

Linguistic complexity in high-school students' EFL writing

Amer Delić & Alma Jahić Jašić

*Faculty of Humanities and Social Sciences
University of Tuzla, Bosnia and Herzegovina*

Abstract

This study examined the syntactic and semantic complexity of L2 English writing in a Bosnian-Herzegovinian high school. Forty texts written by individual students, ten per grade, were quantitatively analyzed by applying methods established in previous research. The syntactic portion of the analysis, based on the t-unit analysis introduced by Hunt (1965), was done using the Web-based L2 Syntactic Complexity Analyzer (Lu, 2010), while the semantic portion, largely based on the theory laid out in systemic functional linguistics (Halliday & Matthiessen, 2014), was done using the Web-based Lexical Complexity Analyzer (Ai & Lu, 2010) as well as manual identification of grammatical metaphors. The statistical analysis included tests of variance, correlation, and effect size. It was found that the syntactic and semantic complexity of writing increases in later grades; however, this increase is not consistent across all grades.

Key words: writing development; syntactic complexity; semantic complexity, lexical density; grammatical metaphor.

1. Introduction

The characteristics of language one produces change as they get older, more educated, or more proficient in that language, native or foreign. While such development occurs in both speech and writing, the latter is naturally easier to observe. The written work of elementary school children is very different from that of college students, and most literate adults should be able to tell the difference even if not paying attention to the content. Word and sentence length could be useful characteristics to look at, for as Hunt (1970: 196) puts it, "Little people speak little sentences and big people speak and write big sentences." Now, as sensible as that idea is, unskilled writers might produce very long sentences simply because they do not know when to use a period instead of a comma, and the skilled ones might have a good reason to use very short and simple sentences. Sentence length may vary for a variety of

reasons, and language proficiency is just one of them. So, if it is at all possible to draw parallels between language proficiency and structure, two questions arise: What to measure, and how to interpret the results? One of the ways language development researchers have been attempting to answer said questions is through the notion of linguistic complexity (Ortega, 2003: 492).

Some studies on linguistic complexity have focused solely on the syntactic elements of complexity and have investigated various issues such as the effect of language proficiency level (Ortega, 2003; Lu, 2010), L1 (Lu & Ai, 2015), genre (Beers & Nagy, 2009), and classroom instruction (Mazgutova & Kormos, 2015; Vyatkina, Hirschmann, & Golcher, 2015) on development of syntactic complexity. On the other hand, some other studies have begun to adopt a more meaning-based approach which is grounded in the theory of Systemic Functional Linguistics (SFL) (Halliday & Matthiessen, 2014). One generally accepted idea is that measurements of syntactic complexity are more relevant to spoken than to written language (Biber et al., 2011: 9). In the case of written language, it is argued that as it becomes more complex, it becomes more lexically dense and more metaphorical, which is achieved, for the most part, through nominalization (Halliday & Matthiessen, 2014: 726-729). Therefore, the studies using the meaning-based approach propose the use of grammatical metaphor (GM) for investigating the nature of complexification for more advanced proficiency levels (Ryshina-Pankova & Byrnes, 2013; Ryshina-Pankova, 2015). This meaning-based approach has been referred to as the study of semantic complexity (Ryshina-Pankova, 2015: 53), and is now seeing increased use in combination with the syntactic approach (Ortega, 2015).

As no research exists on linguistic complexity in writing by students in the local EFL context, we believe that the current study can contribute to the existing pool of research on this topic. Therefore, in line with recent studies investigating linguistic complexity, the present study looks at both, syntactic and semantic complexity (detailed information on both is provided in sections 2 and 3 below) in 40 texts (ten per grade) written by high-school students with the aim to discover how elements of linguistic complexity change with the increase in writing proficiency that is assumed to occur over the course of a four-year high-school education. Furthermore, the study also aims to see if any insights gained from such analyses of linguistic complexity could have practical applications in EFL teaching. To achieve the aims, the study was guided by the following research questions:

RQ1: With respect to the measures of syntactic and semantic complexity suggested by the theory, what are the overall characteristics of writing pro-

duced by the participating students of one grammar-type¹ high-school in Bosnia and Herzegovina?

RQ2: Are there clear, significant differences in both syntactic and semantic complexity of texts produced by students in different grades of high school?

RQ3: Does the analysis reveal a pattern of writing development, and if so, how does it compare to the predictions offered by the theory?

2. Defining syntactic complexity

Syntactic complexity “refers to the range of forms that surface in language production and the degree of sophistication of such forms” (Ortega, 2003: 492). Central to syntactic complexity is the idea that a significant part of language development consists of learning a wide range of syntactic forms and understanding how to use them appropriately (ibid). The key word here is APPROPRIATELY because assuming that high syntactic complexity equals high text quality is simply incorrect. It is important to differentiate between proficiency in a given language and the actual skill of composing written work, where the former can be viewed as a prerequisite to any form of language use, and the latter as a “complex phenomenon that goes well beyond establishing the linguistic abilities of L2 writers” (Ortega, 2003: 494). In other words, syntactic complexity is relevant to, but not the principal factor of, writing quality.

One example of syntactic complexity's relevance to writing quality is its relationship to sentence fluency, which can be defined as “the rhythm and flow of the language, the sound of word patterns, the way in which the writing plays to the ear, not just to the eye” (Northwest Regional Educational Laboratory, 2004, as cited in Beers & Nagy, 2009: 186). To achieve such fluency, a decent grasp of a variety of syntactic structures is necessary (ibid). It has also been argued that being proficient in the use of complex syntax enables one to better express complex relationships between ideas (Coirier, 1996, as cited in Beers & Nagy, 2009: 187). So, as a writer's potential for meaning-making matures, progress in their command of intricate syntactic structures should be expected to follow, in order to facilitate their meaning-making activities.

¹ Grammar schools in Bosnia and Herzegovina are academic secondary schools which provide their students with comprehensive knowledge and prepare them for university education.

2.1. *Measuring syntactic complexity*

Discussing the level of complexity of any given grammatical structure implies the possibility of comparing that structure to its most basic variant. For example, a simple clause consists of a subject, verb, and object or complement, while a simple noun phrase has a determiner and a head noun; therefore, any change to these structures would include additional elements which would increase complexity (Biber et al., 2011: 12). To measure complexity, then, is to measure the number of such additions and the degree of modification. The measures of syntactic complexity that could be used in such analyses include length of production units (sentence and clause length), measures of the amount of coordination and subordination, and measures of clausal and phrasal complexity, to name a few (Yang et al., 2015: 58).

Earlier research has observed that length-based measures usually show an increase as writing becomes more mature (Hunt, 1965, 1970; Cooper, 1976), but this is not always true for every measure. For instance, one's writing might have a high degree of subordination indicating higher complexity; however, overall clause length could be low, indicating the opposite (Hunt, 1965: 20). Sentence length can also be an unreliable measure since it has the downside of being affected by the writer's handling of punctuation, which can be poor in the case of young learners. What this means for a text is that, with less skilled writers, it will likely have fewer stops and commas, in which case, it would likely contain more coordination using AND (in most cases), BUT, and SO (Hunt, 1965: 8-11). As a result, the less proficient writer's text would probably have much longer sentences which would lead to confusing results.

To avoid the abovementioned issues, Hunt (1965) recommends a new unit of measurement called t-unit (minimal terminable unit) which consists of a main clause with all its subordinate clauses (Hunt, 1965: 21). The minimal terminable part of its name refers to the fact that each one of these units has the minimal length to be "grammatically capable of being terminated with a capital letter and a period" (ibid). Therefore, a sentence could be comprised of several t-units which can, in turn, be comprised of several clauses. In such an analysis, excessive coordination is no longer an issue because the length of a t-unit cannot be increased by it. Therefore, the ways to increase the mean length of t-units (MLT) are limited to the use of subordinate clauses and other, non-clausal, elements (i.e. noun, adverbial, and adjectival phrases).

Besides MLT, the present study uses five other measures of syntactic complexity: mean length of sentence (MLS), mean length of clause (MLC),

clauses per t-unit (C/T), t-units per sentence (T/S), and complex nominals per t-unit (CN/T).

MLS is a length of production unit on the sentence level. By itself, a global measure² such as MLS sheds very little light on the ways in which writing becomes more syntactically complex. For example, it might show that the MLS of texts produced by a writer increases over time, but it does not show which particular syntactic structures contributed to that increase. Studies with longer observational periods and larger number of participants have indeed shown statistically significant differences in MLS values with respect to proficiency level (Ortega, 2003) although these values do not always progress linearly (Lu, 2010: 490).

MLC is also a length of production unit; however, it is a local-level measure because it shows the level of complexity within clauses (Yang et al., 2015: 55). It is important to note that what is meant by the term clause in this context is a finite clause, both dependent and independent, while non-finite clauses are considered as non-finite elements (Lu, 2010: 481; Yang et al., 2015: 54). It follows from this that additional non-finite clauses, by not contributing to the total number of clauses, contribute to overall clause length.

C/T is a measure of subordination on the t-unit level, also called the t-unit complexity ratio (Lu, 2010: 479). This local-level measure is considered to be important at intermediate levels (Norris & Ortega, 2009: 563). While C/T has been shown to increase with age (or proficiency), it tends to level off at higher proficiency levels (Norris & Ortega, 2009: 564).

T/S is a sentence coordination ratio. Bardovi-Harlig (1992, as cited in Norris & Ortega, 2009: 558) proposed that coordination could be a measure more suited for investigating complexity at the earliest levels of proficiency.

CN/T is, according to Lu (2010: 479), a measure of a particular structure (the other structure being the verb phrase). According to Cooper (1976: 180), "complex nominals comprise (i) nouns plus adjective, possessive, prepositional phrase, relative clause, participle, or appositive, (ii) nominal clauses, and (iii) gerunds and infinitives in subject position." This measure is, perhaps, the most interesting one to observe, for as Ryshina-Pankova (2015: 55) notes, "the nominal groups are more amenable to being modified by a variety of language structures such as prepositional phrases, attributes, relative clauses, and appositions." For instance, an increased number of complex nominals could mean more finite clauses, and this should result in increases in MLS, MLT, and C/T, but not necessarily MLC. If, however, growth was

² Yang et al. (2015: 55) categorize the mean length of a sentence and t-unit as global complexity measures while those that measure anything below the level of sentence and t-unit length (i.e. subordination, coordination, length of clause, etc.) are categorized as local-level complexity measures.

observed in MLC as well as CN/T, that could indicate a higher level of non-finite elements, adjectives, prepositional phrases, and other elements which increase MLC. Also, a high CN/T value could indicate an increased use of nominalizations which are considered characteristic of advanced levels of writing proficiency (Norris & Ortega, 2009: 562-563).

3. Defining semantic complexity

Although not specifically referring to it as semantic complexity, Norris and Ortega (2009: 563) suggested considering the GM as part of a multidimensional approach to the study of complexity, accuracy, and fluency. More recently, Ryshina-Pankova (2015) and Ryshina-Pankova and Byrnes (2013) explored the idea by applying it in the analysis of texts written by college level L2 German learners and found somewhat promising results.

The concept of GM comes from the distinction SFL makes between two ways of meaning-making, namely the congruent or direct and the incongruent or metaphorical way (Ryshina-Pankova & Byrnes, 2013: 181). Simply put, to congruently express meaning is to use nouns, verbs, and adjectives to represent things, actions (or processes), and qualities, respectively; on the other hand, to express meaning incongruently is to construe actions and qualities as things and present them as nouns (*ibid.*). The GM is therefore defined as a “remapping between the semantics and grammar” (Halliday & Matthiessen, 2004, p. 640), a realization where single elements as nominalizations metaphorically stand for an entire situation typically realized through a clause rather than a phrase” (Ryshina-Pankova, 2015: 54). The following example from Halliday and Matthiessen (2014: 710) illustrates the shift from congruent to incongruent:

Slate is a metamorphic rock. Slate was once shale. But over millions of years, tons and tons of rock *pressed* down on it. The pressure made the shale very *hot*, and the heat and pressure changed it into slate. ... Other metamorphic rocks are made the same way slate is, by heat and pressure.

In the example above, the congruent expressions TO PRESS and HOT are realized incongruently as PRESSURE and HEAT. According to SFL, “these nominalizations are examples of ideational³ metaphors where processes and qualities are construed as if they were entities” (*ibid.*). As can be seen here,

³ SFL theory makes a distinction between two types of GM, namely interpersonal and ideational, where the former refers to metaphorical realizations of expressions of modality and is most relevant to spoken language, while the latter refers to processes and qualities being construed as things by nominalization, and is most relevant to written language (Halliday & Matthiessen, 2014: 686-726). For this reason, the present study only focuses on the ideational variant of GM and all uses of the term GM in the present study refer to the ideational GM.

the GM is, similarly to the lexical metaphor, metaphorical in the sense of having additional meaning beyond the literal meaning. The difference is, as Byrnes (2009: 350) notes, “while for lexical metaphor, the potential for multiple meaning construal of experience (something resembles something else) lies in the vocabulary item, with GM it is located in grammar”, or, more specifically, in the processes of “derivational morphology (to discuss–discussion), agnation (to work–work) or a new lexical item (near–proximity)” (Ryshina-Pankova, 2015: 53).

Halliday and Matthiessen (2014: 730) explain that one of the reasons why GM can be considered more difficult for the language user is due to a loss of information that can occur when a shift is made from congruent to incongruent. They show this in the example of the congruent expression *IS IMPAIRED BY ALCOHOL* becoming incongruent as *ALCOHOL IMPAIRMENT* (ibid). The loss of information here refers to the role of *ALCOHOL* in the incongruent expression becoming less obvious and allowing for other possible, but incorrect, interpretations such as *ALCOHOL IS IMPAIRED* (ibid).

In sum, incongruent, metaphorical modes of expression allow for an expansion in meaning, while also making additional cognitive, morphological, or vocabulary-related requirements of the language user.

3.1. *Measuring semantic complexity*

Besides identifying GMs, this study considers lexical density (LD) as well as the total number of nouns in a given text as potential indicators of semantic complexity. LD, a measure of lexical complexity, represents “the ratio of the number of lexical (as opposed to grammatical) words to the total number of words in a text” (Lu, 2012: 191). In the case of increased LD, Halliday (2002, as cited in Byrnes, 2009: 58–59) suggests that this usually comes as a result of a decrease in the number of grammatical words, rather than just an increase in the number of lexical words. Now, as both verbs and adverbs are in certain forms counted as grammatical words (i.e. modal verbs, auxiliary verbs *BE* and *HAVE*, adverbs ending in *-LY*, etc.), while nouns are always lexical words (Lu, 2012: 192) and are modified by adjectives, one could, perhaps, conclude that this rise in LD is primarily due to an increased occurrence of nouns, nominalizations, and, consequently, GMs. By choosing to measure the semantic complexity of a text in terms of its lexical properties, this study makes the assumption that said lexical properties are functionally connected to the expression of meaning. The need to make this assumption is also mentioned by Ryshina-Pankova and Byrnes (2013: 194) who, while discussing a meaning-oriented theory of language, note that “such a theory would need to assume a ‘natural’ relationship between its lexicogrammatical resources and the world that is being construed through them whereby the organiza-

tion of language itself, its grammar, must be seen as meaningful and analyzed accordingly.”

4. Methodology

4.1. Data

The data used in this study consists of 40 texts written by students of one grammar-type high school in Bosnia and Herzegovina (BiH). While the proficiency level of the students has not been tested, it is assumed, based on the textbooks they use, that they range from pre- to upper-intermediate. The texts are completely authentic since they were written as part of mandatory writing tasks that students are required to do in class twice per year for evaluation purposes and not specifically for this study. For each high-school grade, ten texts were selected from the archive of student-written work based on two criteria: primarily that the texts were rated ‘excellent’⁴ and secondarily that they were on a similar topic, if possible. As the collected texts represent authentic material, it was impossible to ensure that all the texts were of the same genre and on the same topic. For detailed information on genres and topics see Table 1.

Table 1: Information on genre and topics of student texts for each grade.

Grade	Genre	Topic
Grade I	narrative	‘What a horrible/amazing day in my life’
	narrative	‘The day I was very lucky/unlucky’
	narrative	‘The day I got some good/bad news’
Grade II	argumentative	‘The best advice given to the young is: find out what you like doing best and get someone to pay you for doing it’
Grade III	argumentative	‘Technology plays a huge role in our lives today’
	argumentative	‘Luck has nothing to do with success’
	argumentative	‘The media have a great influence on shaping peoples’ ideas’
Grade IV	argumentative	‘Is animal testing necessary’
	argumentative	‘Travelling broadens our horizons and enriches our lives’
	narrative	‘An event that changed my life’

⁴ Since the study used authentic material, text evaluation was unrelated to the study and was done by regular English class teachers.

In preparation for statistical analysis, the original, handwritten texts were transcribed into digital format with minimal corrections to the students' written work. The corrections included adding or removing commas, capitalizing certain words, and spelling corrections and were necessary because, while none of them have an impact on the text's syntactic structure, these small mistakes would interfere with the automatic analysis which does require correct punctuation and spelling to provide reliable results.

4.2. Analysis methods

The analysis begins with a one-way ANOVA test, which was used to see if there are statistically significant differences in syntactic complexity indices between grade groups, and this is accompanied by a Tukey post hoc for specific grade groups. Secondly, the effect sizes for the statistically significant between-level differences in complexity indices are calculated using Cohen's d (Cohen, 1988). The value of d is interpreted according to Cohen's (1988: 25) recommendations for effect size interpretation in behavioral sciences. The effect size is defined as follows: $d = .2$ (small), $d = .5$ (medium), and $d = .8$ (large) (ibid). Finally, the Pearson Correlation Test was used to determine statistically significant links between complexity indices across the four grades.

The syntactic portion of the analysis was conducted using the Web-based L2 Syntactic Complexity Analyzer (Lu, 2010), and the measures of syntactic complexity that were used are MLS, MLT, MLC, C/T, T/S, and CN/T.

Semantic complexity was quantitatively measured in three ways. Firstly, LD was measured using the Web-based Lexical Complexity Analyzer (Ai & Lu, 2010). Secondly, each text's ratio of normal nouns (NN) to the total word count was obtained by a manual count, resulting in the NN/W (normal nouns per word) index. Lastly, instances of GM were counted manually by the lead author, and the number of GMs in a text was then divided by the text's word count resulting in the GM/W (grammatical metaphors per word) index. GMs were coded for type as nominalizations of a process (P) or attribute (A), as well as for the processes of agnation (a) and derivation (d).

5. Results

5.1. Syntactic complexity

Table 2 shows the means and standard deviations for syntactic complexity indices across the four grades. The mean values of MLS, MLT, MLC and CN/T increase between grades I and III with the largest increase occurring

at grade II, while C/T and T/S show inconsistent changes across all grades. All increases or decreases between grades III and IV appear negligible.

Table 2: Syntactic complexity means and standard deviations for all grades

Grade		MLS	MLT	MLC	C/T	T/S	CN/T
Grade I	Mean	12.1770	9.7790	6.7190	1.4400	1.2420	.7040
	Std. Deviation	2.83645	2.25029	1.00330	.13984	.05308	.22994
Grade II	Mean	14.8600	13.1856	8.0333	1.6422	1.1400	1.2378
	Std. Deviation	1.88674	1.58003	.81629	.16612	.12430	.40165
Grade III	Mean	17.4067	14.0689	8.8489	1.5989	1.2233	1.4167
	Std. Deviation	3.29700	1.46736	.73334	.23903	.11769	.23638
Grade IV	Mean	17.6600	14.0411	8.7078	1.6344	1.2589	1.4267
	Std. Deviation	2.01509	1.63686	1.45450	.23896	.13842	.11937

A test of variance shows statistically significant⁵ differences in MLS ($F(3, 33) = 9.419, p = .000$), MLT ($F(3, 33) = 12.719, p = .000$), MLC ($F(3, 33) = 8.448, p = .000$), and CN/T ($F(3, 33) = 15.820, p = .000$). Differences in C/T ($F(3, 33) = 2.181, p = .109$) and T/S ($F(3, 33) = 2.018, p = .131$) do not show statistical significance.

A Tukey post-hoc test shows statistically significant increases in MLS, MLT, MLC, and CN/T, but only when comparing grade I values to those of subsequent grades. For MLC, MLT, and CN/T the increases are significant at grades II, III, and IV, while the MLS increase is significant only at grades III and IV. Lastly, all of the d values for the abovementioned significant increases indicate a large effect size ($d > 0.8$). Table 3 shows the statistical significance (p) and effect size (d) of increases at grades II, III, and IV.

Table 3: Statistical significance (p) and effect size (d) of increases at grade II, III, and IV

Measure	I → II	I → III	I → IV
MLS	<i>Not significant</i>	$p = .001; d = 1.7$	$p = .000; d = 2.2$
MLT	$p = .001; d = 1.7$	$p = .000; d = 2.25$	$p = .000; d = 2.16$
MLC	$p = .045; d = 1.4$	$p = .000; d = 2.4$	$p = .001; d = 1.59$
CN/T	$p = .001; d = 1.6$	$p = .000; d = 3$	$p = .000; d = 3.9$

⁵ Differences are statistically significant if Sig. (p) ≤ 0.05.

5.1.1. Correlation of syntactic complexity measures

As can be seen in Table 4, MLS shows positive correlations of varied strength with all other measures. T/S, on the other hand, has only one significant correlation, a moderate, positive one with MLS. The MLS results are unsurprising given the very general nature of the measure; however, the T/S results show an almost complete lack of significant correlation, which is unexpected. The strongest correlation observed is the one between MLS and MLT. This, along with the weakest correlation being between MLS and T/S, suggests that sentence length is mostly increased by extending t-units, rather than increasing their number.

MLT correlates strongly and positively with MLC, C/T, and CN/T. The strongest correlation is that with CN/T, a measure that also correlates strongly with MLC and C/T. With such results in mind, and considering the fact that complex nominals necessarily increase the length of clauses and may have additional subordinate clauses as part of their construction, it could be said that the increased t-unit complexity, as shown here, is mostly due to the addition and lengthening of clauses that are part of a complex nominal structure.

Grade level shows strong, positive correlations with MLS, MLT, MLC, and CN/T. Interestingly, the correlation with CN/T is slightly stronger than the one with MLS.

Table 4: Pearson Correlation results for syntactic complexity measures

	MLS	MLT	MLC	C/T	T/S	CN/T	Grade
MLS	1	.875**	.656**	.564**	.404*	.734**	.650**
MLT	.875**	1	.749**	.630**	-.078	.864**	.640**
MLC	.656**	.749**	1	-.033	-.054	.671**	.596**
C/T	.564**	.630**	-.033	1	-.031	.507**	.306
T/S	.404*	-.078	-.054	-.031	1	-.094	.118
CN/T	.734**	.864**	.671**	.507**	-.094	1	.689**
Grade	.650**	.640**	.596**	.306	.118	.689**	1

5.2. Semantic complexity analysis

Table 5 shows the means and standard deviations for semantic complexity measures across the four grades. There is observable growth in all three indices up to grade III. As with the syntactic results, the changes observed between grades III and IV appear to be negligible.

A test of variance shows that there are statistically significant between-level differences in NN/W ($F(3, 33) = 5.928, p = .002$) while there are no statistically significant differences in LD ($F(3, 33) = 2.118, p = .117$) and GM/W ($F(3, 33) = .544, p = .656$).

Table 5: Semantic complexity means and standard deviations

Grade		LD	NN/W	GM/W
Grade I	Mean	.4540	.1680	.005280
	Std. Deviation	.02066	.04290	.0053497
Grade II	Mean	.4633	.1844	.005489
	Std. Deviation	.02784	.02128	.0044864
Grade III	Mean	.4856	.2200	.007567
	Std. Deviation	.04065	.02784	.0103637
Grade IV	Mean	.4856	.2100	.007633
	Std. Deviation	.04216	.02000	.0032377

A Tukey post hoc test shows statistically significant increases in NN/W at grades III (.220 +/- .027, $p = .003$) and IV (.210 +/- .020, $p = .022$) when compared to grade I (.168 +/- .042).

The large effect sizes of differences in NN/W for the two grade groups, 1.43 (I-III) and 1.25 (I-IV), indicate high practical significance. The effect size is larger at grade III than at grade IV, which is somewhat similar to the effect size pattern observed in syntactic complexity measures MLT and MLC.

5.2.2. Correlation of semantic complexity measures

The correlation test (Table 6) shows that there is a strong, positive correlation between LD and NN/W while GM/W shows no significant correlation. Grade level shows positive correlation with both LD and NN/W – the correlation with LD is moderate, while the one with NN/W is strong. This result supports the idea that higher proficiency writing would have more content than grammatical words, most of which should be nouns.

Table 6: Pearson Correlation results for semantic complexity measures.

		LD	NN/W	GM/W	Grade
LD	Pearson Correlation	1	.572**	.007	.380*
NN/W	Pearson Correlation	.572**	1	-.038	.524**
GM/W	Pearson Correlation	-.065	-.021	1	.182
Grade	Pearson Correlation	.380*	.524**	.200	1

5.2.3. Analysis of GM

The statistical analysis did not provide any significant findings in the case of GM use, even though the GM/W index shows an increase from .0052 in grade I to .0075 in grade III. The largest change in the value of this index, an increase of 38.8%, occurs between grades II and III while it remains practically unchanged at grade IV. It is likely that the relatively large increase at grade III is flagged as statistically insignificant because of very high standard deviation caused by the fact that some texts exhibited zero instances of GM, while others had as many as six. In fact, the standard deviation at grade I is actually larger than the mean value of GM/W. Nevertheless, the observable increase in GM/W warrants a closer look at the students' use of GM. Table 7 shows all 61 abstract nouns which were identified as potential GMs.

Besides a difference in the number of GMs, there is a noticeable difference in the amount of agnation and derivation used. The first two grades display an almost equal amount of agnation and derivation, while the latter two show a distinct preference for derivation. With regards to the types of GM, all of the texts show a slight tendency towards the process-type.

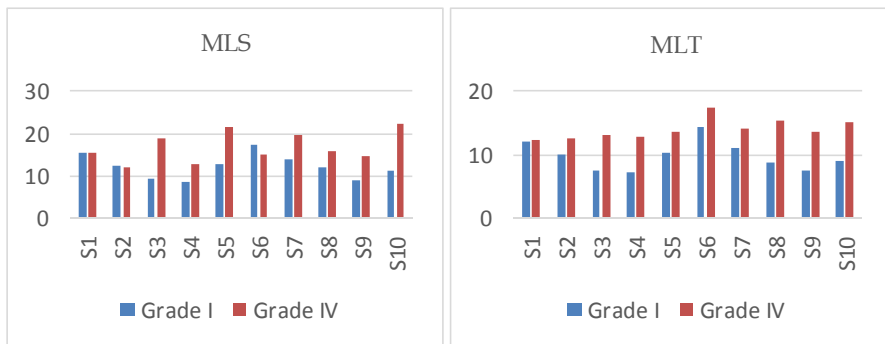
Table 7: List of nouns identified as GMs.

Grade I (13 words)	<i>surprise (Aa); boredom (Ad); disappointment (Ad) ; rental (Ad); lead (Pa); decision (Pd); argument (Pd); motivation (Ad); accomplishment (Pd); love (Pa); hug (Pa); smile (Pa); explanation (Pa)</i>
Grade II (14 words)	<i>comfort (Aa); spelling (Pd); choice x3 (P); fears (Aa); payment (Pd), stability (Ad), creativity (Ad); studies (Pa); experience (Pa); specialties (Pd); sacrifices (Pa); education (Pd)</i>
Grade III (17 words)	<i>happiness (Ad); fulfilment (Pd); suffering (Pd); patience (Ad); curiosity (Ad); responsibility (Ad); journey (Pa); distraction (Pd); communication (Pd); impact (Pa); behavior (Pd); replacement (Pd); efficiency (Ad); improvements (Pd); thought (Pa); manipulation (Pd); importance (Ad)</i>
Grade IV (17 words)	<i>risk (Pa); risk (Aa); torture (Aa); violation (Pd); experience (Pa); outlook (A); necessity (Ad); importance (Ad); behavior (Pd); speech x3 (P); existence (Pd); appearance (Pd); importance (Ad); nonsense (A); freedom (Ad)</i>

6. Discussion

As expected, all three length-of-production measures showed a significant increase in later grades; however, it is important to note that the differences between individual students were perhaps just as significant. For example, the present study's grade I MLS results range from a minimum of 8.5 to a maximum of 15.5, while the results at grade IV range from 12 to 22 words per sentence. Based on these individual results, many of the students from the high-proficiency group could easily be put into the low-proficiency group. It is only when all the MLS values for individual grade levels are averaged and compared that a pattern of increasing sentence length can be observed. However, due to the high variations in MLS values, these differences are statistically significant only when comparing grade I to grades III and IV. Because of this, relying on sentence length to assess a learner's writing development is only somewhat practical in the case of extreme MLS values which would be those that are about two standard deviations larger or smaller than the established mean for a developmental stage, or in this case, grade level. Even then, a finding of particularly low MLS is likely to be more useful, simply because it points at a definite lack of complexity while a particularly high MLS could mean either increased complexity or issues with excessive coordination. Figure 1 shows a visual representation of MLS and MLT values obtained in the present study.

Figure 1: Visual comparison of MLS and MLT values at grades I and IV.



T-unit length also showed a tendency to increase in later grades; however, there are two notable differences. One was that the MLT increase was statistically significant at an earlier point, namely at grade II. Moreover, the effect size of the MLT increase at grade II ($d = 1.75$) is practically equal to the MLS effect size at grade III ($d = 1.70$). Because the overall increase of MLT at grade II was larger than the overall increase of MLS at this point, the results

indicate that the increase in syntactic complexity observed at grade II is not only the result of clause lengthening, which would have an equal effect on both MLS and MLT, but that important changes must have also occurred in the amounts of subordination and coordination.

It was expected that the value of C/T (subordination) would increase at higher grades while T/S (coordination) would show a comparable decrease. The combined effect of these changes would be an increase in MLT. Indeed, C/T and T/S indices show an increase and decrease respectively, but the effect can only be observed at grade II. As a matter of fact, grade II texts had the highest C/T and lowest T/S values of all, meaning that at higher grades there are many cases where C/T actually decreases while T/S increases. This is completely unexpected given the prediction that coordination would decrease in favor of more subordination. The absence of such a pattern could be due to insufficient sample size, but it could also mean that the writers' use of coordination was not excessive at grade I to begin with, in which case, any increases in proficiency were unlikely to result in less overall coordination.

The other observable difference between MLS and MLT is the amount of variation in their values at grades I and IV. Figure 1 shows that the MLT values are more stable at grade IV than at grade I, while MLS maintains a high degree of variation at both stages. This could be understood as the effect of a four-year education in the sense that it made students' writing more uniform with regards to t-unit structure. In any case, MLT showed a more stable pattern of increase than MLS.

The third length of production measure, MLC, showed a similar pattern as MLT, in that the score differences are statistically significant at grade II while they show a slight drop in grade IV. Similar to the abovementioned changes in MLS and MLT, these findings indicate that the most visible changes occur in the first two grades.

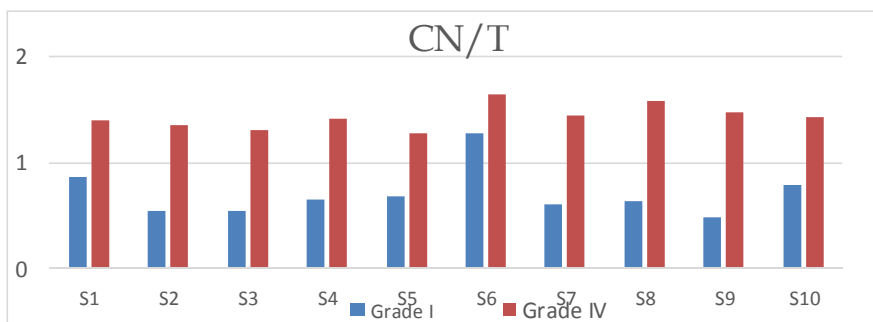
All things considered, the results here suggest that a combination of MLC, C/T, and MLT could provide useful insight because it should give a definite answer to the question of how MLT is being increased. For example, at grade II, the large MLT increase is supported by increases in both MLC and C/T, meaning that the students are writing longer, more complex clauses while at the same time using more subordination. At grade III, the slight MLT increase is supported only by a MLC increase while the C/T value actually shows a drop, leading to the conclusion that the type of complexification occurring at this point is different in the sense that the clauses are getting longer and more complex with a slightly lesser degree of subordination. At grade IV, MLC remains more or less unchanged while C/T increases ever so slightly, indicating that the students are producing clauses of similar

complexity while also finding ways of incorporating more of them within the same t-unit through subordination.

The last syntactic complexity measure left to discuss is CN/T. As expected, the number of complex nominals increased in each consecutive grade following a pattern similar to that of the three length-of-production measures. What makes this measure special is the amount by which it increased. The value of CN/T rose from 0.7 at grade I, to 1.42 at grade IV, which is an increase of 103%. For comparison, the percentages of increase between grades I and IV for MLS, MLT, and MLC are 45%, 44%, and 30%, respectively. The largest CN/T increase was at grade II (75%). At grade III, the value increases but only by roughly 15%, while at grade IV it stagnates with an increase of just below 1%. This is again the same type of pattern that was observed in MLS, MLT, and MLC, and it likely hints at important syntactic changes in students' writing occurring immediately after the first grade, but, at the same time, this also highlights the absence of observable changes between grades III and IV. In any case, when these changes do occur, they appear to be closely linked to the usage of complex nominals.

Changes in CN/T were closely tied to changes in other measures. It had strong positive correlation with MLS, MLT, MLC, and C/T. The only measure that did not correlate positively with CN/T is the coordination measure (T/S) where the correlation was weak, negative, and not statistically significant.

Figure 2: Visual comparison of CN/T values at grades I and IV



Another noticeable characteristic of the CN/T result is the lesser degree of variation of its values at grade IV. A comparison of standard deviations (SD) at grades I and IV shows that the CN/T value differences among first-graders' texts (SD = .229) were much larger than those in texts produced in the final grade (SD = .119), an effect clearly observable in a visual comparison of said values (Figure 2). This suggests that, at least in the case of this

study's participants, the number of complex nominals within a t-unit is a relatively stable variable among writers of similar proficiency.

The observations concerning CN/T, as shown above, indicate a more visible and stable increase than other measures used here. Thus, CN/T appears to be a very important characteristic, at least in the case of written language, which makes it surprising that none of the twenty-one studies examined by Ortega (2003: 496-497) used CN/T as a measure of syntactic complexity in writing. More recently, however, Lu and Ai (2015: 22) noted that CN/T was one of three measures, the other two being MLC and complex nominals per clause (CN/C), that showed significant differences between native and non-native groups of writers. This is similar to the present study's finding – the between level differences were most significant in CN/T, followed by MLC and MLT measures (Table 3). These findings seem to suggest that if there are any valid reasons to draw parallels between syntactic complexity and language proficiency, there is also a reason to include CN/T in such analyses.

In sum, the analysis showed an expected pattern of increasing sentence, t-unit, and clause length while the expected increase of subordination followed by a decrease in coordination was not observed. Grade I students produced on average significantly shorter sentences, t-units, and clauses than most of the students from later grades. The most significant and consistent change occurring in each consecutive grade was the difference in CN/T, a measure that correlated highly and positively with all length-based measurements as well as with the amount of subordination. Therefore, it appears that the primary characteristic of increasing syntactic complexity displayed by this group of students is the degree of noun usage and modification. Levels of coordination and subordination, for the most part, did not prove to be important characteristics in the sense that they did not show consistent or significant changes. Finally, the syntactic measures used here failed to show clear, significant differences between each of the four high school grades. With that said, the clear and significant differences which were observed are only apparent when comparing the first grade to any of the subsequent grades.

In the case of semantic complexity, only the NN/W index displayed changes that were statistically significant. The average value of this particular measure increased from 0.168 (at grade I) to 0.22 nouns per word (at grade III), which is an increase of about 31% (a large effect size (*d*) of 1.43). A higher NN/W value also indicates a lower usage of other types of words, such as verbs. Together with an observable but statistically insignificant LD increase, this increase in the total number of nouns found in each consecutive grade (up to grade III) does support Halliday's (2002, as cited in Byrnes, 2009: 58-59) observation that writing becomes more lexically dense due to a decrease in the number of grammatical words. To illustrate what this differ-

ence in LD and NN/W might look like in practice, below are two excerpts from the texts analyzed in the present study. The first was written in grade I and the second in grade IV. Nouns along with their determiners and modifications are underlined and bolded, while proper nouns are excluded unless they modify another noun.

*I woke up at nine o'clock, and **the day** was starting very well. I had lots of **things to do**. I had **two games of basketball**, one was starting at half-past eleven, and one at two o'clock in **another city**. After these **two games**, I thought that I could get **some rest**, but at six o'clock, something horrible started. At seven o'clock in the morning, **my phone** rang because I had forgotten to turn it off in **the evening**. I tried to get back to **sleep**, but at nine o'clock I gave up. So I got up, had **breakfast**, and, while eating, I looked at **my clock** and it was only 10 o'clock.*

The excerpt above has an LD of 0.42, an NN/W of 0.106, and four complex nominals, one of which is a repeated instance of MY PHONE. The rest of it consists mostly of verbal elements, adverbs of time, and a single adjective modifying the pronoun SOMETHING. As such, and especially if read out loud, the text seems more like a transcription of spoken language rather than a carefully thought out piece of written language. Furthermore, the text offers, arguably, very little in terms of content. It begins with the writer's DAY and the THINGS TO DO in that day. These THINGS TO DO are presented in the complex nominal TWO GAMES OF BASKETBALL with the only additional meaning introduced being the time and place of these games. The writer then follows this with a series of actions, such as GET SOME REST, MY PHONE RANG, GAVE UP SLEEPING, HAD BREAKFAST, LOOKED AT THE CLOCK, all of which are presented in a very congruent, verb-centered manner. Now, compare this to the second excerpt below.

*I have always believed that **musicians** had **a specific point of view**. Thanks to **my parents who gave me the opportunity to play the violin**, I realized how **music** changed **my life**. It happened two years ago when I was in Leipzig with **my friends from the music school**. We were invited to play for **kids without parents**. **The event** was well-organized and very interesting. First, it started with **a song played by some music teachers**. They played **the masterpiece called "Ninth Symphony" by Beethoven**. Then, **one of the organizers** gave us **a short speech** in which he explained why it was important to make it happen.*

This excerpt has an LD of 0.48, an NN/W of 0.168, and nine complex nominals with no repetition. The first sentence introduces the writer's opinion of musicians by employing the noun phrase A SPECIFIC POINT OF VIEW which, while not being the best choice of words, arguably has a higher meaning potential than if it were presented more congruently as THEY THINK

DIFFERENTLY. The rest of the text is also mostly centered around nouns which have their meaning expanded in various ways, such as by relative and infinitive clause (MY PARENTS WHO GAVE ME THE OPPORTUNITY TO PLAY THE VIOLIN), by prepositional phrase (KIDS WITHOUT PARENTS; FRIENDS FROM MUSIC SCHOOL), and by participle (A SONG PLAYED BY). All of these characteristics make the text read like one that has had thought put into the way meaning is constructed, as opposed to the first example which reads more like a spoken language variant.

The two examples examined here are certainly quite different with respect to their syntactic complexity, in that the second example is significantly more complex. While the increased syntactic complexity might as well be interpreted as containing more complex meaning, it could be argued that the main reason the second example is more semantically complex is that it is organized around nominal structures which are simply more meaningful than verbal ones, both in terms of conveying the meaning of a person or object (i.e. MY FRIENDS FROM THE MUSIC SCHOOL), and the meaning of an action or process (i.e. conveying the meaning of THINK DIFFERENTLY as HAVE A SPECIFIC POINT OF VIEW).

The reason GM was not included in the example analysis of the two excerpts discussed above is that this measure produced very inconsistent results across the entire dataset. While the number of GM occurrences did increase ever so slightly in later grades, the total number of GMs identified is very low and varies substantially from student to student, making the results inconclusive. Such results are not completely unexpected due to the age and proficiency level of the study's participants, as well as the rather short length of texts.⁶ With that said, the number of GMs, as shown by the GM/W measure, did increase between the first and last grades, but the only substantial change came at grade III where GM usage rose by 38.8%. This increase, however, is somewhat misleading because of a small total number of GMs, as well as large individual differences in participant GM usage.

Out of the total 17 GMs found in ten third grade texts, one text contained six GMs, two texts had three GMs each, one text had two GMs, three texts had only one GM each, and the two remaining texts had zero instances of GM. Therefore, the apparently large grade III increase in GM/W is largely due to one participant's writing. Here is an excerpt from that text:

*For example, ever since I was a child, I wanted to be a doctor. For me, it's a key aspect of **happiness** and **fulfillment** in life. I feel that my personal mission is to help humanity defeat diseases and overcome **suffering**. But it takes a long time to study medicine. It's very difficult and requires a tremendous amount of **patience** and **hard work**.*

⁶ The analyzed texts were about 250 to 300 words long.

The five bolded words were all identified as GMs based on them being nominalizations of either a process or attribute. The abstract noun SUFFERING has an arguably slightly wider meaning than its congruent, verbal counterpart, but not much more than that. HAPPINESS, FULFILLMENT, PATIENCE, and WORK, on the other hand, are concepts that the writer expands on. HAPPINESS and FULFILLMENT are presented as objects of interest, allowing the writer to explain what these mean to them as well as what it takes to achieve them. The two congruent expressions, TAKES A LONG TIME TO STUDY and VERY DIFFICULT, are summarized as PATIENCE and HARD WORK, which serve the purpose of underlining the main points or ideas the author is trying to convey, or, in other words, the GMs here serve as “a powerful textual resource for managing the creation of text” (Halliday & Matthiessen, 2014: 718). This metaphorical use of nominalization is the type of meaning-making that is arguably beneficial to perceived text quality; however, this is also the only example of such GM use in the analyzed texts.

Setting aside the overall number of GM occurrences, one somewhat interesting find concerning GMs was that the students tended to use more derivation rather than agnation when nominalizing. Generally, one would likely expect to find more of the former due to the additional knowledge of affix usage required by derivation.

An additional observation concerning GM usage was the slight difference in the number of process vs. attribute types of nominalization found in the present study. Participants in grades I, II, and III had twice as many nominalizations of processes (i.e. DECISION, ARGUMENT, EXPLANATION) than those of attributes (i.e. HAPPINESS, RESPONSIBILITY, PATIENCE), while at grade IV, both types are used in equal measure. This prevalence of process type GMs could indicate that a process, rather than an attribute, might be more difficult to construe as a thing or it could simply be that process type GMs are more common in general. However, the present study's dataset is much too small to make any confident conclusions regarding this finding.

To summarize, increases were observed in all three measures used, but only the NN/W measure had statistical significance. The most important characteristic turned out to be the increasing number of nouns in later grades, while the increases in LD were minor, and GM usage was not nearly as significant. Moreover, the low and inconsistent occurrence of GM suggests that this measurement is impractical for analysis when participants are at intermediate proficiency or lower. Nevertheless, a few texts that contained several instances of GM did appear to be of overall higher quality. A tendency to lean towards writing which favors noun phrases instead of verb phrases was also noticed in later grades. Such a shift is in line with previous studies which also observed that academic writing is primarily composed of “embedded noun phrases and prepositional phrases” (Biber et al., 2011: 9).

Finally, as was the case with syntactic measures, these measures also failed to show clear, significant differences between each of the four high school grades, with the only clear difference being between grade I and subsequent grades.

The last point to discuss is the finding of no significant change in any measure between grades III and IV. It seems reasonable to expect that indices would continue to increase until students reach advanced proficiency. If grade IV results are, for example, compared to native English speakers and advanced EFL speakers with Bulgarian as L1 from Ai and Lou's (2015) study, it can be seen that grade IV students lagged behind in all but the T/S measure which, as discussed earlier, is not expected to increase with proficiency. Therefore, the lack of continuous growth between grades III and IV is not an expected finding. The reason for such a situation is open to speculation. One possibility is that this is due to the lack of sufficiently advanced input. Schools in BiH generally do not offer explicit writing instruction (Jahić, 2009: 174) so student's writing seems to develop, for example, due to exposure to reading material found in their textbooks and the occasional writing tasks students have to complete. Another possible reason could be the type of writing task. All of the grade III texts were argumentative while four, out of ten, grade IV texts were narrative. Regarding the effect of the writing task on syntactic complexity, Lu and Ai (2015: 18) note that lower syntactic complexity is generally found in the narrative type. However, this study's findings do not corroborate this since comparing the mean values of grade IV narrative-type texts with the argumentative texts reveals almost the opposite – the argumentative texts were equally, or even slightly less syntactically complex.

7. Conclusion

In conclusion, the syntactic analysis turned out to be more informative in the sense that most of the measures used showed a pattern of varying increase. Measures of production length showed a somewhat consistent increase, while the most notable change was an increase in complex nominal usage. On the other hand, the semantic analysis was rather inconclusive. It showed that the students produced very few instances of nominalization and even fewer of those that can be termed actual GMs. Comparison of the highest to lowest proficiency texts, so as to understand the effects of a higher degree of nominalization, showed that higher proficiency texts had on average more nouns, nominalizations, and higher lexical density. As for nominalizations, the students tended to nominalize actions and processes more frequently than attributes, and they used more derivation than agnation to do that. Lastly, measures of both types of complexity revealed that the students'

writing does not really change after grade III. This is interesting because the data shows a drastic change in measurable complexity after grade I and a slight change after grade II, whilst the values obtained at grades III and IV are practically equal.

The obtained results should, of course, be observed through the lens of the limiting factors affecting this study, namely the observation period, approximate proficiency level, genre and topic homogeneity, and the size of the corpus. As this study analyzed texts written by different, individual students at each high-school grade, the language development occurring in later grades relative to earlier ones is assumed. Furthermore, the analyzed texts were selected from the school archives of student-written work rendering it impossible to evaluate with certainty students' English language proficiency, as students themselves were not personally involved in the study. In addition, some studies (Beers & Nagy, 2009; Yang et al., 2015) have demonstrated that the characteristics of syntactic complexity vary with genre and topic. As the present study used authentic texts produced by high-school students for their regular written assessment task, it was not possible to secure the texts of the same genre written on the same topic. Finally, the number of texts analyzed should be higher. The small number of texts, along with high variation within each group, was likely a major reason why some of the statistical findings were weak or inconclusive.

The above-mentioned limitations invite further research. Thus, ideally, a study of this type should be longitudinal, tracking the writing development of a select group of participants over a certain period in the course of their education. This select group of participants should have their English language proficiency measured using a standardized proficiency test and they should all produce texts of the same type and on the same topic. Lastly, the sample of student's texts should be large enough so as to allow for more accurate findings. All things considered, the general impression is that an analysis such as this can indeed yield practical insights into a language learner's writing development. This should especially be true for the language teacher who might use such insights to judge if their instruction is generally having the desired effect; however, this is not to say that such analysis can or should be used in any form of proficiency testing or grading because the results of such a quantitative text analysis are suggestive at best.

8. Pedagogical implications

The additional linguistic awareness, offered by the analysis of linguistic complexity, may be the main benefit for a language teacher who can then use such knowledge to be more precise in identifying a learner's potential

area of improvement. However, because such analysis also provides insights into the characteristics of high-quality or advanced writing, some studies have also pointed out the benefits of practically applying such knowledge to writing instruction in the language classroom.

Mazgutova and Kormos (2015) found that a one-month intensive English for Academic Purposes course did have an impact on the lexical and syntactic characteristics of university level L2 English writing. The course did not focus specifically on syntax and vocabulary but included writing practice, written feedback, and individual writing tutorials where participants received suggestions for further improvements (Mazgutova & Kormos, 2015: 6). The results were promising since the learners' writing showed more syntactic characteristics of academic writing, such as increased noun phrase complexity, increased use of complex nominals and nominalizations, especially in the lower proficiency group (Mazgutova & Kormos, 2015: 10).

It would be interesting to see how high-school students respond to formal writing instruction that draws attention to this rather abstract and noun-heavy style of academic writing, especially as previous research has shown, at least in the case of college-level L2 German speakers (Byrnes, 2009), that such writing instruction is effective. Not only that, but the students who attended such classes adopted GM as nominalization to such an extent that they would even get carried away and overuse it, thus making the writing nearly unintelligible (Byrnes, 2009: 63). Now, while the goal is certainly not to train students to produce complex gibberish, it might be beneficial to draw students' attention to these types of structures and their effects on meaning-making.

Personally, while working on the present study, I (the lead author) started to increasingly pay attention to the syntactic structure of my high-school students' writing. More specifically, I decided to spend more time than usual covering noun modification and subordination with one second-grade classroom prior to their obligatory writing test. The tasks were fairly simple, and they included description tasks. The students were instructed to avoid coordination of independent clauses where possible. For example, describing the day they said goodbye to their best friend, one student wrote the sentence I LOOKED OUT THE WINDOW, AND SHE WAS STANDING NEXT TO THE CAR, I would ask "where was the window in relation to you?", "how big was the window?", or "what can you say about the car?". The student would then come up with the additional information and I would help them integrate it into the text, hopefully leading to a sentence such as WHEN I LOOKED OUT THE SMALL WINDOW TO THE LEFT, I SAW HER NEXT TO THE CAR THAT WOULD TAKE HER TO HER NEW HOME. Many of these exercises produced comically complex sentences as a result, but, more often than not, I was left with the impression that many of the more proficient students were using those modifying struc-

tures more frequently, and, at times, even skillfully which had a positive effect on the quality of the students' written work.

References

- Ai, Haiyang, & Xiaofei Lu (2010). *A web-based system for automatic measurement of lexical complexity*. Paper presented at the 27th Annual Symposium of the Computer-Assisted Language Consortium (CALICO-10). Amherst, MA. June 8–12.
- Beers, Scott F., & William E. Nagy (2009). Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre? *Read Writ* 22: 185–200.
- Biber, Douglas, Bethany Gray, & Kornwepa Poonpon (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly* 45(1): 5–35.
- Byrnes, Heidi (2009). Emergent L2 German writing ability in a curricular context: A longitudinal study of grammatical metaphor. *Linguistics and Education* 20: 50–66.
- Cohen, Jacob (1988). *Statistical Power Analysis for the Behavioural Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Cooper, Thomas C. (1976). Measuring written syntactic patterns of second language learners of German. *The Journal of Educational Research* 69(5): 176–183.
- Halliday, M. A. K., & Christian M.I.M. Matthiessen (2014). *Halliday's Introduction to Functional Grammar* (4th ed.). London, New York: Routledge.
- Hunt, Kellogg W. (1965). *Grammatical Structures Written at Three Grade Levels*. Campaign, IL: National Council of Teachers of English.
- Hunt, Kellogg W. (1970). Do sentences in the second language grow like those in the first? *TESOL Quarterly* 4(3): 195–202.
- Jahić, Alma (2011). An Insight into Textual Borrowing Practices of University-Level Students in Bosnia and Herzegovina. *Journal of Academic Writing* 1(1): 173–179.
- Lu, Xiaofei (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics* 15(4): 474–496.
- Lu, Xiaofei (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal* 96(2): 190–208.
- Lu, Xiaofei, Haiyang Ai (2015). Syntactic Complexity in College-level English Writing: Differences among Writers with Diverse L1 Backgrounds. *Journal of Second Language Writing* 29: 16–27.
- Mazgutova, Diana, & Judit Kormos (2015) Syntactic and lexical development in an intensive English for Academic Purposes Programme. *Journal of Second Language Writing* 29: 3–15.
- Norris, John M., & Lourdes Ortega (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics* 30(4): 555–578.
- Ortega, Lourdes (2003). Syntactic complexity measures and their relationship to L2 Proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics* 24(4): 492–518.

- Ortega, Lourdes (2015). Syntactic complexity in L2 writing: Progress and expansion. *Journal of Second Language Writing* 29: 82-94.
- Ryshina-Pankova, Marianna (2015). A meaning-based approach to the study of complexity in L2 writing: The case of grammatical metaphor. *Journal of Second Language Writing* 29: 51-63.
- Ryshina-Pankova, Marianna, & Heidi Byrnes (2013). Writing as learning to know: Tracing knowledge construction in L2 German compositions. *Journal of Second Language Writing* 22: 179-197.
- Vyatkina, Nina, Hagen Hirschmann, & Felix Golcher (2015). Syntactic modification at early stages of L2 German writing development: A longitudinal learner corpus study. *Journal of Second Language Writing* 29: 28-50.
- Yang, Weiwei, Xiaofei Lu, & Sara C. Weigle (2015). Different topics, different discourse: Relationships among writing topic, measures of syntactic complexity, and judgments of writing quality. *Journal of Second Language Writing* 28: 53-67.

Authors' addresses:

Amer Delić
Fatmić 2/19, 72270 Travnik
Bosnia and Herzegovina
e-mail: amer_d_86@hotmail.com

Alma Jahić Jašić
Faculty of Humanities and Social Sciences
University of Tuzla
Dr. Tihomila Markovića 1
75000 Tuzla
e-mail: alma.jahic@untz.ba

Received: February 19, 2019

Accepted for publication: March 15, 2019