Exploring linguistic complexity of English First Additional Language Grade 12 and Grade 11 exam texts

Sibeko, Johannes
Nelson Mandela University
Johannes.sibeko@mandela.ac.za

Abstract

This article investigates consistency and levelness of texts administered for English First Additional Language exams in grade 11 and grade 12. The analysis is based on a data set comprising 80 texts extracted from 20 question papers from the year 2011 to the year 2020. The exam texts are separated into reading comprehension, summary writing and language and editing texts. A total of 18 indices were analysed. Aspects covering lexical diversity, syntactical complexity, text readability, the use of connectives and surface level complexity were analysed. The findings indicate that grade 12 exam texts are consistently more linguistically complex than grade 11 exam texts. Grade 12 exam texts are also found to be more lexically diverse and syntactically complex than grade 11 exam texts. Furthermore, grade 12 exam texts are harder to read than grade 11 exam texts. For both grades, summary writing texts are easier to read.

Keywords: Linguistic complexity; English first additional language; Lexical diversity; Text readability

1 Introduction

Desirably, texts used for language teaching and assessment should increase in complexity as grades advance. South African school grades are separated into four phases, namely, (i) the Foundation phase (composing pre-school to grade 3), (ii) the Intermediate phase (grade 4 to 6), (iii) the Senior phase (grade 7 to 9), and (iv) the Further Education and Training phase (grade 10 to 12) (Molepo, 2014; Moodley, 2013).

English, like all South African official language subjects, is taught at three different levels. First, the home language (EHL) is designed for learners that start school with English linguistic competencies (DBE, 2011a). Second, the First Additional Language (EFAL) curriculum is designed for learners who start school with minimal linguistic competency in English (2011b). Most learners in South African schools are enrolled for EFAL (Kaiser et al., 2010:53; Mophosho et al., 2019:60). Third, the Second Additional Language (ESAL) curriculum is designed for learners that start school with no prior exposure to, or very limited competencies in English (DBE, 2011c). By the end of high school, learners in both EHL and EFAL are expected to have mastery of English and to be able to use their proficiency in workplaces and higher education (DBE, 2011a:8-9; DBE, 2011b:8-9). In addition to other efforts, teachers are tasked with exposing learners to different text types (Zano, 2022:2013), the lengths of which are prescribed by respective Curriculum and Assessment Policy Statements (Sibeko, 2021:54).

English is esteemed in South African basic education (Zano and Phatudi, 2019:16). However, learners doing EFAL in different grades are not performing well (Khosa, 2019; DBE 2013; DBE 2014; Zano, 2020; Molotja and Themane, 2018).

Exam formats in lower grades tend to follow those of the school exit exams (Van der Walt (2018). As such, exam paper formats from the 10th grade are likened to the grade 12 exam. Therefore, if grade 11 EFAL exam texts are selected through the same process as the grade 12 exam texts, similarities should be observed in the exam texts for the two grades. However, an increment in linguistic complexity should also be expected. The aims of this article are two-fold. First, similarities between the grade 11 and grade 12 exam texts are explored. Second, the differences in linguistic complexity levels between the grade 11 and grade 12 exam texts are highlighted.

2 Literature review

Although linguistic complexity is a contemporary science topic (Mitchell, 2009), an agreed upon definition of the concept is yet to be coined (Bulté and Housen, 2012:22). Linguistic complexity considers specific properties of language forms that are used in the production of language and their levels of sophistication (Ortega, 2003:492; Ai & Lu, 2013:249). Evaluations of linguistic complexity include amongst others calculating

averages of lengths and frequencies of sentences, clauses, utterances, subordinations and other features (Ortega, 2012:117).
It is generally accepted that there are two main branches in linguistic complexity, namely, the (i) absolute and the (ii) relative complexity. Relative complexity is characterised by its focus on processing and learning difficulties (Bulté & Housen, 2012; Dahl, 2009; Miestamo, 2008, Miestamo, 2009). In this study, linguistic complexity is approached from a quantitative point of view (De Clerq and Housen, 2017:316). Lexical, syntactical and part of speech measures are discussed below.

2.1 Lexical diversity

The Coh-Metrix evaluates four lexical diversity measures, namely, the type-token ratio for content words and for all words, the MTLD and the *Vocd*. Token-Type Ratio (TTR) is calculated through counting the total token types and dividing them by the total number of tokens (Jarvis, 2002; Covington and McFall, 2008). Since TTR is easy to calculate, more complex computational metrics are used significantly less (Cvrček and Chlumská, 2015:315). Even so, there is scholarship highlighting the TTR method's shortcomings. Even so, Cvrček and Chlumská, (2015:315) identify two major flaws of the TTR namely (i) the TTR is extremely sensitive to text length making it unusable when comparing texts of significant varied lengths and (ii) the likelihood of texts to stop introducing new token types as the text lengthens as opposed to when it starts, thereby resulting in longer texts indicating lower TTR. In other words, shorter texts will indicate a higher linguistic diversity while longer texts indicate lower linguistic diversity (Covington and McFall, 2008). Subsequently, new methods have been developed to augment these shortcomings (Jarvis, 2002:57; Cvrček and Chlumská, 2015:315).

One such method is the D measure (Malvern et al., 2004; Richards and Malvern 1997; MacWhinney 2000). The D measure is based on the predicted decline of the TTR as the text lengthens (Johansson, 2009:64). It is calculated using the *vocd* computer program which can calculate linguistic diversity through either random sampling or via sequential sampling (McCarthy and Jarvis, 2007:461). The random sampling method sets *vocd* apart and grants it validity over other methods of measuring linguistic diversity (Malvern et al. 2004). The *vocd* method involves taking 100 sample sets of 35-tokens each from the whole text and calculating the TTR of each set and averaging the 100 outcomes. Incrementally, one more token is added, and the process is repeated. This process is continued until each sample consists of 50 tokens (McCarthy and Jarvis, 2007:464). Finally, the mean of these values determines the overall linguistic diversity. This process is repeated three times (Malvern et al., 2004:57; McCarthy and Jarvis, 2007:464). The outcomes of the three processes is averaged to arrive at the final text's linguistic variation (Malvern et al., 2004:57; McCarthy and Jarvis, 2007:464). In this way, the factor of text length is countered, and texts of different lengths can be evaluated. Even so, the *vocd* has been challenged for its sensitivity text length (Daller et al., 2003; McCarthy and Jarvis, 2007).

Another method designed to counter the effects of text length on lexical diversity is the measure of textual lexical diversity (MTLD). According to McCarthy (2005), the MTLD is a reliable lexical diversity measure as it does not correlate with text length. Even so, it works better when used together with other sophisticated measures of lexical diversity (McCarthy, 2005).

2.2 Syntactic complexity

Absolute complexity is measured through syntactic complexity (Bulté and Housen 2012). Syntactic complexity refers to "the range and degree of sophistication of syntactic structures that surface in language production," (Ai and Lu, 2013:249). According to Kyle (2016:08) syntactic complexity refers to the formal properties of a text. There are no guidelines for the number of syntactic measures that should be investigated. In fact, they can be as few as one (Lu 2011:38; Lu and Ai, 2015:19; Wolfe-Quintero et al. 1998). Sentence, clause and phrase measures are considered in this article.
Many studies regard clauses as having both a subject and a finite verb (Bulté and Housen, 2012;

Bulté 2013:86). According to Bulté and Housen (2012:40) this definition entails the exclusion of non-finite clauses and subject ellipsis clauses which should also be included in the calculation of clausal complexity. To this end, they suggest using measures that focus on intra-phrasal complexity like the number of dependents per phrasal head and the number of words per phrase. Syntactic structures affect working memory. Working memory load is evaluated with left embeddedness. That is, the number of words before the main verb of the main clause in a sentence (Polio and Yoon, 2018). Higher numbers of words before the main verb of the main clause make the sentence more dense and ambigous (Graessar et al., 2004; Khushik and Huhta, 2020). According to Delage and Frauenfelder (2019), longer distances result in more items needing to be stored in working memory. To this end, left embeddedness and number of modifiers per noun phrase result in higher complexity through negatively affecting the working memory.

2.3 Word information

The Coh-Metrix evaluates 22 indices for word information. These calculations are completed per 1000 words through the Charniak Parser (Charniak, 1997). The identification of these phrases depend on the correct POS classification. The Coh-metrix uses the Penn Treebank (Marcus et al., 1993) to identify over 50 POS tags (McCarthy et al., 2007:07). The Brill POS Tagger (Brill, 1995), assigns the tags to each word (McCarthy et al., 2007:07; Graessar et al., 2004:197). The Coh-Metrix can identify POS tags for new words (Graessar et al., 2004:197). Tagging new words depends on the syntactical context. During tagging, only one POS tag is attached to each word (Graessar et al., 2011:225). When more than one tag can be attached to a word, the Brill POS Tagger chooses one tag based on the syntactical context (Graessar et al., 2004:197).

3 Methodology

3.1 Material

Three exam papers are written in the EFAL final exams in the FET phase (c.f. Sibeko and van Zaanen, 2021; Sibeko 2021). Learners are examined on comprehension and language structures (first paper), literature (second paper) and creative and transactional writing (third paper) (DBE 2011a; DBE, 2011b). This article investigates the exam texts in the first paper. The data set comprises 20 exam texts composed of 10 EFAL grade 11 November exam texts and 10 EFAL grade 12 November exam texts from 2011 to 2020. Grade 12 exam question papers were downloaded from the South African DBE's online repository [1]. The grade 11 exam question papers were downloaded from the Eastern Cape Exam Resources repository [2]. The question papers were downloaded in PDF format. Texts were extracted using Python. Both the grade 12 and the grade 11 exam papers, are divided into three sections, namely, reading comprehension (section A), summary writing (section B), and (iii) language and editing (section C). For this investigation, all visual texts, instructions and questions were excluded. The question paper texts were then manually spliced into three separate documents. In the end a dataset with 80 sets of texts was created. The resultant dataset is presented in Table 1. Reading comprehension texts are labelled as Text A, summary texts are labelled as Text B, and the language convention texts are labelled as Text C. The combined exam texts are indicated as All.

3.2 Procedure

The Coh-Metrix 3.0 online tool was used to assess the data sets. The Coh-Metrix automatically tokenizes, tags the parts of speech and parses the texts during the evaluation. The outcomes are output into a text document. The Coh-metrix has improved a lot from its sole focus on cohesion markers as was the original focus in initial development (Graessar et al., 2011:224). More studies have seen an improvement and addition of metrics that can be analysed by the Coh-metrix (c.f. Graesser and McNamara, 2011; Graesser et al., 2004; McNamara et al., 2010; McNamara et al., 2012). A total of 106 indices are evaluated in the tool.

Table 1: Number of words per grade, year and text

Year	Grade 11				Grade 12			
	All	Text A	Text B	Text C	All	Text A	Text B	Text C
2011	1141	723	277	141	1254	702	288	264
2012	997	524	268	205	1317	672	287	358
2013	1141	553	237	351	1143	594	309	240
2014	1047	645	232	170	1121	641	255	225
2015	969	549	233	187	1127	683	261	183
2016	1125	699	259	167	1148	707	260	181
2017	1048	616	240	192	1164	733	254	177
2018	1003	586	234	183	1065	643	250	172
2019	1059	610	236	213	1048	614	260	174
2020	1060	610	238	212	1067	649	263	155
Total	10590	6115	2454	2021	11454	6638	2687	2129

3.3 Analyses

This article analyses 18 indices. Three descriptive indices are analysed, namely, the average sentence length, average word length measured by number of letters in a word and the average word length measures by mean number of syllables per word. Four lexical diversity indices are analysed, namely, the TTR for all words, TTR for content words, the MTLD, and *vocd*. One connective index is investigated. Four syntactic complexity indices are analysed, namely, the left embeddedness, number of modifiers per noun phrase, sentence syntax similarity for adjacent sentences and sentence syntax similarity for across the text.

4 Results

Density plots (Figure 1) indicate that the scores for the different indices are generally normally distributed.

4.1 Correlation Analyses

Correlations between the different index subgroups were investigated. First correlations between descriptive indexes were investigated.

Table 2: Descriptive indices features

Features		Grade		
1	2	11	12	combined
asl	asyl	0.26	0	0.07
asl	awl	0.25	0.06	0.1
asyl	awl	0.87	0.9	0.9

Table 2 presents the correlation analyses between average surface level features through average sentence length (asl), average word length measured by number of syllables (asyl), and word length based on letters per word (awl).

Although the grade 11 (r=0.26) and the combined scores (r=0.07) indicate low correlation between average sentence lengths and average word lengths by syllable counts, the grade 12 texts indicate no correlation. The asl and awl scores for the grade 11 (0.25), grade 12 (0.06) and combined scores (r=0.1) show weak and insignificant correlations at (p>.05). The asyl and awl grade 11 (r=0.87), grade 12 (r=0.9) and combined scores (r=0.9) show significant and strong correlations at (p<.001).

Second, correlations between lexical diversity measures were investigated.

Table 3: Lexical diversity correlations

Features		Grade		
1	2	11	12	combined
ttrc	ttr	0.88	0.89	0.88
ttrc	mtld	-0.25	0.23	0.05
ttrc	*vocd*	-0.56	-0.09	-0.3
mtld	ttr	-0.14	0.25	0.11
mtld	*vocd*	0.72	0.8	0.76
vocd	ttr	-0.52	-0.13	-0.3

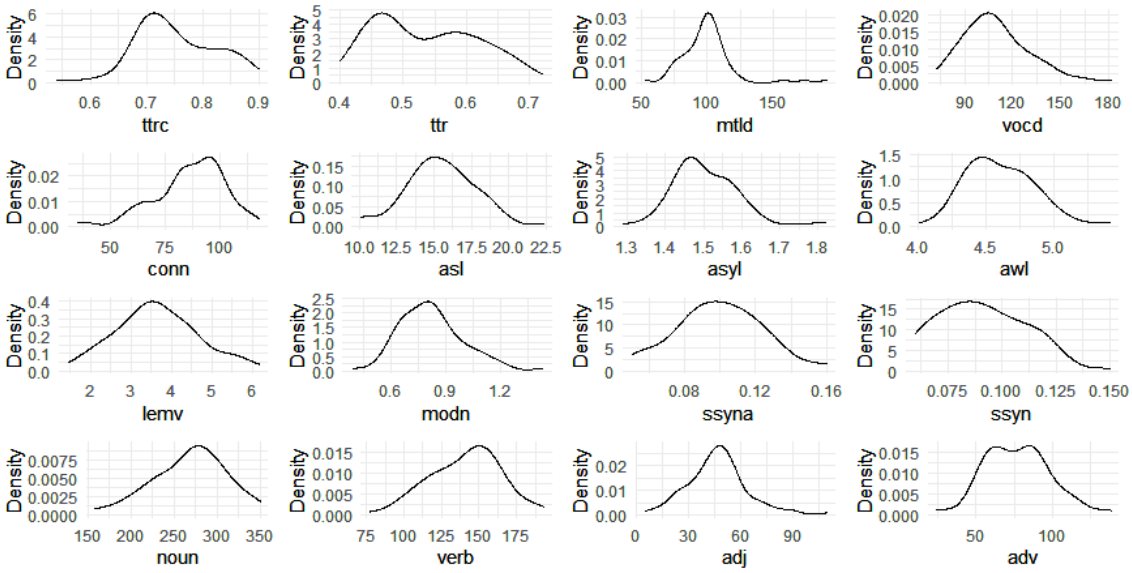

Figure 1: Density plots

Four lexical diversity indices were considered. The results are presented in Table 3.

All results for type token ratio for content word lemmas (ttrc) and type token ratio for all words (ttr) are strongly correlated. These correlations are significant at (p<.0001). The grade 11 ttrc and MTLD show negative and weak correlations (r=-0.25). Similarly, all ttrc and *vocd* results are negatively correlated. The MTLD and ttr scores across the text sets are not strongly correlated. The results do not show significant correlations (p>.05). The grade 11 texts show a strong negative correlation between the *vocd* and the ttrc (r=-0.56). The MTLD and *vocd* scores across the three text sets show significant strong positive correlations at (p<.0001). The *vocd* and ttr results for grade 12 (r=-0.13) and combined scores (r=-0.3) are negatively correlated. The grade 11 and the combined scores indicate significant correlations at (p<.005).

Table 4: Syntactic complexity correlations

Features		Grade		
1	2	11	12	combined
lemv	modn	0.38	0.15	0.26
lemv	ssyna	0.17	-0.2	0.01
lemv	ssyn	0.01	-0.13	-0.02
modn	ssyna	0	0.24	0.2
modn	ssyn	-0.02	0.24	0.2
ssyna	ssyn	0.61	0.85	0.76

The *Vocd* scores indicate that the grade 11 texts (M= 129.26, SD=14.01) are less lexically diverse than the grade 12 texts (M= 130, SD=12.88). In contrast, the MTLD for the grade 11 scores (M= 104.2, SD=10.35), is lower than the grade 12 scores (M=102,9, SD=9.32). Similarly, the ttrc for the grade 11 (M=0.7, SD=0.02) is higher than the grade 12 ttrc (M=0.68, SD=0.03).

Third, correlations between syntactical complexity measures were investigated. Four syntactical complexity indices were analysed, namely, left embeddedness (lemv), modifiers per noun phrase (modn), syntax for adjacent sentences (ssyna), and syntax similarity for the whole text (ssyn). The results are presented in Table 4.

The lemv and modn scores for the three text sets are not strongly correlated. The grade 11 texts are moderately positively correlated (r=0.38). The grade 12 scores (r=0.15) and the combined scores (r=0.26) are weakly correlated. The lemv and the mean sentence syntax similarity for adjacent sentences (ssyna) show very weak correlations. The lemv and ssyna grade 12 scores are negatively correlated (r=-0.2). The grade 11 scores (r=0.17) and combined scores (r=0.01) show weak positive correlations. The lemv and ssyn scores for grade 12 (r=-0.13), and combined scores (-0.02) are negatively correlated. The lemv and ssyn for the grade 11 texts are weakly correlated (r=0.01). The

Table 5: Parts of Speech

Features		Grades		
		11	12	combined
noun	verb	-0.42	-0.42	-0.43
noun	adj	-0.58	-0.66	-0.61
noun	adv	-0.05	-0.01	0
verb	adj	0.02	0.09	0.03
verb	adv	-0.39	-0.18	-0.34
adj	adv	0.29	-0.12	0.09

modn and ssyna for the grade 11 texts show no correlation at all. Weaker correlations are observed for the modn and ssyna grade 12 (r=0.24) and combined scores (r=0.2). Equally, the modn and ssyn for the grade 12 texts (r=0.24) and combined scores (r=0.2) show a consistent and weak correlation. Furthermore, the modn and ssyn for the grade 11 texts indicate negative correlations (r=-0.02). None of these corelations are significant. Finally, the ssyna and ssyn for grade 11 (r=0.61), grade 12 (r=0.85) and combined scores (r=0.76) show significant strong and positive correlations at (p<.0001).

Fourth, correlations between parts of speech were investigated. Correlation results are presented in Table 5. All text sets indicate significant moderate and negative correlations between the nouns and verbs at (p<.05). Similarly, the nouns and adjectives show significant negative correlation at (p<.001). No significant correlations are observed for the nouns and adverbs for the grade 11 texts (r-0.05), the grade 12 texts (r-0.66). The combined scores show no correlation between nouns and adverbs. The verbs and adjectives for the grade 11 scores (r=0.02), the grade 12 scores (r=0,09), and the combined scores (r=0.03) show small correlations. Even so, the combined scores indicate significant correlations between the verbs and the adjectives at (p<.05). The verbs and adverbs for grade 11 scores (r=-0.39), the grade 12 scores (r=-0.18) and the combined scores (r=-0.34) are negatively correlated. The grade 11 and the combined scores indicate significant correlations at (p<.05). Finally, the adjectives and adverbs for the grade 11 (r=0.29) and combined scores (r=0.09) indicate a positive and low degree of correlation while the

Table 6: Connectives

Features		Grades		
		11	12	combined
conn	asyl	-0.13	-0.45	-0.31
conn	awl	-0.12	-0.4	-0.25
conn	modn	-0.14	-0.52	-0.37
conn	ssyn	-0.11	-0.4	-0.25
conn	noun	-0.23	-0.48	-0.38
conn	verb	0.03	0.44	0.26
conn	lemv	-0.19	0.12	-0.01
conn	ttr	-0.16	-0.12	-0.14
conn	adj	0.54	0.27	0.38

grade 12 scores indicate a small negative correlation (r=-0.12). None of these results show significant correlations at (p>.05).

The nouns for the grade 12 texts (M=277.39), are higher than the verbs (132.47), the adjectives (M=82.15) and the adverbs (46.72). Similarly, the grade 11 nouns (M=265.57), are higher than the verbs (148.14), the adjectives (72.65) and the adverbs (45.67).

Fifth, one connective index was investigated. The results for all connective incidences (conn) are presented in Table 6.

Negative correlations between connectives and asyl are observed for three score sets. Grade 12 scores indicate significant correlations between connectives and asyl (r=-0.45, p<.01). Combined scores show significant moderate correlation (r=-0.31, p<.01) while grade 11 texts show low correlation (r=-0.13). Grade 12 texts show a moderate negative correlation between connectives and awl (r=-0.04, p<.05), while grade 11 scores (r=0.12) and combined scores show small negative correlations (r=-0.31) at (p<.05). Grade 12 scores indicate significant moderate negative correlation between the connectives and the modn (r=-0.52, p<.001), and between connectives and nouns (r=-0.48, p<.05). Combined scores indicate significant moderate negative correlation between the connectives and the modn (r=-0.37), and between connectives and nouns (r=-0.38) at (p<.001). Grade 12 scores indicate a significant negative correlation between connectives and ssyn (r=-.04, p<.05). Grade 11

Table 7: Flesch Reading Ease

Year	11All	11A	11B	11C	12All	12A	12B	12C
2011	67.15	69.74	66.94	53.23	59.67	52.61	64.54	72.41
2012	62.35	57.87	62.34	71.97	65.84	64.96	65.09	67.65
2013	66.62	63.66	70.14	68.31	61.27	55.44	65.72	66.01
2014	69.16	70.42	62.45	72.18	55.00	49.89	46.18	78.08
2015	72.82	67.26	73.46	84.71	60.63	58.86	70.07	53.38
2016	63.99	60.39	75.65	60.21	65.76	62.99	62.80	77.75
2017	64.18	56.14	74.87	75.32	57.91	53.55	67.98	60.05
2018	66.30	66.50	69.18	61.61	57.85	56.71	67.26	44.54
2019	57.31	46.88	73.71	65.93	54.33	52.32	68.33	39.11
2020	68.75	70.78	69.49	62.04	59.47	53.84	61.94	73.51
Mean	65.86	62.96	69.82	67.55	59.77	56.12	63.99	63.25
SD	4.24	7.69	4.80	8.97	3.88	4.84	6.75	13.69

scores indicate a significant low negative correlation between connectives and ssyn (r=-0.11, p<.05). Combined scores indicate low significant low positive correlation between connectives and ssyn (r=-0.25, p<.05).

Grade 11 scores show small negative correlations between connectives and modn (r=-0.14) and between connectives and nouns. All three text scores indicate small and negative correlations between connectives and lemv and between connectives and ttr. Positive correlations are observed between connectives and verbs and between connectives and adjectives. Grade 12 connectives and verbs show significant medium correlations (r=0.44) at (p<.005). Grade 11 scores indicate significant strong correlations between the connectives and adjectives (r=0.54, p<.001). Combined scores indicate negative and small correlations between the connectives and lemv (r=-0.01). Grade 12 scores for connectives and lemv are positively correlated (r=0.12, p<.05). Combined scores indicate moderate correlation between connectives and adjectives (r=0.38, p<.001). Grade 12 scores show low correlations between connectives and adjectives (r=0.27).

Finally, correlations between the surface level features and parts of speech were also investigated. Nouns show strong positive correlations with the awl, asyl and modn at (p<.001). Nouns are weakly and positively correlated with the lemv (r=0.24, p<.05), and negatively correlated with connectives (r=-0.38, p<.001). Verbs show significant negative correlations to asyl (r=0.32), awl (r=-33) and modn (-0.39) at (p<01). Significant positive correlations are observed between verbs and connectives (r=0.26, p<.05). Adjectives show significant negative correlations with awl (r=-0.22, p<05) and modn (-0.39, p<.001). Significant positive correlations are observed between adjectives and connectives (r=0.38, p<.001). Finally, the adverbs show significant positive correlations with asyl (r=0.43), awl (r=0.58) and modn (r=0.35) at (p<.005).

4.2 Readability

The Coh-Metrix investigates text readability using three indices, the Flesch Reading Ease (FRE), the Flesch-Kincaid Grade Level (FKGL) and the Coh-Metrix L2 readability. The FRE and FKGL indices are analysed below.

The FRE classical readability metric indicates scores between 0 and 100. Higher scores indicate higher ease of readability while lower scores indicate lower ease of readability. The scores per text are indicated in Table 7.

The mean grade for the combined grade 11 scores indicate that texts are easily understood and are suitable for 13 to 15 year old language learners. Grade 12 scores indicate fairly difficult to read texts. Therefore, grade 12 exam texts are harder to read than grade 11 texts. Additionally, summary texts are the most difficult section to read with the highest average grades. Furthermore, paired t-tests

Table 8: Flesch-Kincaid Grade Level

YEAR	11All	11A	11B	11C	12All	12A	12B	12C
2011	7.83	7.35	7.75	10.84	9.09	10.39	8.23	6.80
2012	8.91	9.26	10.11	7.14	7.71	8.17	7.42	7.24
2013	8.08	8.67	8.01	7.37	8.95	10.45	8.61	7.19
2014	7.55	7.44	7.94	7.88	9.74	10.99	10.34	6.10
2015	6.54	8.14	5.53	4.54	8.35	8.74	6.73	9.31
2016	8.44	9.29	6.28	8.63	8.22	9.22	8.30	5.41
2017	7.65	9.16	5.53	5.93	9.37	10.34	7.28	8.88
2018	7.85	7.62	7.49	9.21	8.82	9.67	7.15	9.56
2019	9.31	11.46	5.91	8.13	9.57	10.32	7.13	11.13
2020	7.28	7.00	7.26	8.12	8.75	10.36	7.89	5.57
Mean	7.94	8.54	7.18	7.78	8.86	9.86	7.91	7.72
SD	0.80	1.33	1.42	1.73	0.63	0.89	1.05	1.91

indicate that combined text scores for the grade 11 and grade 12 texts are significantly different at ($p<.05$). Even so, the texts A and texts C do not indicate any significant differences at ($p>.05$). The text B scores indicate significant differences between the grade 11 and 12 scores at ($p<.005$).

The FKGL outputs grade levels based on American school grades. The FKGL scores indicate an approximate grade level. The scores per text are presented in Table 8.

The FKGL for grade 11 indicate an overall 7th grade while the grade 12 scores indicate the 8th and 9th grade. Significant differences in grades can be observed between the grade 11 and grade 12 overall scores ($p<.05$). Reading comprehension texts also indicate significant differences in grades at ($p<.05$). Texts B and C do not indicate significant differences in grades. Consistent with FRE, section B texts are indicated as easier to read.

5 Discussion

Although the ttrc is significantly and strongly correlated with the ttr, it shows weak correlations with the MTLD and negative correlations with the *vocd*. Text length plays a negative role in these scores in that ttr and ttrc scores show varied results from the other scores. As mentioned earlier, *vocd* and MTLD were developed to counter the effects of text length on lexical diversity. As such, their significant strong correlations are more indicate of text diversity. Although the vocd and MTLD results are inconsistent, it can be concluded that grade 12 texts are more lexically diverse than grade 11 texts based on the fact that the ttr, ttrc, and MTLD scores.

The significant correlations observed for the syntax similarity in adjacent sentences and across the whole text indicate consistent sentence syntax. Although four POS tags were considered in this discussion, it is indicated that there is consistency in the texts chosen for exam texts. Since there is consistency in the average number of parts of speech observed from the grade 11 and the grade 12 exam texts. As such, the syntax similarity can be expected.

Correlations between connectives and sentence lengths may be an indication that short connecting words were used. The correlations observed are not significant. However, the relationship between the connectives and other parts of speech is interesting. The connectives are positively correlated with verbs, adjectives and adverbs and negatively correlated with the nouns. This coincides with the observation that nouns are negatively correlated with other parts of speech observed in this article.

Surface level features indicate positive correlations between average word lengths calculated through syllables and average word length based on letters per word. That is, as syllable counts increase, the number of letters also increase. It is surprising that the grade 12 scores show no relationship between average sentence lengths and average word lengths

by syllable counts. Unfortunately, this surprising result was not investigated further.

Text readability scores indicate that the grade 12 texts are harder to read than the grade 11 texts. The reading comprehension texts are easier to read in both grade 11 and grade 12 texts. From this finding, it can be concluded that grade 12 texts are more complex than grade 11 texts. Although this article only investigates the EFAL exam texts, the consistency in these findings confirm Sibeko and van Zaanen's (2021) conclusion that exam texts are consistent in readability.

6 Conclusion

This article set out to investigate the relationship between the grade 11 and grade 12 texts in terms of linguistic complexity. Overall, the relationship between the EFAL exam texts is positive in that there is consistency in the linguistic complexity properties in the two grades. The second aim of this article was to determine whether there is any increment in the linguistic complexity of exam texts from the grade 11 exam texts to the grade 12 exam texts. The lexical diversity and readability indexes indicate higher linguistic complexity in the grade 12 exam texts.

This article analyses only select indexes from the Coh-Metrix. Future research could investigate other indices not explored in this article. This would provide a more generalised analysis. Finally, this article presents the first analysis of language conventions section of the EFAL grade 11 and grade 12 texts. However, the focus has not been placed on these texts. Future research could pay special attention to these texts. Furthermore, exam texts for lower grades will provide a clear picture on the linguistic complexity of EFAL as s whole.

Notes

[1] https://www.education.gov.za/Curriculum/NationalSeniorCertificate(NSC)Exams/NSCPastExampapers.aspx

[2] https://www.ecexams.co.za/ExamPapers.htm

References

Ai, H & Lu, X 2013, 'A corpus-based comparison of syntactic complexity in NNS and NS university students' writing' In *Automatic Treatment and Analysis of Learner Corpus Data,* John Benjamins, p. 249-264.

Bulté, B & Housen, A 2012, 'Defining and operationalising L2 complexity. Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA', John Benjamins Publishing, p. 23-46.

Bulté, B 2013, '*The development of complexity in second language acquisition-A dynamic systems approach'*, Vrije Universiteit Brussel, PhD Thesis.

Charniak, E, 1997, '*Statistical parsing with a context-free grammar and word statistics, In Proceedings of the Fourteenth National Conference on Artificial Intelligence'*, AAAI/MIT Press, Menlo Park, Ca.

Covington, M & McFall, JD 2008, 'The Moving-Average Type-Token Ratio', view, 26 April, 2022, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.248.5206&rep=rep1&type=pdf.

Cvrček, V & Chlumská, L 2015, 'Simplification in translated Czech: a new approach to type-token ratio', *Linguist*, vol. 39, pp.309–325.

Dahl, Ö 2009 'Testing the assumption of complexity invariance: the case of Elfdalian and Swedish'. In: G. Sampson, D. Gil, and P Trudgill (eds.), *Language Complexity as an Evolving Variable*, Oxford: Oxford University Press, pp. 50–63.

Daller, H, Van Hout, R & Treffers-Daller, J 2003, 'Lexical richness in the spontaneous speech of bilinguals', *Applied linguistics*, vol. 24, no. 2, pp. 197-222.

DBE 2011a, *'English first additional language, Further Education and Training Phase Grades 10–12: Curriculum and Assessment Policy Statement (CAPS)'*, Government printing works, Pretoria.

DBE 2011b, *'English home language, Further Education and Training Phase Grades 10–12: Curriculum and Assessment Policy Statement (CAPS)'*, Government Printing Works, Pretoria.

DBE 2011c, *'English second additional language, Further Education and Training Phase Grades 10–12: Curriculum and Assessment Policy Statement (CAPS)'*, Government Printing Works, Pretoria.

De Clercq, B & Housen, A 2017, 'A Cross-Linguistic Perspective on Syntactic Complexity in L2 Development: Syntactic Elaboration and Diversity', *The Modern Language Journal*, vol. 101, no. 2, pp. 315-334.

Delage, H & Frauenfelder, UH 2019, 'Syntax and working memory in typically-developing children: Focus on syntactic complexity', *Language, Interaction and Acquisition,* vol. 10, no. 2, pp.141-176.

DBE 2014, '*Report on the Annual National Assessment (ANA) of 2014. Grades 1 to 6 and 9'*, Government Printing works, Pretoria.

DBE 2013, '*Report on the Annual National Assessment (ANA) of 2013. Grades 1 to 6 and 9'*, Government Printing works, Pretoria.

Graesser, AC & McNamara, DS 2011, 'Computational analyses of multilevel discourse comprehension', *Topics in Cognitive Science*, vol. 3, no. 2, pp. 371–398.

Graesser, AC, McNamara, DS & Kulikowich, JM 2011, 'Coh-Metrix: Providing Multilevel Analyses of Text Characteristics', *Educational Researcher*, vol. 40, no. 5, pp 223–234.

Graesser, AC, McNamara, DS, Louwerse, MM & Cai, Z 2004, 'Coh-Metrix: Analysis of text on cohesion and language', *Behavioral Research Methods, Instruments, and Computers*, vol. 36, pp. 193–202.

Jarvis, S 2002, 'Short texts, best-fitting curves and new measures of lexical diversity', *Lang. Test*, vol. 19, no. 1, pp. 57-84.

Johansson, V 2009, 'Lexical diversity and lexical density in speech and writing: A developmental perspective', *Lund Working Papers in Linguistics*, vol. 53, no. 1, pp. 61-79.

Kaiser, K, Reynecke, M & UYS, M 2010, 'Eating soup with a fork-why the EFAL syllabus cannot promote learning across the curriculum', *Journal for Language Teaching*, vol. 44, no. 2, pp. 52–67.

Khosa, M 2019, '*Exploring visual literacy development through films in senior phase English first additional language'*. (Master of Education), Rhodes University, Grahamstown.

Khushik, GA & Huhta, A 2020, 'Investigating Syntactic Complexity in EFL Learners' Writing across Common European Framework of Reference Levels A1, A2, and B1', *Applied Linguistics*, vol. 41, no. 4, pp. 506-532.

Kyle, K 2016, '*Measuring syntactic development in L2 writing: Fine grained indices of syntactic complexity and usage-based indices of syntactic sophistication'*, (PhD Thesis), Georgia State University.

Lu, X 2011, 'A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development', *TESOL quarterly*, vol. 45, no. 1, pp. 36-62.

Lu, X & Ai, H 2015, 'Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds', *Journal of Second Language Writing*, vol. 29, pp. 16–27.

MaCarthy, PM & Jarvis, S 2007, 'Vocd: A theoretical and empirical evaluation', *Language Testing*, vol. 24, no. 4, pp. 459–488.

MacWhinney, B 2000, '*The CHILDES Project: Tools for Analyzing Talk'*, Mahwah, NJ, Lawrence Erlbaum Associates.

Malvern, DD. Richards, BJ, Chipere, N & Durán, P 2004, '*Lexical diversity and language development: Quantification and assessment'*, Houndmills, Hampshire: Palgrave Macmillan.

Marcus, M, Santorini, B & Marcinkiewicz, M 1993, 'Building a large annotated corpus of English: The Penn Treebank', *Computational Linguistics*, vol. 19, no. 1, pp. 313-330.

McCarthy, PM 2005, '*An assessment of the range and usefulness of lexical diversity measures and the potential of the measure of textual, lexical diversity (MTLD)*'. (PhD dissertation), University of Memphis.

McCarthy, PM, Lehenbauer, BM, Hall, C, Duran, ND, Fujiwara, Y & McNamara, DS 2007, 'A Coh-Metrix analysis of discourse variation in the texts of Japanese, American, and British

Scientists', *Foreign Languages for Specific Purposes*, vol. 6, pp.46-77.

McNamara, DS, Crossley, SA & McCarthy, PM 2010, 'Linguistic features of writing quality', *Written Communication*, vol. 27, pp. 57–86.

McNamara, DS, Graesser, AC & Louwerse, MM 2012, 'Sources of text difficulty: Across genres and grades', *Measuring up: Advances in how we assess reading ability*, pp.89-116.

Miestamo, M 2009, 'Implicational hierarchies and grammatical complexity', In Sampson, G, Gil, D & Trudgill, P (eds.), *Language complexity as an evolving variable (Studies in the Evolution of Language 13)*, Oxford: Oxford University Press, pp. 80-97

Miestamo, M 2008, 'Grammatical complexity in a cross-linguistic perspective', *Language complexity: Typology, contact, change*, vol. 23, pp. 1-41.

Mitchell, M 2009, '*Complexity: A guided tour*', Oxford University Press.

Molepo, VM 2014, '*Implementation of the curriculum and assessment policy statements in selected primary schools in Limpopo Province'*, (MA dissertation), University of South Africa, Pretoria.

Molotja, TW & Themane, M 2018, 'Enhancing learners' reading habits through reading bags at secondary schools', *Reading & Writing,* vol. 9, no. 1, doi.org/ 10.4102/rw.v9i1.185.

Moodley, G 2013, '*Implementation of the curriculum and assessment policy statements: Challenges and implications for teaching and learning'*, (MA dissertation), University of South Africa, Pretoria.

Mophosho, M, Khoza-Shangase, K & Sebole, LL 2019, 'The reading comprehension of Grade 5 Setswana-speaking learners in rural schools in South Africa: Does home language matter?', *Per Linguam*, vol. 35, no. 3, pp. 59–73.

Ortega, L 2003, 'Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing', *Applied Linguistics,* vol. 24, no. 4, pp. 492-518.

Polio C & Yoon HJ 2018, 'The reliability and validity of automated tools for examining variation in syntactic complexity across genres', *International Journal of Applied Linguistics*, vol. 2018, no. 28, pp. 165–188.

Richards, BJ & Malvern, D 1997, *'Quantifying lexical diversity in the study of language development'*, University of Reading: Faculty of Education and Community Studies.

Sibeko, J & van Zaanen, M 2021, 'An analysis of readability metrics on English exam texts', In *Proceedings of the International Conference of the Digital Humanities Association of Southern Africa*, December 2021, pp. 1-11.

Sibeko, J 2021, 'A comparative analysis of the linguistic complexity of Grade 12 English Home Language and English First Additional Language exam papers'. *Per Linguam*, vol. 37, no. 2, pp. 50-64.

Van der Walt, C 2018, 'The category Language Structures and Conventions in the CAPS for English First Additional Language: a critical analysis', *Journal for Language Teaching*, vol. 52, no. 1, pp. 170–200.

Wolfe-Quintero, K, Inagaki, S & Kim, H 1998, *'Second language development in writing: Measures of fluency, accuracy, and complexity'*, Honolulu, HI: University of Hawaii at Manoa.

Zano, K & Phatudi, N 2019, 'Relationship between vocabulary knowledge and reading comprehension of South African EFAL high school learners', *Per Linguam*, vol. 35, no. 3, pp.16-28.

Zano, K 2020, 'Reading comprehension strategies of the efal learners in the fet phase: Teacher perspectives', *e-BANGI*, vol. 17, no. 7, pp.1-12.

Zano, K 2022, 'Breadth and depth-vocabulary knowledge and reading comprehension in an English first additional language context', *Journal of Languages and Language Teaching*, vol. 10, no. 2, pp.223-233.