



A Comparison of Linguistic Aspects in 2015-2019 NAEA English Listening Texts Between School Levels and Between Language Forms*

Jiyoung Lee

Pusan National University

ARTICLE INFO

Received: 31 October 2022

Revised: 27 November 2022

Accepted: 01 December 2022

Examples in: English

Applicable Languages: English

Applicable Levels: Secondary

KEYWORDS

*National Assessment
of Educational Achievement/
English listening text/
Coh-Matrix*

국가수준 학업성취도 평가/

영어 듣기 텍스트/

코메트릭스

ABSTRACT

Lee, Jiyoung. (2022). A comparison of linguistic aspects in 2015-2019 NAEA English listening texts between school levels and between language forms. *Modern English Education*, 23(4), 1-11.

The current study aims to investigate English listening texts of the National Assessment of Educational Achievement (NAEA) at word, sentence, and text levels. It classifies NAEA listening texts into two school levels (middle school and high school) and two forms of spoken language (dialogues and monologues). It is necessary to examine changes of linguistic aspects in the NAEA for both middle and high school students to figure out whether text difficulty systematically rises as school level increases. The corpus for this study includes 169 listening texts from five academic years from 2015 to 2019. In this period, tests were implemented as paper-based tests. Two research questions are about linguistic differences of NAEA listening texts between middle school and high school and between dialogues and monologues at word, sentence and text levels. Coh-Matrix analyzed descriptive measures related to words, lexical diversity, and word information for word-level analysis, descriptive measures related to sentences and syntactic complexity for sentence-level analysis, and readability and referential cohesion for text-level analysis. Results and educational implications could increase the quality of NAEA listening texts.

I. INTRODUCTION

The National Assessment of Educational Achievement (NAEA) is implemented in order to inspect the practice of the national curriculum in the education field in Korea. Likewise, many countries examine the academic achievement in the national level including the National Assessment of Educational Progress (NAEP) in the United States and the National Curriculum Assessment (NCA) in England. These tests, which are not entrance exams, also play an important role to improve the quality of teaching

and learning in class, and to find the way to support the education field (I. Lee et al., 2015). For the education of middle school, only the third graders of middle school take the NAEA of Korean, math, English, science and social studies. For the education of high school, only the second graders of high school take the NAEA of Korean, math and English.

English tests of the NAEA include reading comprehension and listening comprehension in order to examine students' proficiency related to spoken language and written language. The existing Coh-Matrix studies on the NAEA

*This work was developed based on the presentation, *A Coh-Matrix Analysis of English Listening Texts in the National Assessment of Educational Achievement for Middle and High School Students*, in the 2022 MEESO International Conference.

Jiyoung Lee (PhD)

Department of English Language Education, Pusan National University, 2, Busandaehak-ro 63beon-gil, Geumjeong-gu, Busan, 46241, South Korea

Email: jylee1915@gmail.com / ISNI: 0000 0005 0718 4049

researched only reading comprehension (J. Kim, 2018; J. Lee, 2021a). The two studies focused on the change of text difficulty because the previous studies considered the proper text difficulty of testing materials to be important. The study conducted by J. Kim (2018) investigated whether the 2015-2017 NAEA for middle school students maintained text difficulty within the same school level. The study conducted by J. Lee (2021a) examined whether the 2015-2019 NAEA became systematically more difficult as the school level increased by comparing text difficulty between the different school levels.

The levels of difficulty in NAEA listening texts are also needed to be researched as communicative competence is one of the goals in English language education. Especially, the comparison between middle school and high school has not been actively conducted for the discrimination between the different school levels in the NAEA. Thus, this study tries to compare the NAEA listening texts between middle school and high school.

English listening tests, such as the NAEA, the College Scholastic Ability Test (CSAT) and TOEIC, include two forms of spoken language, dialogues and monologues. When analyzing listening materials, the different forms of spoken language need to be considered. It is necessary to identify the differences between dialogues and monologues in order to understand listening comprehension tests (Papageorgiou et al., 2012). Dialogues and monologues are differentiated in accordance with the number of speakers. In terms of monologues, one speaker participates in announcements, speeches, lectures, and explanation (Brown, 2007; H. Jung, 2010). Two or more speakers in dialogues make social relationships and interactions between the speakers (Brown, 2007). The previous studies on the CSAT analyzed the differences between dialogues and monologues (J. Kim & H. Kim, 2021; J. Lee, 2021b). In addition, the two previous studies found the clearly different linguistic aspects between dialogues and monologues. Unlike the existing CSAT research, the NAEA studies hardly analyzed the differences between the two forms of spoken language.

The current study tries to investigate the linguistic aspects of the NAEA listening texts by classifying into the two different school levels (middle school and high school), and into the two forms of spoken language (dialogues and monologues). This study could identify whether dialogues and monologues change as the school levels are different. This study can improve the quality of future NAEA listening texts. Thus, the present study focuses on the following two research questions.

- 1) Are NAEA listening texts for middle and high school students different at word, sentence and text levels?
- 2) Are dialogues and monologues of NAEA listening texts different at word, sentence and text levels?

II. THEORETICAL FRAMEWORK AND LITERATURE REVIEW

Research on the discrimination of different grades of academic texts mainly focused on textbooks. Especially, the comparison has been usually conducted between different grades within the same school level. The discrimination of reading texts in textbooks for different graders was analyzed (M. Jeon, 2011, 2015; S. H. Lee, 2013; J. Ryu & M. Jeon, 2020a). Middle school English 1 and 2 textbooks were well-distinguished in descriptive measures such as word counts, sentence counts and sentence length, readability such as the Flesch Reading Ease (FRE) and the Flesch-Kincaid Grade Level (FKGL), and word information such as word imaginability, word concreteness and age of acquisition. However, the continuity was not well-established in word frequency (M. Jeon, 2011). Middle school English 1, 2 and 3 textbooks were well-distinguished in descriptive measures, word information, readability, referential cohesion, semantic cohesion, syntactic similarity and situational model according to the grades, whereas the textbooks of the three different grades were not significantly different in word frequency and lexical diversity (J. Ryu & M. Jeon, 2020a). The discrimination of the middle school English 1, 2, and 3 textbooks was controlled properly for descriptive measures, age of acquisition, subject density and the FKGL, but not for other measures among descriptive measures, word frequency, readability, syntactic density, lexical diversity, referential cohesion, semantic cohesion, pronouns, and connectives (M. Jeon, 2015). High school textbooks between different grades were also compared by S. H. Lee (2013). The study denoted that the text difficulty between high school English 1 and 2 textbooks was not distinguished except for word counts and the FKGL.

Listening texts in textbooks were also analyzed in order to compare the level of texts between different grades within the same school level (J. Kim & M. Jeon, 2013; J. Ryu & M. Jeon, 2020b). The two previous studies revealed that some linguistic features properly distinguished the textbooks of different grades, whereas other linguistic features did not. More specifically, J. Kim and M. Jeon's (2013) study indicated that the continuity was well-established in word counts, sentence counts, word frequency, readability and left embeddedness. However, noun density, lexical diversity and referential cohesion were not significantly different. In the research conducted by J. Ryu and M. Jeon (2020b), the continuity of listening texts in middle school English 1, 2 and 3 textbooks was well-established based on word counts, sentence counts, sentence length, word information, and the FKGL. However, lexical diversity and word frequency did not significantly distinguish the listening texts of textbooks among the different grades of middle school.

For the analysis on listening texts in the CSAT, two existing studies (J. Kim & H. Kim, 2021; J. Lee, 2021b) analyzed the differences between dialogues and monologues. J. Kim and H. Kim (2021) analyzed dialogues and monologues in the 1994-2020 CSAT with reading texts. The results revealed that monologues were significantly different

from dialogues in terms of lexical characteristics such as lexical diversity and word frequency, syntactic complexity such as left embeddedness and modifiers per Noun Phrase (NP), and discourse features such as noun overlap, argument overlap and stem overlap. J. Lee (2021b) compared dialogues and monologues in the 2014-2021 CSAT. The study indicated that monologues were significantly more difficult than dialogues since monologues were textually difficult, syntactically complex and lexically difficult with more words. Although the comparison between dialogues and monologues exists, the research on NAEA listening texts has rarely compared the two forms of spoken language.

Text analysis on the NAEA usually adopted reading texts. In order to investigate the consistency of text difficulty among 2015, 2016 and 2017 academic years within the same school level, J. Kim (2018) conducted Coh-Metrix analysis on NAEA reading texts for middle school students. As a result, NAEA reading texts of middle school remained the level of text difficulty among the three different academic years as for descriptive measures, lexical characteristics, syntactic complexity and cohesion. However, sentence structures were getting more complex. In order to examine the discrimination of the NAEA reading texts between the different school levels, J. Lee (2021a) compared text difficulty of the 2015-2019 NAEA reading texts between middle school and high school. The study found that the NAEA reading texts of high school were significantly more difficult than those of middle school from the results of descriptive measures, word information and readability. However, the NAEA reading passages cannot be distinguished between the two different school levels from the results of lexical diversity, syntactic complexity and referential cohesion.

Even when comparing the NAEA and textbooks, listening texts were not the main focus. As for the analysis on texts of middle school, O. I. Ha and Y. Cho (2016) compared reading texts between the middle school English textbooks and the NAEA of middle school in order to analyze lexical difficulty and readability. The results revealed that the NAEA reading texts showed higher lexical diversity values than the textbooks. As for the analysis on texts of high school, H. Lee (2021) analyzed the distribution of words in high school English textbooks and the 2019 NAEA of high school in order to identify how vocabulary items of the 2015 Revised National Curriculum were realized in both listening and reading texts of textbooks and tests. The results revealed that the NAEA English texts needed to contain more difficult words since the 2019 NAEA included significantly easier words than the textbooks.

Existing studies have rarely compared academic texts between different school levels. However, the levels of tests need to be distinguished in accordance with their targets' proficiency in order to give proper feedback on their academic achievement. In addition, research on the NAEA did not analyze listening texts by comparing dialogues and monologues. Thus, the NAEA listening texts need to be analyzed by dividing into different school levels and different forms of spoken language.

III. METHOD

1. Corpus Collection

NAEA English listening texts of the five consecutive academic years from 2015 to 2019 were collected. In this period, the NAEA was implemented through the paper-based test. The NAEA aims at the third grade of middle school and the second grade of high school. The corpus consists of 169 NAEA listening texts in total. The 169 texts include 90 listening texts for middle school students and 79 listening texts for high school students. The corpus is divided into the two forms of spoken language: dialogues ($n = 108$) and monologues ($n = 61$). Table 1 suggests the classification of NAEA listening texts in accordance with the academic years, the school levels, and the two forms of spoken language.

Except for 2019 NAEA listening texts for high school

TABLE 1
The Classification of NAEA Listening Texts

Language form	School level	Year	<i>n</i>	Item No.
Dialogue (<i>n</i> = 108)	Middle (<i>n</i> = 51)	2015	11	1, 2, 3, 6, 7, 8, 9, 12, 13, short answer Q1 and Q3
		2016	10	1, 2, 3, 6, 7, 8, 9, 12, 13 and short answer Q1
		2017	10	
		2018	10	1, 2, 3, 5, 7, 8, 9, 12, 13 and short answer Q1
		2019	10	
	High (<i>n</i> = 57)	2015	13	1, 2, 3, 5, 7, 8, 9, 10, 12, 14, 15, short answer Q1 and Q3
		2016	13	
		2017	13	
		2018	13	
		2019	5	
Mono- logue (<i>n</i> = 61)	Middle (<i>n</i> = 39)	2015	7	4, 5, 10, 11, 14, 15 and short answer Q2
		2016	8	4, 5, 10, 11, 14, 15, short answer Q2 and Q3
		2017	8	
		2018	8	4, 6, 10, 11, 14, 15, short answer Q2 and Q3
		2019	8	
	High (<i>n</i> = 22)	2015	5	4, 6, 11, 13 and short answer Q2
		2016	5	
		2017	5	
		2018	5	
		2019	2	3 and 4

students, all the academic years include 18 listening texts. Each test consists of 15 multiple-choice items and three constructed-response items. Only the 2019 NAEA of high school has seven listening texts. The 2019 NAEA of high school includes five multiple-choice items and two constructed-response items. In this study, regardless of the item types, all the listening scripts were extracted. The words M (man) and W (woman) were not included for analysis although the listening scripts contain speakers' information.

2. Coh-Metrix

The current study adopts the online version of Coh-Metrix 3.0 (<http://141.225.61.35/cohetrix2017>) as the tool to assess the linguistic aspects of the 2015-2019 NAEA listening texts. Among 106 Coh-Metrix indices, this study analyzes 16 indices by dividing into word, sentence and text levels. Table 2 displays the selected Coh-Metrix indices according to word, sentence and text levels. These 16 linguistic aspects can show the level of text difficulty.

1) Word Level

In order to analyze the linguistic aspects of NAEA listening texts at the word level, this study uses two descriptive indices, one lexical-diversity index and three word-information indices. Descriptive measures indicate information about words, sentences and paragraphs in the

corpus. Among eleven indices of descriptive measures, word counts and word length, which are related to the word level, are analyzed by calculating the number of words, and the number of syllables per word. High values of descriptive measures are highly correlated to high values of text difficulty (M. Jeon, 2011).

Coh-Metrix estimates lexical diversity with the type-to-token ratio for all words and content words. Among the two indices of the type-token ratio, the type-token ratio for content words is adopted in the current study. High values of the type-token ratio suggest that the given corpus consists of various words (McNamara et al., 2014).

Coh-Metrix provides word information including word frequency, age of acquisition and word familiarity. These indices suggest the level of lexical difficulty in the corpus. Word frequency for content words indicates the amount of high frequency words based on the CELEX database from the Dutch Center for Lexical Information (McNamara et al., 2014). High values of word frequency indicate that the given corpus contains easier words. The age-of-acquisition index suggests the age when children can acquire the words in the given corpus based on the Medical Research Council (MRC) Psycholinguistics database (McNamara et al., 2014). The word-familiarity index indicates how familiar the words in the given corpus are for adult speakers based on the MRC Psycholinguistics database (McNamara et al., 2014).

2) Sentence Level

In order to analyze the linguistic aspects of NAEA listening texts at the sentence level, the current study utilizes two descriptive indices and two syntactic complexity indices. Among eleven indices of descriptive measures, sentence counts and sentence length, which are related to basic information about the patterns of sentences in the corpus, are analyzed by calculating the number of sentences, and the number of words per sentence. High values of descriptive measures are highly correlated to high values of text difficulty (M. Jeon, 2011).

This study uses two indices of syntactic complexity: left embeddedness and modifiers per NP. Left embeddedness is estimated by calculating the average number of words before the main verb. High values of left embeddedness denote complex sentence structures in the given text. In addition, more numbers of modifiers per NP indicate more complex sentence structures in the given text. The high value from syntactic complexity is related to difficult texts (McNamara et al., 2014).

3) Text Level

In order to analyze the linguistic aspects of NAEA listening texts at the text level, the current study adopts two readability indices and four referential-cohesion indices. Readability indices identify information about text complexity (McNamara et al., 2014). Coh-Metrix provides three readability indices such as the Flesch Reading Ease (FRE),

TABLE 2

Selected Coh-Metrix Indices

Level	Category	Coh-Metrix index
Word level	Descriptive measures	Word count
		Word length
	Lexical diversity	Type-token ratio
		Word frequency
	Word information	Age of acquisition
		Word familiarity
Sentence level	Descriptive measures	Sentence count
		Sentence length
	Syntactic complexity	Left embeddedness
		Modifiers per NP
Text level	Readability	FRE
		FKGL
	Referential cohesion	Noun overlap
		Argument overlap
		Stem overlap
	Content word overlap	

the Flesch-Kincaid Grade Level (FKGL) and Coh-Metrix L2 readability. Among the three indices, the present study uses the FRE and the FKGL as the two indices are based on the traditional approach. High values of the FRE denote easy texts, whereas high values of the FKGL denote difficult texts (McNamara et al., 2014). Specifically, the FRE scores can be divided into seven levels such as the 'very difficult' level (0-30), the 'difficult' level (30-50), the 'fairly difficult' level (50-60), the 'standard' level (60-70), the 'fairly easy' level (70-80), the 'easy' level (80-90), and the 'very easy' level (90-100) (Schmitt & Prestigiacomo, 2013).

Referential cohesion estimates overlapped linguistic elements such as nouns, arguments, content words, and stems within adjacent sentences or all sentences. This study analyzes how overlapped same nouns, arguments, content words and stems are within adjacent sentences. High values indicate easy texts in noun overlap, argument overlap, content word overlap, and stem overlap as same linguistic elements more repeatedly appear in the give text (Crossley & McNamara, 2011; McNamara et al., 2014).

3. Data Coding and Statistical Analysis

After the corpus was collected, the results of Coh-Metrix measures were coded. In order to analyze the results of Coh-Metrix measures, the statistical program, Jamovi 2.2.5, was utilized. For the comparison between the two school levels, and the two different forms of spoken language, an independent two sample *t*-test was used.

IV. RESULTS AND DISCUSSION

1. Differences of Linguistic Aspects in NAEA English Listening Texts Between School Levels

Research question 1 concerns about the linguistic differences of NAEA listening texts for middle and high school students in accordance with word, sentence and text levels. An independent *t*-test is utilized for analyzing differences of 169 NAEA listening texts between the two school levels.

1) Word Level

Table 3 presents the results from the linguistic aspects of 169 NAEA listening texts at the word level with an independent two sample *t*-test including two descriptive indices, one lexical-diversity index and three word-information indices. In terms of descriptive measures, both of the NAEA dialogues and monologues of high school include significantly more words than those of middle school, like the overall forms of listening texts. However, the number of syllables per word in the NAEA dialogues and monologues does not show significant differences between middle school and high school, like in the overall forms. The results are partly different from those of the previous research on NAEA English reading passages (J. Lee, 2021a) in that the NAEA reading passages of high school included significantly more words,

and significantly longer words. Like the research on listening texts of middle school textbooks (J. Kim & M. Jeon, 2013; J. Ryu & M. Jeon, 2020b), the NAEA adjusts the number of words for their different targets. Manipulating numbers of words is an explicit way to change text difficulty.

In terms of lexical diversity, the NAEA listening texts for middle and high school students are not significantly different in both dialogues and monologues, like in the overall forms of listening texts. The results of lexical diversity indicate that the NAEA listening texts of high school do not contain significantly more unique words than those of middle school. The result is similar with the research on NAEA English reading passages (J. Lee, 2021a) in that the NAEA reading passages for middle and high school students did not show significantly different lexical-diversity values. Like the current study, the studies on listening texts of middle school textbooks (J. Kim & M. Jeon, 2013; J. Ryu & M. Jeon, 2020b) revealed that lexical diversity was not used to discriminate listening texts based on students' grades.

In terms of word information, the NAEA dialogues of middle school and high school do not show significant differences in word frequency, age of acquisition and word familiarity, like the overall forms of listening texts. The results indicate that the NAEA dialogues of high school contain similar levels of words with those of middle school. In contrast, monologues show that the NAEA listening texts of high school consist of significantly less high frequency words and less familiar words than those of middle school. The results of monologues are similar with the research on NAEA English reading passages (J. Lee, 2021a) in that the NAEA reading passages of high school included significantly less high frequency words and less familiar words than those of middle school.

2) Sentence Level

Table 4 demonstrates the results from the linguistic aspects of 169 NAEA listening texts at the sentence level with an independent two sample *t*-test including the two descriptive indices and the two syntactic-complexity indices. In terms of descriptive measures, the NAEA listening texts of high school include significantly more sentences than those of middle school. However, sentence length does not show significant differences between the two school levels. The results are partly different from those of the research on NAEA English reading passages (J. Lee, 2021a) in that the NAEA reading passages of high school included significantly more sentences, and significantly longer sentences. Like the research on listening texts of middle school textbooks (J. Kim & M. Jeon, 2013; J. Ryu & M. Jeon, 2020b), the NAEA adjusts the number of sentences for their different targets. Manipulating numbers of sentences is an explicit way to change text difficulty.

In terms of syntactic complexity, differences between the two school levels are not found in left embeddedness and modifiers per NP although the NAEA listening texts of high school are supposed to include more complex sentence structures than those of middle school. The results

correspond to the analysis on the NAEA reading passages (J. Lee, 2021a) in that the NAEA reading passages of high school did not contain significantly more complex sentence structures than those of middle school. Sentence structures in the NAEA need to be adjusted to their targets' proficiency in that complex sentence structures are hard to process (McNamara et al., 2014).

3) Text Level

Table 5 displays the results from the linguistic aspects of 169 NAEA listening texts at the text level with an independent two sample *t*-test including the two readability indices and the four referential-cohesion indices. At the text level, dialogues and monologues show the same patterns of the statistical significance in all six Coh-Metrix indices, like the overall forms of listening texts.

In terms of the readability, the FRE and the FKGL do not show upward trends in accordance with the school level. Differences were not found in the FRE and the FKGL. The results indicate that the NAEA listening texts of high

school were not significantly more difficult than those of middle school. These results are totally different from the previous research on NAEA English reading passages (J. Lee, 2021a) in that the NAEA reading passages of high school were significantly more difficult than those of middle school. In addition, research on listening texts of middle school textbooks also indicated that different grade levels had significantly different readability levels (J. Kim & M. Jeon, 2013; J. Ryu & M. Jeon, 2020b). The levels of difficulty are supposed to be significantly different according to the school levels like the previous research on NAEA English reading passages. However, the current study finds that the NAEA listening texts do not have proper text difficulty for their targets.

In terms of referential cohesion, dialogues and monologues show the same levels of repetition in same nouns, arguments, content words and stems within adjacent words between the two school levels. As the overlapping linguistic elements could lower text difficulty (McNamara et al., 2014), the tests between the two school levels have similar levels of text difficulty.

TABLE 3
T-test for the Word Level

Language form	Category	Index	Middle school <i>M (SD)</i>	High school <i>M (SD)</i>	<i>t</i>	<i>df</i>	<i>p</i>
Overall forms (<i>n</i> = 169)	Descriptive measures	Word count	81.50 (10.83)	101.17 (12.37)	-11.02	167	.000***
		Word length	.62 (.13)	.61 (.12)	.47	167	.637
	Lexical diversity	Type-token ratio	.85 (.06)	.85 (.05)	.08	167	.934
		Word frequency	2.40 (.16)	2.36 (.15)	1.44	167	.151
	Word information	Age of acquisition	299.66 (41.39)	302.19 (41.16)	-.40	167	.691
		Word familiarity	586.32 (6.83)	584.36 (7.11)	1.83	167	.070
Dialogues (<i>n</i> = 108)	Descriptive measures	Word count	83.86 (11.79)	102.96 (12.89)	-8.00	106	.000***
		Word length	1.28 (.08)	1.29 (.08)	-.89	106	.378
	Lexical diversity	Type-token ratio	.86 (.06)	.85 (.05)	.68	106	.501
		Word frequency	2.36 (.13)	2.38 (.14)	-.49	106	.626
	Word information	Age of acquisition	289.13 (42.00)	295.97 (41.40)	-.85	106	.396
		Word familiarity	585.83 (7.43)	584.90 (7.52)	.64	106	.521
Monologues (<i>n</i> = 61)	Descriptive measures	Word count	78.41 (8.64)	96.50 (9.69)	-7.51	59	.000***
		Word length	1.36 (.11)	1.38 (.13)	-.57	59	.571
	Lexical diversity	Type-token ratio	.85 (.05)	.85 (.06)	-.64	59	.524
		Word frequency	2.44 (.18)	2.33 (.16)	2.39	59	.020*
	Word information	Age of acquisition	313.44 (36.70)	318.30 (36.68)	-.50	59	.621
		Word familiarity	586.98 (5.99)	582.98 (5.87)	2.52	59	.015*

p* < .05, **p* < .001

TABLE 4
T-test for the Sentence Level

Language form	Category	Index	Middle school <i>M (SD)</i>	High school <i>M (SD)</i>	<i>t</i>	<i>df</i>	<i>p</i>
Overall forms (<i>n</i> = 169)	Descriptive measures	Sentence count	10.73 (3.58)	13.94 (3.49)	-5.87	167	.000***
		Sentence length	8.38 (2.74)	7.72 (2.09)	1.75	167	.083
	Syntactic complexity	Left embeddedness	1.64 (1.09)	1.44 (.77)	1.33	167	.185
		Modifiers per NP	.58 (.19)	.56 (.18)	.63	167	.532
Dialogues (<i>n</i> = 108)	Descriptive measures	Sentence count	13.29 (2.36)	15.58 (2.10)	-5.32	106	.000***
		Sentence length	6.40 (.88)	6.67 (.82)	-1.66	106	.099
	Syntactic complexity	Left embeddedness	1.03 (.48)	1.15 (.46)	-1.42	106	.159
		Modifiers per NP	.53 (.14)	.51 (.15)	.88	106	.380
Monologues (<i>n</i> = 61)	Descriptive measures	Sentence count	7.39 (1.52)	9.68 (2.70)	-4.27	59	.000***
		Sentence length	10.96 (2.13)	10.41 (1.95)	1.00	59	.323
	Syntactic complexity	Left embeddedness	2.43 (1.16)	2.18 (.92)	.89	59	.376
		Modifiers per NP	.65 (.23)	.71 (.16)	-1.17	59	.247

*** *p* < .001

TABLE 5
T-test for the Text Level

Language form	Category	Index	Middle school <i>M (SD)</i>	High school <i>M (SD)</i>	<i>t</i>	<i>df</i>	<i>p</i>
Overall forms (<i>n</i> = 169)	Readability	FRE	86.99 (9.85)	87.49 (9.19)	-.34	167	.733
		FKGL	3.19 (1.92)	2.95 (1.65)	.87	167	.388
	Referential cohesion	Noun overlap	.13 (.16)	.10 (.13)	1.46	167	.147
		Argument overlap	.36 (.26)	.29 (.19)	1.93	167	.056
		Stem overlap	.17 (.18)	.13 (.14)	1.60	167	.111
		Content word overlap	.09 (.06)	.08 (.05)	1.08	167	.280
Dialogues (<i>n</i> = 108)	Readability	FRE	91.89 (6.66)	90.51 (6.24)	1.11	106	.270
		FKGL	2.00 (1.08)	2.26 (1.00)	-1.29	106	.199
	Referential cohesion	Noun overlap	.06 (.07)	.05 (.06)	.52	106	.602
		Argument overlap	.22 (.12)	.24 (.14)	-.71	106	.481
		Stem overlap	.07 (.08)	.07 (.07)	.06	106	.949
		Content word overlap	.06 (.04)	.07 (.04)	-.89	106	.377
Monologues (<i>n</i> = 61)	Readability	FRE	80.58 (9.72)	79.66 (10.98)	.34	59	.738
		FKGL	4.74 (1.66)	4.73 (1.67)	.02	59	.985
	Referential cohesion	Noun overlap	.22 (.19)	.22 (.18)	.16	59	.870
		Argument overlap	.54 (.28)	.43 (.22)	1.64	59	.107
		Stem overlap	.29 (.20)	.27 (.18)	.41	59	.683
		Content word overlap	.12 (.07)	.10 (.06)	1.06	59	.292

2. Differences of Linguistic Aspects in NAEA English Listening Texts Between Language Forms

Research question 2 focuses on the linguistic differences between dialogues and monologues at word, sentence and text levels. An independent *t*-test is conducted for analyzing differences of 169 NAEA listening texts between the two forms of spoken language.

1) Word Level

Table 6 remarks the results from the linguistic aspects of 169 NAEA listening texts at the word level with an independent two sample *t*-test including two descriptive indices, one lexical-diversity index and three word-information indices. In terms of descriptive measures, the NAEA listening texts of the two school levels show that dialogues contain significantly more words than monologues, like all texts. In addition, the two school levels show that monologues include significantly more syllables per word than dialogues, like all texts. The results are partly different from the previous research on the CSAT (J. Lee, 2021b) in that the CSAT monologues contained more words and longer words than the CSAT dialogues.

In terms of lexical diversity, the NAEA dialogues and monologues are not significantly different in both middle school and high school. The results indicate that dialogues and monologues in the NAEA listening texts contain similar numbers of unique words. The results are totally different from the existing study on the CSAT (J. Kim & H. Kim, 2021) in that the CSAT monologues contained significantly more diverse words than the CSAT dialogues.

In terms of word information, the NAEA listening texts of the middle school and high school have the same patterns of the statistical significance in age of acquisition and word familiarity, but not in word frequency. Only the NAEA listening texts of middle school show that monologues contain significantly more high frequency words than dialogues, whereas the NAEA listening texts of high school and all texts do not. The NAEA monologues show significantly higher levels of the age-of-acquisition index than the NAEA dialogues in both middle school and high school, like in all texts. The levels of word familiarity in dialogues and monologues are not significantly different. The results of word frequency are different from the previous research on the CSAT (J. Kim & H. Kim, 2021) in that the CSAT monologues included significantly less high frequency words than the CSAT dialogues.

TABLE 6
T-test for the Word Level

School level	Category	Index	Dialogues <i>M</i> (<i>SD</i>)	Monologues <i>M</i> (<i>SD</i>)	<i>t</i>	<i>df</i>	<i>p</i>
All (<i>n</i> = 169)	Descriptive measures	Word count	93.94 (15.61)	84.93 (12.53)	3.86	167	.000***
		Word length	1.29 (.08)	1.37 (.11)	-5.46	167	.000***
	Lexical diversity	Type-token ratio	.85 (.06)	.85 (.06)	.48	167	.631
		Word frequency	2.37 (.14)	2.40 (.18)	-1.27	167	.204
	Word information	Age of acquisition	292.74 (41.63)	315.19 (36.46)	-3.52	167	.000***
		Word familiarity	585.34 (7.46)	585.54 (6.20)	-.18	167	.859
Middle school (<i>n</i> = 90)	Descriptive measures	Word count	83.86 (11.79)	78.41 (8.64)	2.43	88	.017*
		Word length	1.28 (.08)	1.36 (.11)	-4.18	88	.000***
	Lexical diversity	Type-token ratio	.86 (.06)	.85 (.05)	.94	88	.349
		Word frequency	2.36 (.13)	2.44 (.18)	-2.38	88	.019*
	Word information	Age of acquisition	289.13 (42.00)	313.44 (36.70)	-2.87	88	.005**
		Word familiarity	585.83 (7.43)	586.98 (5.99)	-.79	88	.432
High school (<i>n</i> = 79)	Descriptive measures	Word count	102.96 (12.89)	96.50 (9.69)	2.13	77	.037*
		Word length	1.29 (.08)	1.38 (.13)	-3.65	77	.000***
	Lexical diversity	Type-token ratio	.85 (.05)	.85 (.06)	-.40	77	.693
		Word frequency	2.38 (.14)	2.33 (.16)	1.26	77	.213
	Word information	Age of acquisition	295.97 (41.40)	318.30 (36.68)	-2.22	77	.030*
		Word familiarity	584.90 (7.52)	582.98 (5.87)	1.07	77	.287

p* < .05, *p* < .01, ****p* < .001

TABLE 7

T-test for the Sentence Level

School level	Category	Index	Dialogues M (SD)	Monologues M (SD)	t	df	p
All (n = 169)	Descriptive measures	Sentence count	14.50 (2.50)	8.21 (2.29)	16.19	167	.000***
		Sentence length	6.55 (.86)	10.76 (2.06)	-18.60	167	.000***
	Syntactic complexity	Left embeddedness	1.09 (.47)	2.34 (1.07)	-10.41	167	.000***
		Modifiers per NP	.52 (.14)	.67 (.21)	-5.72	167	.000***
Middle school (n = 90)	Descriptive measures	Sentence count	13.29 (2.36)	7.39 (1.52)	13.63	88	.000***
		Sentence length	6.40 (.88)	10.96 (2.13)	-13.85	88	.000***
	Syntactic complexity	Left embeddedness	1.03 (.48)	2.43 (1.16)	-7.85	88	.000***
		Modifiers per NP	.53 (.14)	.65 (.23)	-3.05	88	.003**
High school (n = 79)	Descriptive measures	Sentence count	15.58 (2.10)	9.68 (2.70)	10.30	77	.000***
		Sentence length	6.67 (.82)	10.41 (1.95)	-12.04	77	.000***
	Syntactic complexity	Left embeddedness	1.15 (.46)	2.18 (.92)	-6.58	77	.000***
		Modifiers per NP	.51 (.15)	.71 (.16)	-5.48	77	.000***

p < .01, *p < .001

TABLE 8

T-test for the Text Level

School level	Category	Index	Dialogues M (SD)	Monologues M (SD)	t	df	p
All (n = 169)	Readability	FRE	91.16 (6.45)	80.25 (10.11)	8.56	167	.000***
		FKGL	2.14 (1.04)	4.74 (1.65)	-12.57	167	.000***
	Referential cohesion	Noun overlap	.05 (.07)	.22 (.19)	-8.46	167	.000***
		Argument overlap	.23 (.13)	.50 (.26)	-8.92	167	.000***
		Stem overlap	.07 (.07)	.28 (.19)	-10.24	167	.000***
		Content word overlap	.07 (.04)	.11 (.07)	-5.44	167	.000***
Middle school (n = 90)	Readability	FRE	91.89 (6.66)	80.58 (9.72)	6.54	88	.000***
		FKGL	2.00 (1.08)	4.74 (1.66)	-9.49	88	.000***
	Referential cohesion	Noun overlap	.06 (.07)	.22 (.19)	-5.71	88	.000***
		Argument overlap	.22 (.12)	.54 (.28)	-7.40	88	.000***
		Stem overlap	.07 (.08)	.29 (.20)	-7.21	88	.000***
		Content word overlap	.06 (.04)	.12 (.07)	-4.63	88	.000***
High school (n = 79)	Readability	FRE	90.51 (6.24)	79.66 (10.98)	5.52	77	.000***
		FKGL	2.26 (1.00)	4.73 (1.67)	-8.07	77	.000***
	Referential cohesion	Noun overlap	.05 (.06)	.22 (.18)	-6.13	77	.000***
		Argument overlap	.24 (.14)	.43 (.22)	-4.54	77	.000***
		Stem overlap	.07 (.07)	.27 (.18)	-6.99	77	.000***
		Content word overlap	.07 (.04)	.10 (.06)	-2.57	77	.012*

*p < .05, ***p < .001

2) Sentence Level

Table 7 displays the results from the linguistic aspects of 169 NAEA listening texts at the sentence level with an independent two sample t-test including the two descriptive indices and the two syntactic-complexity indices. At the sentence level, middle school and high school show

the same patterns of the statistical significance in all four Coh-Metrix indices, like all texts. In terms of descriptive measures, dialogues include significantly more sentences than monologues, whereas monologues contain significantly longer sentences. The results correspond to the previous research on the CSAT (J. Lee, 2021b).

In terms of syntactic complexity, both middle school

and high school show that monologues include the significantly more number of words before the main verb, and the significantly more number of modifiers per noun phrase than dialogues, like all texts. In general, sentence structures of the NAEA monologues are significantly more complex. The results are similar with the analysis on the CSAT (J. Kim & H. Kim, 2021; J. Lee, 2021b) in that the CSAT monologues were significantly more complex than the CSAT dialogues.

3) Text Level

Table 8 illustrates the results from the linguistic aspects of 169 NAEA listening texts at the text level with an independent two sample *t*-test including the two readability indices and the four referential-cohesion indices. At the text level, middle school and high school show the same patterns of the statistical significance in all six Coh-Metrix indices, like all texts. In terms of readability, monologues are significantly more difficult than dialogues. Based on Schmitt and Prestigiacomio's (2013) study, the NAEA dialogues of middle school and high school are at the 'very easy' level, whereas monologues of middle school and high school are at the 'easy' level and at the 'fairly easy' level, respectively. The results correspond to the previous study on the CSAT (J. Lee, 2021b) in that monologues were significantly more difficult than dialogues in the CSAT. The significant differences between monologues and dialogues could come from the different characteristics. Monologues tend to be more formal and deal with more difficult topics than dialogues because monologues include lectures and announcements (Brown, 2007; H. Jung, 2010).

Monologues are significantly more referentially cohesive than dialogues. More specifically, the NAEA monologues include more numbers of overlapped nouns, arguments, stems and content words than the NAEA dialogues. The results are similar with the analysis on the CSAT (J. Kim & H. Kim, 2021; J. Lee, 2021b) in that the CSAT monologues had significantly higher levels of referential cohesion than the CSAT dialogues.

V. CONCLUSION

The current study aims to investigate linguistic differences of NAEA listening texts between middle school and high school, and between the two forms of spoken language in order to identify whether the text difficulty of dialogues and monologues increases as the school level increases. Few studies analyzed the interconnection of differences between the two different school levels and between the two forms of spoken language. Thus, the current study was needed.

For this linguistic analysis, this study collected the 2015-2019 NAEA listening texts which were conducted just before the system of the NAEA changed from the paper-based testing into the computer-based testing. The linguistic aspects of the collected corpus were analyzed at

word, sentence and text levels. This study used descriptive measures related to words, lexical diversity and word information for word-level analysis, descriptive measures related to sentences and syntactic complexity for sentence-level analysis, and readability and referential cohesion for text-level analysis with Coh-Metrix.

The two research questions focused on the differences of the NAEA listening texts between middle school and high school, and the differences of the NAEA listening texts between dialogues and monologues. In terms of the comparison between middle school and high school, the listening texts of high school were not lexically, syntactically and textually more difficult than those of middle school, whereas numbers of words and sentences discriminated the two school levels properly. In addition, the NAEA monologues of high school were lexically more difficult than those of middle school unlike the NAEA dialogues. When comparing the results of the current study from those of the existing study on the NAEA reading texts (J. Lee, 2021a), the listening test of the NAEA did not discriminate the level of the text difficulty in accordance with the school levels.

In terms of the comparison between dialogues and monologues, monologues were syntactically more complex and textually more difficult than dialogues. The NAEA listening texts of the two different school levels showed the same trends of the significant differences between dialogues and monologues except for word frequency. Monologues contained significantly more high frequency words only in the NAEA listening texts of middle school. The result of word frequency in high school was different from that of the previous research on the CSAT listening texts (J. Kim & H. Kim, 2021).

The present study could give the educational implications related to future NAEA English listening texts. In order to increase the quality of the future NAEA English listening, both NAEA dialogues and monologues need to adjust the level of difficulty as the paper-based NAEA English listening texts were not distinguished in the word, sentence and text levels according to the school levels except for word counts and sentence counts. At the word level, the NAEA dialogues of high school need to contain more difficult words than those of middle school. At the sentence level, the NAEA listening texts of high school need to include more complex sentence structures than those of middle school. At the text level, the NAEA listening texts of high school have to be more difficult than those of middle school.

This study could have the limitation related to the collection of the corpus. The current study analyzed the paper-based NAEA listening texts although the computer-based test started to be implemented. For future studies on the NAEA, the computer-based NAEA English listening texts could be analyzed by classifying into different school levels and different language forms. Future studies can compare the paper-based and computer-based testing systems in the NAEA.

REFERENCES

- Brown, H. D. (2007). *Teaching by principles: An interactive approach to language pedagogy* (3rd ed.). Longman.
- Crossley, S. A., & McNamara, D. S. (2011). Understanding expert ratings of essay quality: Coh-Metrix analyses of first and second language writing. *International Journal of Continuing Engineering Education and Life Long Learning*, 21(2-3), 170-191.
- Ha, Ok I, & Cho, Yunkyong. (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.
- Jeon, Moongee. (2011). A corpus-based analysis of continuity of the reading materials in middle school English 1 and 2 textbooks with Coh-Metrix. *The Journal of Linguistic Science*, 56(1), 201-218.
- Jeon, Moongee. (2015). An analysis of the continuity among middle school English textbooks with an automated language analysis program. *Modern English Education*, 16(1), 195-218.
- Jung, Haeng. (2010). An analysis of spoken English test items in Korean CSAT. *Studies in English Education*, 15(2), 177-197.
- Kim, Jien, & Jeon, Moongee. (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, Jin-Young, & Kim, Hae-Dong. (2021). A comparison of the lexical, syntactic, and discourse features between the listening and the reading texts in the College Scholastic Ability Test. *Modern English Education*, 22(1), 45-56.
- Kim, Joonkoo. (2018). A corpus-based investigation of reading passages in the National Assessment of Educational Achievement English test using Coh-Metrix. *Secondary English Education*, 11(2), 27-51.
- Lee, Hyunoo. (2021). Distribution of graded vocabulary in the high school English textbooks and K-11 NAEA English test based on the 2015 Revised National Curriculum. *Secondary English Education*, 14(1), 22-44.
- Lee, Inho, Lee, Sang Il, Kim, Seung Hyun, Lee, Jungwoo, Seo, Min-cheol, Jo, Yun Dong, Lee, Kwang Sang, Kim, Hyun-Kyung, Dong, Hyo Kwan, Pae, Jue-Kyoung, Kim, Sung Hye, Jung, Ki Moon, Kwon, Kyung Phil, & Lee, Ho Kyu. (2015). *A study on the development of the 2015 National Assessment of Educational Achievement (NAEA)*. Korea Institute for Curriculum and Evaluation. <https://doi.org/10.23000/TRKO201800023011>
- Lee, Jiyoung. (2021a). A comparative analysis of NAEA English reading passages for middle and high school students using Coh-Metrix. *Modern English Education*, 22(3), 12-23.
- Lee, Jiyoung. (2021b). A comparison of text difficulty in 2014-2021 CSAT English listening materials between testing types and between language forms. *Journal of the Korea English Education Society*, 20(4), 65-86.
- Lee, Seung-Hwan. (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.
- McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.
- Papageorgiou, S., Stevens, R., & Goodwin, S. (2012). The relative difficulty of dialogic and monologic input in a second-language listening comprehension test. *Language Assessment Quarterly*, 9(4), 375-397.
- Ryu, Jisu, & Jeon, Moongee. (2020a). An analysis of text difficulty across grades in Korean middle school English textbooks using Coh-Metrix. *The Journal of Asia TEFL*, 17(3), 921-936.
- Ryu, Jisu, & Jeon, Moongee. (2020b). 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.
- Schmitt, P. J., & Prestigiacomo, C. J. (2013). Readability of neurosurgery-related patient education materials provided by the American Association of Neurological Surgeons and the National Library of Medicine and National Institutes of Health. *World Neurosurgery*, 80(5), e33-e39.