. (2016). A critical evaluation of text difficulty development in ELT textbook series: A corpus-based approach using variability neighbor clustering. *System, 58*, 64-81.

Cho, W. (2003). An analysis of the English materials of the middle school English textbooks for the 2nd year students based on the 7th Curriculum. *The British and American Language and Literature Association, 19*(1), 151-169.

Choi, J. (2015). Cultural aspects on the Korean and Japanese middle school English textbooks: Based on the critical pedagogy. *Journal of Language Sciences, 22*(3), 219-253.

Cirilo, R. (1981). Referential coherence and text structure in story comprehension. *Journal of Verbal Learning and Verbal Behavior, 20*, 358-367.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology, 33*, 497-505.

Go, G., & Park, J. (2007). A corpus-based study of continuity in Korean middle school English textbooks. *English Language and Linguistics, 24*, 27-45.

Graesser, A. C., Jeon, M., Cai, Z.,& McNamara, D. S. (2008). Automatic analyses of language, discourse, and situation models. In J. Auracher & W. van Peer (Eds.), *New Beginnings in Literary Studies* (pp. 72-88). Cambridge: Cambridge Scholars Publishing.

Graesser, A. C., Jeon, M., Yan, Y., & Cai, Z. (2007). Discourse cohesion in text and tutorial dialogue. *Information Design Journal, 15*(3), 199-213.

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai. Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavioral Research Methods, Instruments, and Computers, 36*(2), 193-202.

Ha, O., & Cho, Y. (2016). A comparative analysis of the reading passages of National Assessment of Educational Achievement and middle school English textbooks: Focus on lexical difficulty and readability. *Foreign Languages Education, 23*(3), 191-213.

Haberlandt, K. & Bingham, G. (1978). Verbs contribute to the coherence of brief narrative: Reading related and unrelated sentence triples. *Journal of Verbal Learning and Verbal Behavior, 17*, 419-425.

Im, I., & Jeon, M. (2013). A corpus-based analysis of the listening materials in middle school English 1 textbooks. *Korean Journal of English Language and Linguistics, 13*(1), 105-134.

Jang & Chang (2018). Corpus-based lexical analysis of reading and listening data of high school English textbooks. *Journal of Studies in Language, 34*(2), 261-286.

Jang, N., & Kim, J. (2018). An analysis of cultural contents in elementary school English textbooks for grade 3 and 4 based on the 2015 revised Korean National Curriculum. *The Journal of the Korea Contents Association, 18*(8), 1-8.

Jeon, M. (2011). A corpus-based analysis of the continuity of the reading materials in middle school English 1 and 2 textbooks with Coh-Metrix. *The Journal of Linguistics Science, 56*, 201-218.

Jeon, M. (2014). An analysis of the continuity of the reading materials in high school English textbooks. *Journal of the Korean Data Analysis Society, 16*(2), 925-938.

Jeon, M. (2015). An analysis of the continuity among middle school English textbooks with an automated language analysis program. *Modern English Education, 16*(1), 195-218.

Jeon, M., & Im, I. (2009). A corpus-based analysis of middle school English 1 textbooks with Coh-Metrix. *English Language Education, 21*(4), 265-292.

Jeon, M., & Im, I. (2010). A corpus-based linguistic analysis of high school English textbooks. *Studies in Foreign Language Education, 17*(1), 209-233.

Jung, H., & Kim, H. (2017). An analysis of grammar activities in middle school English textbooks and learners' perspectives. *Journal of the Korea English Education Society*, *16*(1), 129-146.

Jurafsky, D., & Martin, J. H. (2008). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, NJ: Prentice-Hall.

Kang, S., & Hong, K. (2016). Cultural content in 5th-& 6th-grade English textbooks on 2009 National Curriculum. *The Mirae Journal of English Language and Literature, 21*(1), 339-365.

Kim, E. (2001). A study on the listening and speaking activities in the middle school English textbooks. *English Teaching, 56*(2), 219-243.

Kim, I. (2018). An analysis of writing activities in the 3rd and 4th grade English textbooks based on 2015 National English curriculum. *The Jungang Journal of English Language and Literature, 60*(3), 339-359.

Kim, J., & Chun, Y. (2008). A study on the continuity of elementary and middle school English textbooks through a corpus-based analysis: the 6th grade elementary school English textbook and the 1st grade middle school English textbook. *ENGLISH TEACHING, 63*(2), 361-394.

Kim, J., & Jeon., M. (2013). A corpus-based analysis of the continuity of the listening materials in middle school English textbooks for Coh-Metrix measures. *Journal of the Korean Data Analysis Society, 15*(4), 1987-2000.

Kim, S., & Jeon, M. (2016). An analysis study of English writing of elementary school 6th grade English language learners using Coh-Metrix. *Modern English Education, 17*(3), 263-287.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Boston: Addison-Wesley Longman Ltd.

Kwon, Y., & Kang, M. (2017). A comparative analysis of grammar types: focusing on the 2nd grade middle school English textbooks. *Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology, 7*(3), 765-774.

Landauer, T. K. (2007). LSA as a theory of meaning, In T. K. Landauer, D. S. McNamara & W. Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 3-34). Mahwah, NJ: Lawrence Erlbaum Associates.

Lee, J., & Hyun, T. (2013). A comparative analysis of the reading materials in the middle school English textbooks for the 3rd year students. *Modern English Education, 14*(4), 263-286.

Lee, S. (2013). A corpus-based analysis of the difficulty of the reading materials in high school English 1 and 2 textbooks using Coh-Metrix. *Studies in Foreign Language Education, 27*(2), 131-148.

Lee, S., & Lee, H. (2017). The distribution of English vocabulary in middle school English III textbooks. *Secondary English Education, 10*(4), 121-142.

Lee, S., & Lee, H. (2019). An analysis of the distribution of vocabulary in middle school English 2 textbooks based on 2015 Revised English National Curriculum. *Secondary English Education, 12*(3), 71-93.

Meyer, C. F. (2002). *English corpus linguistics: An introduction*. Cambridge: Cambridge University Press.

Millis, K., & Just, M. (1994). The influence of connectives on sentence comprehension. *Journal of Memory and Language, 33*, 128-147.

Murray, J. (1997). Connective and narrative text: The role of continuity. *Memory & Cognition, 25*(2), 227-236.

Plakans, L., & Bilki, Z. (2016). Cohesion features in ESL reading: Comparing beginning, intermediate and advanced textbooks. *Reading in a Foreign Language, 28*(1), 79-100.

Ryu, J., & Jeon, M. (2020). A Coh-Metrix analysis of the first grade middle school English textbooks revised by the 2015 National Curriculum. *Korean Journal of English Language and Linguistics, 20*(1), 89-121.

Shin, C. (2017). An analysis of vocabulary authenticity in high school English textbooks. *The Mirae Journal of English Language and Literature, 22*(3), 167-193.

Son, M., & Lee, J. (2005). An analysis of writing activity materials in the middle school English textbooks. *English Language & Literature Teaching, 11*(3), 139-164.

Song, H. (2013). Exploring the connectedness according to grades in high school English textbooks. *English Language & Literature Teaching, 19*(4), 279-305.

Song, J., & Yoon, I. (2003). An analysis of the sequencing and grading in elementary school English textbooks: Focused on communicative functions and their exponents in the 7th National Curriculum. *Primary English Education, 9*(2), 149-188.

Vygotsky, L. S. (1978). The interaction between learning and development. In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.), *Mind in society: The development of higher psychological processes* (pp. 79-91). Cambridge, MA: Harvard University Press.

Yu, S., & Kim, H. (2015). Distributions of tense and aspect in readings from Korean high school textbooks. *The Journal of Foreign Studies, 34*(4), 77-98.

Examples in: English
Applicable Language: English
Applicable Level: Secondary

Ryu, Jisu (Instructor)
Department of English Language & Literature
Konkuk University
120 Neungdong-ro, Gwangin-gu, Seoul, Korea
Email: jsryu0508@konkuk.ac.kr

Jeon, Moongee (Professor)
Department of English Language & Literature
Konkuk University
120 Neungdong-ro, Gwangin-gu, Seoul, Korea
Email: mjeon1@konkuk.ac.kr

Received: July 13, 2020
Reviewed: August 24, 2020
Revised version received: September 5, 2020