

Research Paper

Investigating the Patterns of Syntactic Complexity Predicting High-Quality Writing: A Corpus-Based Study of the Written Text Production at the B2+ English Language Examination at a Hungarian University

Karlygash Adamova¹

Recommended citation:

Adamova, K. (2025). Investigating the Patterns of Syntactic Complexity Predicting High-Quality Writing: A Corpus-Based Study of the Written Text Production at the B2+ English Language Examination at a Hungarian University. *Central European Journal of Educational Research*, 7(1), 76–90. <https://doi.org/10.37441/cejerr/2025/7/1/15027>

Abstract

This pilot study explores the predictive role of syntactic complexity in assessing L2 writing proficiency, with a focus on its potential contribution to validating a high-stakes English language examination. Drawing on prior research that highlights the importance of syntactic complexity in writing evaluation, the study aims to identify specific syntactic measures that reliably distinguish between low-rated and high-rated L2 texts. The analysis is based on a corpus of written texts for the B2+ level for so-called 'Basic' English Language Examination (BLE) administered at a Hungarian university. Although labeled "Basic" the BLE represents a mandatory proficiency examination (B2+ level according to the CEFR) required for academic advancement. Rather than examining inter-rater reliability, the research centers on contribution to validation by investigating linguistic features associated with rated writing quality. Grades assigned by human raters were used solely to group texts and build the corpus for analysis. A total of 60 syntactic complexity indices were extracted using the Multidimensional Analysis Tagger (MAT) (Nini, 2019) and the Coh-Metrix 3.0 software package (Graesser, McNamara & Kulikowich, 2011). These indices include measures of clausal, phrasal, and overall structural complexity. The findings are expected to inform ongoing validation efforts for the BLE and contribute to more robust, evidence-based practices in L2 writing assessment by identifying linguistic patterns that correlate with writing proficiency.

Keywords: L2 writing assessment; syntactic complexity; writing proficiency; corpus-based analysis; language test validation

Introduction

A growing body of literature recognizes the importance of syntactic complexity in analyzing L2 learners' writing. Numerous studies suggest that syntactic complexity indices are among the key elements characterizing writing proficiency and quality (e.g. Ai & Lu, 2013; Crossley & McNamara, 2014; Taguchi, Crawford & Wetzel, 2013; Lu, 2017). The current study seeks to explore syntactic complexity of L2 writing and attempts to identify the predictive patterns of syntactic complexity for writing quality characterized by the syntactic complexity variables.

The structure of the pilot study is led by two research questions, starting the investigation from identifying syntactic complexity measures characterizing the low-rated and high-rated texts, followed by the investigation of the more specific syntactic complexity measures predicting high-quality writing as viewed by human evaluators. Altogether, the results are expected to add to the investigation of the large-scale project of the validation process of the 'Basic' English Language examination at a Hungarian university (Adamova, 2021; Reményi, 2024).

¹ Doctoral School of Linguistics, Pázmány Péter Catholic University, Budapest, Hungary; adamova1403@gmail.com

Two research questions lead the piloting study and are addressed to the investigation of patterns of syntactic complexity predicting high-rated and low-rated writing as measured by the syntactic complexity variables and as judged by human assessors, respectively: (1) *What are the predictive patterns of syntactic complexity for writing quality characterized by the syntactic complexity variables of high-rated and low-rated texts?* (2) *What are the predictive patterns of syntactic complexity for writing quality characterized by score levels assessed by human ratings?*

Literature review: Syntactic complexity in L2 research

Syntactic complexity is widely recognized as a valuable measure for analyzing second language (L2) proficiency (Ortega, 2003; Bulté & Housen, 2012). According to the taxonomic model of L2 complexity of Bulté and Housen (2012), syntactic complexity represents a fundamental component of absolute complexity and consists of three developmental levels: sentence, clausal, and phrasal. This multimodal construct serves as an essential indicator of L2 learners' proficiency and performance, reflecting varied dimensions of linguistic ability (Housen et al., 2012).

Biber, Gray and Staples (2016) demonstrates a foundational perspective on the relationship between syntactic complexity and language proficiency, distinguishing between spoken and written modes. Their findings reveal that clausal complexity characterizes spoken language, while phrasal complexity is indicative for written texts. Specific phrasal (e.g. nominalizations, prepositional phrases, attributive adjectives, and noun phrase modifiers) and clausal features (e.g. passive verbs and that-clauses) were identified as frequent components of high-proficiency written discourse.

Crossley and McNamara (2014) and Taguchi, Crawford and Wetzel (2013) further established similar findings, identifying phrasal complexity as a notorious aspect of high-quality writing. Their outcomes highlight the role of phrasal-level complexity in distinguishing highly rated from lower-rated texts.

Mazgutova and Kormos (2015) determined an additional measure portraying complexity, which is the syntactic structure similarity index. This indicator acts as one of the predictors of syntactic richness in L2 writing through evaluating similarity in syntactic structures production. Reduction of similar features determines the rise of variety and thus, the degree of syntactic richness.

In addition to the establishment variables, the introduction of methodological approaches is equally significant. Two primary approaches to analyze syntactic complexity have been identified in the literature: the traditional T-unit-based approach and Multi-Dimensional Analysis (MDA). The first mentioned method focuses on "traditional holistic measures" (Biber et al., 2016, p. 652) and corroborate length of production units as significant predictors of high-quality writing.

According to Biber, Gray and Staples (2016), and Biber et al. (2020), while a traditional T-unit based framework facilitates the analysis of L2 performance and development, it lacks sufficient attention to the functional characteristics of syntactic complexity features in the texts. To address these limitations, Biber (1992) introduces an alternative approach – Multi-Dimensional Analysis. MDA investigates co-occurrence of the individual grammatical features and their functions in texts. This approach also implements examination of distribution of grammatical features across genres, registers and tasks. Therefore, the MDA framework allows for a more comprehensive analysis of L2 learners' performance by integrating both structural and functional dimensions of complexity.

Taken together, the current study combines the presented traditional T-unit approach and the MDA. By implementing both structural and functional variables, this research intends to indicate whether the existing measures of high-quality writing in the L2 literature can serve as indices of the B2+ level of proficiency in specific assessment contexts. The selected comprehensive approach addresses the gap in existing literature and tends to provide valuable insights to the process of the validation of the language Basic English Examination at a Hungarian university.

Research methods

Research setting

This pilot study examines the case of the English language examination at a Hungarian university, which stands as an obligatory assessment procedure for all English majors. This high-stakes examination is administered at the end of students' first year of their studies and aims to determine whether they have reached the B2+ level of English proficiency according to the Common European Framework of Reference for Languages (Council

of Europe, 2001). The CEFR is an international framework developed by the Council of Europe, which provides a standardized approach to language proficiency assessment. Six levels of proficiency are outlined in the CEFR (i.e. A1, A2, B1, B2, C1, C2), each of them corresponding to a particular level ranging from basic to proficient. For instance, B2 level is indicated as upper intermediate level and is attributed to ability of producing detailed and complex texts, to engage fluently in various interactions, express thought effectively (Council of Europe, 2001). B2+ in the current context refers to the level of language proficiency expected to be performed, which is beyond B2 (upper-intermediate) level.

The so-called 'Basic' English Examination (BLE) has significant implication for students, as passing is mandatory to be able to continue their academic journey. Despite its name, "Basic" does not correspond to proficiency level, rather, it denotes a mandatory examination that plays a vital role in determining students' academic progress.

The BLE is a sit-in high-stakes examination, which consists of four parts: the Use of English, reading comprehension, written text production, and spoken components. The present study is focused exclusively on its written text production part. For this part of the examination, it is required to write a 180–200-word text in 45 minutes on one of three offered topics that can be selected by students. The texts correspond to three genres (a review, a formal letter, or a narrative), the latter changing each examination. The genre of each examination varies, and the database of the present study was developed on an examination occasion where the genre assigned was a formal letter. Specifically, the students were asked to write a formal letter on one of three topics, including a letter about a degree program, or about available summer camp jobs, or about a language school.

The assessment of the written component of the BLE is built on an analytic rating scheme. Students' texts are evaluated based on four criteria: (1) task achievement; (2) coherence and cohesion; (3) grammar, and (4) vocabulary. Each of the above-mentioned competencies is scored on a scale from one to five points by the human raters. The scoring rubric for the writing section of the BLE is structured around descriptors assigned to each score level, with evaluation criteria specifically developed for this examination. Since the present study is aimed at investigating the syntactic complexity analysis of the written text production, the grammar facet of the scoring rubric is taken into consideration.

To provide clarity, the grammar aspect in the rating scheme evaluates range, complexity and accuracy of grammatical structures. A scoring rubric includes five descriptions assigned to each of the points on a five-score scale (Alapvizsga Writing, n.d.). A score of one (1) point is given to the texts containing no range of structures and showing frequent inaccuracies. There is no specific descriptor for two (2), but to get this point students are required to perform at the transitional level between one and three. A score three (3) indicates a limited range of grammatical structures, with frequent inaccuracies that may significantly disrupt communication. A descriptor for four (4) indicates a performance showing a good range of structures with occasional inaccuracies. Conversely, a writing demonstrating a wide range of structures with minor inaccuracies that do not disrupt communication is scored with five (5) points.

The written component of the BLE is evaluated by two independent blind reviewers (raters), who are typically university professors or lecturers actively engaged in teaching preparatory English language courses. In order to pass this section, students are required to obtain at least 40% of the total available points, equivalent to a minimum of 16 out of 40. Initially, each rater assigns a score independently. These preliminary evaluations are then followed by a discussion between the two raters to determine a consensus score for each written script. In cases when two raters are unable to reach a consensus, a third rater is brought into the process to ensure the assignment of a fair and reliable grade.

It is important to note that reliability analysis was not the focus of this study. Rather than evaluating the consistency of scoring across evaluators, the primary aim was to explore the validity of the examination in terms of its linguistic construct, particularly syntactic complexity. The assigned grades were not treated as dependent variables for measuring inter-rater reliability; instead, they served solely as a means of constructing a representative corpus of student writing. Texts were grouped based on these grades to facilitate the identification of potential predictive patterns in syntactic complexity that could support the validation of the test's construct.

Data collection

The database for the analysis has been created from the existing L2 learner corpus of the May 2017 examination papers (Radnai, 2017; Adamova, 2021), of which 14 were selected for this piloting study. Thus, to meet the aims of the current research, the selected texts were tailored in a small sub-corpus.

The corpus is designed to align with the objectives and research questions of the present study. Corpus design and development contain several steps to create a systematic sampling of the data, including selection of representative texts, classifying them based on specific grammar points, and ensuring the data was prepared for analysis. Following ethical considerations, all texts were anonymized and were labeled as H1-H7 for high rated texts and L1-L7 for low rated ones. Furthermore, considering implementation of computational tools for data analysis, the accuracy of the data was manually checked to ensure their accuracy for the automated analysis. Altogether, a corpus of 3121 words was created for this pilot analysis, providing a solid foundation for the syntactic complexity analysis.

Data analysis

The methodological approach adopted in this study is a quantitative method with the implementation of the Multidimensional Analysis Tagger (MAT) (Nini, 2019), also known as the Biber-tagger, as the major research tool for analyzing and tagging the syntactic complexity variables in the texts. In addition, the Coh-Metrix 3.0 software package was also employed to investigate one of the selected grammatical measures. Coh-Metrix is an automated tool that facilitates the analysis of discourse and language elements on different levels, including coherence and cohesion (Graesser, McNamara & Kulikowich, 2011). Specifically, this study includes the analysis of the syntactic structure similarity index evaluated with the help of the Coh-Metrix. For data analysis, SPSS Statistics software is used to process and analyze the collected data.

Introduction to key variables

The MAT is a program created as a multidimensional analysis tool to examine English texts (Nini, 2019). Following Biber's (1988) approach, the current analysis involves the set of 67 variables first represented in his study. For research perspectives, the full set of linguistic features were classified into sixteen sub-groups adopting Biber's (1988, p. 73-75) method of allocation within grammatical and functional categories (Appendix 1).

The decision on the implementation of linguistic features for the computational analysis was mainly based on their reliability, which had been proved by previous studies, including Biber (1988), Biber (1992), Biber, Gray and Poonpon (2011), and Biber, Gray and Staples (2016).

The analysis revealed that seven out of 67 variables were not identified in any of the texts of the dataset. Furthermore, the texts were also manually checked on the occurrence of the linguistic features tagged with zero frequency by the MAT. Since those linguistic features occurred to be problematic and thus, it seemed to be difficult to distinguish by the computational tool, they were initially excluded from the analysis.

The list of grammatical features excluded from the analysis comprise, for instance, hedges, contracted forms, discourse particles (e.g., the words now, anyhow, anyways), that relative clauses in subject position, past participial clauses, split infinitives and the words seem/appear. A possible explanation for these results might be related to the priority of the students that was given to other constructions within this context. The choice of the appropriate lexical items and syntactic structures is apparently determined by the genre of formal letters within which the written texts were produced. This includes, for instance, contracted forms, which are inappropriate for the formal register.

Coh-Metrix is another computational tool implemented to the analysis. Aiming to analyze linguistic and discourse features of texts, it provides insights into cohesion, coherence, and complexity (McNamara et al., 2014). It enables researchers to examine various aspects of syntactic complexity by measuring sentence structure, word frequency and cohesion.

To investigate L2 writing complexity, the syntactic structure similarity index has been involved in the current research. This index evaluates the degree of variation in syntactic patterns used by learners, with lower similarity scores indicating a more diverse range of structures. A decrease in syntactic similarity suggests greater syntactic flexibility, reflecting ability to employ varied sentence constructions.

Altogether, the present study is focused on 60 textual and syntactic features, syntactic structure similarity index and grammar points in order to investigate syntactic complexity analysis of the high and low rated writing. Implementation of diverse range of grammatical features tend to provide a more nuanced understanding of L2 learners' writing proficiency beyond traditional length-based measures.

Results

Descriptive statistics and t-test: Research question 1

The grammatical features portrayed above serve as the key variables examined in this study. To begin, patterns indicative of writing quality was identified. To explore this aspect, descriptive statistical measures, including the mean and standard deviation, were utilized. Additionally, an independent samples t-test was conducted to determine whether there were statistically significant differences in the distribution of linguistic features between the texts that assessed with the highest grammar scores (H1-H7) and those with the lowest grammar scores (L1-L7).

Descriptive statistics and t-test

Table 1. presents only the statistically significant findings obtained from the descriptive statistical analysis, comparing the two groups: texts that received the highest and lowest grammar scores.

Table 1. Mean values and statistically significant differences for selected variables

Name of the variable	Mean values in the texts with the highest marks	Mean values in the texts with the lowest marks	t-test results (Levene's Test for Equality of Variances)
Agentless passives	.687	.399	.007*
By passives	.111	.000	.001*
Concessive adverbial subordinators	.000	.074	.034*
Conditional adverbial subordinators	.827	.544	.036*
Emphatics	.757	.463	.044*
Present participial clauses	.069	.000	.034*
Present participial WHIZ deletion relatives	.071	.000	.034*
Pronoun it	.704	1.207	.021*
Subordinator that deletion	.000	.233	.000**
Syntactic similarity structure	.088	.057	.006*
Synthetic negation	.137	.000	.000**
That verb complement	.064	.000	.034*
Wh relative clauses on object position	.000	.141	.000**
Grammar points	=5	<5, =1.5-2	

(*p < .05, **p < .01)

The outcomes indicate that the writing of high and low rated students differ significantly in the distribution of specific linguistic features. Higher-rated examination papers exhibited a significantly greater presence of passives, conditional adverbial subordinators, emphatics, that-verb complements, agentless passives, present participial clauses, synthetic negation, and present participial WHIZ deletion relatives. This suggests that these linguistic structures are characteristic of more proficient writing.

Contrastively, the lower-rated texts demonstrated a preference for use of concessive adverbial subordinators, pronoun it, subordinator that deletion, and WH-relative clauses in object position. These grammatical features appear to be more frequent in less advanced writing, highlighting their potential as linguistic indices of lower-rated writing.

The table also demonstrates the statistically significant ($p = .006^*$) difference in the number of syntactically similar structures with mean values of $m = .088$ for higher-rated texts and $m = .057$ for lower-rated ones. Notably, this finding suggests that students with stronger grammar scores tend to use more syntactically similar structures in their writing.

Non-parametric correlation analysis (Spearman's coefficient): Research question 2

The second research question of this study addresses the predictive patterns of writing quality as judged by human raters. To explore this, the correlation analysis was performed to investigate the relationship between the linguistic features and the grammar points. Given the nature of the data, a non-parametric correlation analysis using Spearman's coefficient was used to identify the predicting patterns of the B2+ level characterized by the grammar points given by the human raters. Correlation analysis employed to the study identifying the extent to which the frequency of selected grammatical features correlates with the human ratings.

Non-parametric correlation analysis (Spearman's coefficient)

The findings from the correlation analysis reveal a strong correlation between the ratio of other adverbial subordinators ratio and the grammar points ($r=.589$, $p=.027^*$). Some illustrative cases from the corpus highlight this connection: "Since not only am I interested in assisting at camp, but I also intend to explore the country while I am there." (from the H1 text); "I am writing my letter to apply for your cultural exchange program, since I am planning to spend a month and work in the USA." (from the H3 text); "Besides free time courses could you tell me please what other study options Erasmus students have while their stay at your university?" (from the H7 texts). These significant correlations suggest that the frequency of adverbial subordinators had an influence on the score, meaning an increase in the use of these subordinators is associated with higher grammar points.

None of the other measures showed a statistically significant correlation with grammar scores. Overall, the findings presented in this chapter suggest that among all variables examined, other adverbial subordinators exhibit the strongest and most significant positive correlation with grammar scores. The following chapter will provide a detailed discussion and interpretation of these results, followed by the conclusion.

Discussion and conclusion

This chapter, which synthesizes and interprets the findings of the study, is guided by the research questions of the present study. The major purpose of the current study was to contribute to the investigation of syntactic complexity analysis of the written texts at the BLE. The first research question of the study sought to determine indices of high and low rated writing, as measured by the syntactic complexity variables.

The results presented in Chapter 4.2 demonstrate significant statistical differences in the distribution of 12 linguistic measures between texts with the highest and lowest grammar scores. Descriptive statistics indicate that high-rated texts are characterized by specific grammatical features, such as: by passives, conditional adverbial subordinators, emphatics, that verb complements, agentless passives, present participial clauses, synthetic negation and present participial WHIZ deletion relatives.

Conversely, lower-rated texts are associated with linguistic features, including concessive adverbial subordinators, pronoun it, subordinator that deletion and wh- relative clauses on object position. These features, identified in the analysis, are found to be indicative of lower syntactic complexity, as performed by L2 writers. Importantly, the interpretation of these results must consider contextual factors - particularly the genre of the writing task and the scoring rubric applied during assessment. Since the task required the production of a formal letter, genre conventions inevitably influenced the structures that appeared. Formal letters demand a polite, respectful, and professionally appropriate tone. As a result, the students might tend to favor certain grammatical choices, such as passive constructions and conditional clauses, because these forms help writers express requests indirectly, present information impersonally, and maintain a formal register.

For example, the frequent use of conditionals demonstrates not only grammatical sophistication but also an awareness of the interpersonal conventions typical of formal written communication. Likewise, passive voice allows writers to focus on actions rather than agents, a strategy that aligns with the impersonal tone expected in formal correspondence.

On the other hand, over reliance on features like the pronoun it or deletion of complementizers, which was common in lower-rated texts, might reflect a more casual style that does not meet the formality expectations of the task. These features might simplify sentence structure but may reduce precision or weaken the tone, which more likely plays a role in lower assessments of writing quality.

The results also highlight an important finding: genre conventions can limit the range of syntactic structures. More precisely, formal letters often follow a predictable structure - including openings, polite requests, explanations, and closings, which can be also used repeatedly. Interestingly, this means that a lower variety of

syntactic structures does not necessarily indicate lower proficiency. Instead, it may reflect strategic choices that optimize clarity, politeness, and appropriateness for the communicative context.

This observation aligns with the present results, which show that high-rated texts, despite sometimes demonstrating less syntactic variety, still achieve higher scores. This suggests that examiners value controlled, context-appropriate complexity over structural diversity. It also implies that assessments of writing should not equate complexity solely with variety, appropriateness, accuracy, and functional use are equally important. The results further confirm previous findings in the literature. Studies such as Grant and Ginther (2000) and Ferris (1994) have reported that features like adverbial subordination and passive voice are markers of advanced writing. Similarly, Ferris (1994) emphasizes that frequent use of complex syntactic structures, such as present participial clauses and that-verb complements, reflects more mature writing. Moreover, the results from one of the previous piloting studies also revealed that passive constructions are among the indices of higher-rated writing.

The analysis also examined the similarity of syntactic structures across texts. High-rated texts demonstrated greater structural consistency, which suggests that more proficient writers rely on controlled patterns of complexity rather than experimenting widely with different forms. As Mazgutova and Kormos (2015) observed, advanced learners do not necessarily use a wider range of structures but instead achieve sophistication through careful, accurate use of selected forms.

Another noteworthy point is that the correlation between syntactic complexity and writing quality was weaker than expected. Although previous research (e.g., Taguchi, Crawford & Wetzel, 2013; Crossley & McNamara, 2014) has shown that complexity measures at the phrasal level are strongly linked to writing quality, the present study found only one feature with a statistically significant correlation with grammar scores. A possible reason for this is the BLE scoring rubric, which appears to reward effective and appropriate use of grammar rather than the presence of specific complex forms. Similar conclusions were identified by Biber, Gray and Staples (2016), who reported that TOEFL iBT scores were not strongly linked to syntactic complexity, likely due to the weighting of accuracy over complexity in scoring criteria.

From a broader perspective, this study has important implications for language testing field. If assessments focus too heavily on accuracy, they may fail to capture a learner's ability to use complex structures in meaningful, context-appropriate ways. The findings reveal that BLE stakeholders might consider revising their scoring rubrics to evaluate not only correctness but also the range, appropriateness, and communicative effectiveness of syntactic choices. In the case of formal letter writing, rubrics could explicitly reward the use of linguistic strategies that support politeness, precision, and clarity.

The importance and originality of this study is that it investigates the patterns of syntactic complexity predicting high-quality writing as measured by the syntactic complexity variables in the frame of the writing part of the BLE. However, it must be noted that the unexpected outcomes might lay in a major limitation of the current piloting study, which is a relatively small sample size. For this piloting stage, 14 texts, including seven texts with the highest grammar points and seven texts with the lowest grammar points, were selected for the analysis. Given these findings, future language assessments should consider the potential impact of genre on syntactic complexity. Furthermore, scoring rubrics could be refined to better capture the full spectrum of syntactic competence, considering not only the range and accuracy of structures but also the effective use of complex, more precise language that may be overlooked in traditional assessments focused solely on accuracy. In conclusion, the findings of this study offer valuable insights into how syntactic complexity interacts with genre, scoring criteria, and writing quality in L2 assessment. They show that complexity is not merely a matter of structural variety but is deeply shaped by communicative context and rhetorical purpose. For tasks like formal letter writing, success depends not just on demonstrating grammatical knowledge, but on deploying that knowledge strategically - using syntactic structures to fulfil the functional, stylistic, and interpersonal demands of the genre. Future language assessments should therefore integrate genre-sensitive approaches into their rubrics creation and consider complexity as a multidimensional construct that includes not only form but also appropriateness, precision, and communicative effectiveness.

Acknowledgement: We thank Johnathan Dabney for the English language editing.

References

- Adamova, K. (2021, November 18–19). *Patterns of syntactic complexity predicting high-quality writing: A corpus-based study of the written text production at the Basic English Language Exam at the Hungarian University* [Conference abstract]. Szeged, Hungary.
- Ai, H., & Lu, X. (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. In A. Díaz-Negrillo, N. Ballier & P. Thompson (Eds.), *Automatic treatment and analysis of learner corpus data* (pp. 249–264). John Benjamins.
- Alapvizsga writing. (n.d.). Alapvizsga writing: Descriptors for the exam [Manuscript]. Pázmány Péter Catholic University.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge University Press.
- Biber, D. (1992). The multi-dimensional approach to linguistic analyses of genre variation: An overview of methodology and findings. *Computers and the Humanities*, 26(5–6), 331–345.
- Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, 45(1), 5–35. <https://doi.org/10.5054/tq.2011.244483>
- Biber, D., Gray, B., & Staples, S. (2016). Predicting patterns of grammatical complexity across language exam task types and proficiency levels. *Applied Linguistics*, 37(5), 639–668. <https://doi.org/10.1093/applin/amu059>
- Biber, D., Gray, B., Staples, S., & Egbert, J. (2020). Investigating grammatical complexity in L2 English writing research: Linguistic description versus predictive measurement. *Journal of English for Academic Purposes*, 46. <https://doi.org/10.1016/j.jeap.2020.100869>
- Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 21–46). John Benjamins.
- Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26, 42–65. <https://doi.org/10.1016/j.jslw.2014.09.005>
- Conrad, S., & Biber, D. (2001). *Variation in English: Multi-dimensional studies*. Pearson.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press.
- Crossley, S. A., & McNamara, D. S. (2014). Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing*, 26, 66–79. <https://doi.org/10.1016/j.jslw.2014.09.006>
- Ferris, D. (1994). Lexical and syntactic features of ESL writing by students at different levels of L2 proficiency. *TESOL Quarterly*, 28(3), 414–420. <https://doi.org/10.2307/3587446>
- Graesser, A., McNamara, D., & Kulikowich, J. M. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational Researcher*, 40(5), 223–234. <https://doi.org/10.3102/0013189X11413260>
- Grant, L., & Ginther, A. (2000). Using computer-tagged linguistic features to describe L2 writing differences. *Journal of Second Language Writing*, 9(2), 123–145. [https://doi.org/10.1016/S1060-3743\(00\)00019-9](https://doi.org/10.1016/S1060-3743(00)00019-9)
- Greenbaum, S., & Quirk, R. (1995). *A student's grammar of the English language*. Longman.
- Haegeman, L. (1994). *Introduction to government and binding theory* (2nd ed.). Blackwell.
- Housen, A., Kuiken, F., & Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurement and research. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 1–20). John Benjamins.
- Lu, X. (2017). Automated measurement of syntactic complexity in corpus-based L2 writing research and implications for writing assessment. *Language Testing*, 34(4), 493–511. <https://doi.org/10.1177/0265532217710675>
- Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for academic purposes programme. *Journal of Second Language Writing*, 29, 3–15. <https://doi.org/10.1016/j.jslw.2015.06.004>
- McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511894664>
- Nini, A. (2014). *Multidimensional Analysis Tagger (Version 1.1)* [Computer software manual]. <https://sites.google.com/site/multidimensionaltagger/>
- Nini, A. (2019). The Multi-Dimensional Analysis Tagger. In T. Berber Sardinha & M. Veirano Pinto (Eds.), *Multi-dimensional analysis: Research methods and current issues* (pp. 67–94). Bloomsbury Academic.
- Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24(4), 492–518. <https://doi.org/10.1093/applin/24.4.492>
- Reményi, A. A. (2024). Evaluating written L2-English learner texts through manual and automated quantitative analysis. In A. Dósa, Á. Magnuczne Godó, R. L. Nagano, & A. Schäffer (Eds.), *Crossing boundaries: Space, identity and discourse in Anglophone studies*.
- Taguchi, N., Crawford, W., & Wetzel, D. Z. (2013). What linguistic features are indicative of writing quality? A case of argumentative essays in a college composition program. *TESOL Quarterly*, 47(2), 420–430. <https://doi.org/10.1002/tesq.91>

Appendix

Table 2. Classification and description of linguistic variables tagged with MAT

Variable	Description	Examples
a. tense and aspect markers		
Past Tense	“As most commonly used, the past tense combines two features of meaning: (a) The event/state must have taken place in the past, with a gap between its completion and the present moment. (b) The speaker or writer must have in mind a definite time at which the event/state took place” (Quirk et al., 1985, p. 183)	“I also heard...” (L4)
Perfect Aspect	“Aspect is a grammatical category that reflects the way in which the meaning of a verb is viewed with respect to time” (Greenbaum & Quirk, 1995, p. 51)	“After I have thoroughly examined your website...” (H5)
Present Tense	“The terms PRESENT TENSE and PAST TENSE have this justification: that the tenses they name typically have reference to present and past time respectively” (Quirk et al., 1985, p. 176).	“I am a student...” (H2)
b. place and time adverbials		
Place Adverbials	“Any item in this list: aboard, above, abroad, across, ahead, alongside, around, ashore, astern, away, behind, below, beneath, beside, downhill, downstairs, downstream, east, far, hereabouts, indoors, inland, inshore, inside, locally, near, nearby, north, nowhere, outdoors, outside, overboard, overland, overseas, south, underfoot, underground, underneath, uphill, upstairs, upstream, west” (Nini, 2014, p. 18)	“I want to make wonderful memories abroad...” (H5)
Time adverbials	“Any item in this list: afterwards, again, earlier, early, eventually, formerly, immediately, initially, instantly, late, lately, later, momentarily, now, nowadays, once, originally, presently, previously, recently, shortly, simultaneously, subsequently, today, to-day, tomorrow, to-morrow, tonight, to-night, yesterday” (Nini, 2014, p. 18)	Looking forward to having your early reply (H2)
c. pronouns and pro-verbs		
First Person pronouns	“Any item of this list: I, me, us, my, we, our, myself, ourselves” (Nini, 2014, p. 18)	I hope to receive your answers as soon as possible (H1).
Second person pronouns	“Any item of this list: you, your, yourself, yourselves, thy, thee, thyself, thou” (Nini, 2014, p. 18)	I am entirely thankful for your help (H1).
Third person pronouns	“Any item of this list: she, he, they, her, him, them, his, their, himself, herself, themselves” (Nini, 2014, p. 18)	“...programmes for staff members on their day off” (H1)
Pronoun it	Any occurrence of a pronoun “it, its and itself” (Nini, 2014, p. 18)	“...or is it my responsibility to cover the cost of my flight and accommodation (H1)”.
Demonstrative pronouns	“The words those, this, these followed by a verb, or auxiliary verb, or a punctuation mark, or a WH	These are the questions for which I seek answers (H5).

	pronoun, or the word and (Nini, 2014)”	
Indefinite pronouns	Pronouns that “lack the element of definiteness which is found in the personal, reflexive, possessive, and demonstrative pronouns, and to some extent also in the wh-pronouns” (Quirk et al., 1985, p. 376)	As for the tasks, will we get a list on which everything has to be done, or can we choose from several duties that we would like to be responsible for? (H2)
Pro-verb do	“Any form of DO that is used as main verb” (Nini, 2014, p. 19)	“...so I am really looking forward to hearing about all the activities I will be able to do in my free time”. (H5)
d. questions		
Direct WH-questions	“Any punctuation followed by a WH word and followed by any auxiliary verb (modal verbs, forms of do, forms of have or be) (Nini, 2014, p. 19)”	What do these organise programs involve? (L4)
e. nominal forms		
Nominalizations	a structure with a noun head, which is related morphologically to a verb or an adjective (Quirk et al., 1985, p. 1288).	I would also like to ask about the requirements of these courses. (H4)
Gerunds	“The program tags as gerunds any nominal form (N) that ends in -ing or -ings (Nini, 2014, p. 20)”	For instance, visiting museums or outdoor activities, like hiking. (H4)
Total other nouns	Any noun that has not been tagged as nominalization or gerund (Nini, 2014, p. 20)	I hope to receive your answers as soon as possible. (H1)
f. passives		
Agentless passive	One of the two following patterns are tagged as agentless passives: “(1) any form of be followed by a participle plus one or two optional intervening adverbs or negations; (2) any form of be followed by a nominal form and a participle” (Nini, 2014, p. 20)	“Are airplane tickets provided...” (H4)
By-passives	The program tags any agentless passive constructions followed by the preposition BY as by-passives (Nini, 2014, p. 20)	As I mentioned, I am really amazed by the British culture... (H3)
g. stative forms		
Be as main verb	Be, functioning as a main verb in patterns, like: “be followed by a determiner, or a possessive pronoun, or a preposition, or an adjective” (Nini, 2014, p. 20)	“I am happy to participate in any kind of camp” (L4)
Existential there	“There of existential sentences differs from there as an introductory adverb in lacking stress, in carrying none of the locative meaning of the place-adjunct there, and in behaving in most ways like the subject of the clause” (Quirk et al., 1985, p. 1405)	Is there any chance for me to working in a public or a private camp? I also would like to ask about the available tasks (L1).
h. subordination features		
That adjective complement	The category of complementation of an adjective by a that-clause (Quirk et al., 1985, p. 1220).	“...but I am aware that there are just as many expectations towards me.” (H7)
That verb complement	The program tags the following pattern: that is (1) preceded by and, nor, but, or, also or any punctuation mark and followed by a determiner, a pronoun, there, a plural noun or a proper noun; (2) preceded by a public, private or suasive verb or a form of seem or appear and followed by any word that is NOT a verb,	In terms of available tasks I have seen that I can be a counselor at the camp... (H2)

	auxiliary verb, a punctuation or the word and; (3) preceded by a public, private or suasive verb or a form of seem or appear and a preposition and up to four words that are not nouns.	
WH-clauses	“Any public, private, or suasive verb followed by any WH word, followed by a word that is not an auxiliary” (Nini, 2014, p. 21)	I haven’t decided where to go. (L4)
WH relative clause on object position	“Any word that is not a form of the words ASK or TELL followed by any word, followed by a noun, followed by any word that is not an adverb, a negation, a verb or an auxiliary verb” (Nini, 2014, p. 23)	The man who Sally likes (Nini, 2014)
WH relative clauses on subject position	“Any word that is not a form of the words ASK or TELL followed by a noun, then a WH pronoun, then by any verb or auxiliary verb, with the possibility of an intervening adverb or negation between the WH pronoun and the verb” (Nini, 2014, p. 23).	Is there any trip for the Erasmus students which shows us the city of Edinburgh? (H4)
Infinitives	The Program finds all the occurrences of “to followed by a subordinator, a cardinal number, a determiner, an adjective, a possessive pronoun, WH words, a pre-determiner, a noun, or a pronoun and tags the remaining instances of to as infinitive markers” (Nini, 2014, p. 21)	To explore (H2)
Past participial WHIZ deletion relatives	The following tag is assigned to WZPAST: “a noun or a quantifier pronoun followed by a past participial form of verb followed by a preposition or an adverb or a form of be” (Nini, 2014, p. 22)	“...free-time programs organized for the workers” (L1)
Past Participial clauses	The structures when participial forms of verb occur in subject position, which is -ed verb form (Quirk et al., 1985)	“Given these characteristics, it is not surprising that...” (Conrad, S. & Biber, D., 2001, p. 18-19).
Present participial clauses	The structures when participial forms of verb occur in subject position, which is -ing verb form (Quirk et al., 1985)	“Screaming with rage, he ran up the stairs” (Conrad, S. & Biber, D., 2001, p. 18-19)
Present participial WHIZ deletion relatives	“Present participial form of a verb preceded by a noun” (Nini, 2014, p. 22)	I kindly ask for a detailed description including the things I have just mentioned (H7).
That relative clause on object position	“The occurrences of that preceded by a noun and followed by a determiner, a subject form of a personal pronoun, a possessive pronoun, the pronoun it, an adjective, a plural noun, a proper noun or a possessive noun followed by a genitive marker” (Nini, 2014, p. 22).	“...or can we choose from several duties that we would like to be responsible for?” (H2)
That relative clauses on subject position	“The occurrence of that preceded by a noun or followed by an auxiliary verb or a verb” (Nini, 2014, p. 22)	The dog that bit be (Nini, 2014)
That verb complement	The program tags the following pattern: that is (1) preceded by and, nor, but, or, also or any punctuation mark and followed by a determiner, a pronoun, there, a plural noun or a proper noun; (2) preceded by a public, private or suasive verb or a form of seem or appear and followed by any word that is NOT a verb,	In terms of available tasks I have seen that I can be a counselor at the camp... (H2)

	auxiliary verb, a punctuation or the word and; (3) preceded by a public, private or suasive verb or a form of seem or appear and a preposition and up to four words that are not nouns.	
Pied-piping relative clauses	“The phenomenon where the preposition is moved along with the complement NP is referred to as pied-piping” (Haegeman, 1994, p. 375)	Are there any obligatory programs in which the workers have to take part? (L2).
Sentence relatives	When a punctuation mark is followed by the word which (Nini, 2014, p. 23)	Are the workers paid per hour or is there a general number, which is fit? (L4)
Causative adverbial subordinators	The program tags “any occurrence of the word because” (Nini, 2014, p. 23)	It is a very essential information to me because it does matter what I am going to do every day for a month (L5).
Concessive adverbial subordinators	Concessive adverbial subordinators are depicted as although and its more formal variant though (Quirk et al., 1985, p. 1097).	Although I do not have much working experiences, I am intend to gain some during the summer (L1)
Sentence relatives	When a punctuation mark is followed by the word which (Nini, 2014, p. 23)	Are the workers paid per hour or is there a general number, which is fit? (L4)
Conditional adverbial subordinators	The program tags “any occurrence of the words if and unless” (Nini, 2014, p. 23)	I would like to improve my English language skills and would like to know if you can offer appropriate courses at my level (H3).
Other adverbial subordinators	“This tag identifies any occurrence of the words: since, while, whilst, whereupon, whereas, whereby, such that, so that (followed by a word that is neither a noun nor an adjective), such that (followed by a word that is neither a noun nor an adjective), inasmuch as, forasmuch as, insofar as, insomuch as, as long as, as soon as” (Nini, 2014, p. 24)	Since not only am I interested in assisting at camp, but I also intend to explore the country while I am there (H1).
i. prepositional phrases, adjectives and adverbs		
Total prepositional phrase	“This tag identifies any occurrence of the prepositions listed by Biber (1988) under this category” (Nini, 2014, p. 24)	“...provided with flight tickets.” (H2)
Attributive adjectives	One of the functions of an adjective premodifying a noun (Quirk et al., 1985)	“the highest standard”, “cultural knowledge” (H1).
Predicative adjectives	adjectives functioning as subject (1), or object complement (2) are regarded as having a predicative function (Greenbaum & Quirk, 1990, p. 132)	Is it essential to write a language level test before entering the courses? (H3)
Total adverbs	All the adverbs in the text	“entirely thankful” (H2)
j. lexical classes		
Conjuncts	Linguistic features that are related to “the speaker's comment in one quite specific respect: his assessment of how he views the -connection between two linguistic units”. (Quirk et al., 1985, p. 632)	moreover (L1)
Downtoners	The program tags the words “almost, barely, hardly, merely, mildly, nearly, only, partially, partly, practically, scarcely, slightly, somewhat” as downtoners (Nini, 2014, p. 26)	only (L1)

Hedges	Words used to express a “tentative meaning” (Quirk et al., 1985, p. 1113).	Maybe, at about, something like, etc. (Nini, 2014)
Amplifiers	The following words are tagged as amplifiers: “absolutely, altogether, completely, enormously, entirely, extremely, fully, greatly, highly, intensely, perfectly, strongly, thoroughly, totally, utterly, very” (Nini, 2014, p. 26)	greatly (H1)
Demonstrative pronouns	“The words those, this, these followed by a verb, or auxiliary verb, or a punctuation mark, or a WH pronoun, or the word and (Nini, 2014)”	These are the questions for which I seek answers (H5).
Emphatics	The following words are assigned to emphatics: “just, really, most, more, real+adjective, so+adjective, any form of do followed by a verb, for sure, a lot, such a” (Nini, 2014, p. 26).	really (H3)
Demonstratives	“A demonstrative is found when the words that, this, these, those have not been tagged as either DEMP, TOBJ, TSUB, THAC, or THVC” (Nini, 2014, p. 27)	Could you inform me about these activities? (L1)
k. modals		
Necessity modals	The words must, should, ought (Nini, 2014, p. 27)	Should we find them uploaded on the site of the University of Edinburgh? (L6)
Possibility modals	The program tags the following modals as possibility modals: “can, may, might, could” (Nini, 2014, p. 27)	“...I have not found any information as to what types of camps I may work at”. (H1)
Predictive modals	The program tags the following modals as possibility modals: “will, would, shall and their contractions” (Nini, 2014, p. 27)	I would like to know more about the courses available and the knowledge they offer. (H2)
I. specialized verb classes		
Private verbs	“This tag finds any of the items listed by Quirk et al. (1985, p.1181): accept, anticipate, ascertain, assume, believe, calculate, check, conclude, conjecture, consider, decide, deduce, deem, demonstrate, determine, discern, discover, doubt, dream, ensure, establish, estimate, expect, fancy, fear, fears, feel, find, foresee, forget, gather, guess, hear, hold, hope, imagine, imply, indicate, infer, insure, judge, know, learn, mean, note, notice, observe, perceive, presume, presuppose, pretend, prove, realize, reason, recall, reckon, recognize, reflect, remember, reveal, see, sense, show, signify, suppose, suspect, think, understand” and their forms (Nini, 2014, p. 28)	hope (H2)
Public verbs	“This tag finds any of the items listed by Quirk et al. (1985, p.1180). Among them, for instance, acknowledge, add, admit, affirm, agree, allege, announce, argue, assert, bet, boast, certify, claim, comment, complain, concede, confess, confide, confirm, contend, convey, declare, deny, disclose, exclaim, explain,	writing (H1)

	forecast, foretell, guarantee, hint, insist, maintain, mention, object, objects, predict, proclaim, promise, pronounce, prophesy, protest, remark, repeat, reply, report, say, state, submit, suggest, swear, testify, vow, warn, write” and their forms (Nini, 2014, p. 27)	
Suasive verbs	“This tag finds any of the items listed by Quirk et al. (1985, p. 1182). Among them, for instance, agree, allow, allows, arranges, ask, beg, command, concede, decide, decree, demand, desire, determine, enjoin, ensure, entreat, grant, insist, instruct, intend, move, ordain, order, pledge, pray, prefer, pronounce, propose, recommend, request, require, resolve, rule, stipulate, suggest, urge, vote” and their forms (Nini, 2014, p. 29)	ask (L5)
Seem/appear	“Any occurrence of the verbs seem and appear, and their forms” (Nini, 2014, p. 29)	Seem, appear (Nini, 2014)
m. reduced forms		
Split auxiliaries	auxiliary verb followed by one (or two) adverbs and a verb base form (Nini, 2014)	“I am always open...” (H1).
Split infinitives	“An infinitive marker to followed by one or two adverbs and a verb base form” (Nini, 2014, p. 30)	He wants to convincingly prove that... (Nini, 2014)
Stranded preposition	A deferred preposition standing in a post-verbal position (Quirk et al., 1985)	“...what types of camps I may work at.” (H2).
Subordinator that deletion	The program tags one of the following patterns as THATD: “(1) a public, private or suasive verb followed by a demonstrative pronoun or a subject form of a personal pronoun; (2) a public, private or suasive verb is followed by a pronoun or a noun and then by a verb or auxiliary verb; (3) a public, private or suasive verb is followed by an adjective, an adverb, a determiner or a possessive pronoun and then a noun and then a verb or auxiliary verb, with the possibility of an intervening adjective between the noun and its preceding word” (Nini, 2014, p. 30)	Please let me know if there is any additional information about courses or cultural activities (that) I should know about (L6).
Contractions	“Phonologically reduced or simplified forms which are institutionalized in both speech and writing.” (Quirk et al., 1985, p. 123)	N’t (Nini, 2014)
n. coordination		
Independent clause coordination	“The word and in one of the following patterns: (1) preceded by a comma and followed by it, so, then, you, there + BE, or a demonstrative pronoun or the subject forms of a personal pronouns; (2) preceded by any punctuation; (3) followed by a WH pronoun or any WH word, an adverbial subordinator or a discourse particle or a conjunct” (Nini, 2014, p. 30)	“My name is XY, and I am about to hand in my application to your company...” (L4)
Phrasal coordination	Any occurrence of and “preceded and followed by the same tag and when	“flight tickets and accommodation” (H2)

	this tag is either an adverb tag, or an adjective tag, or a verb tag or a noun tag” (Nini, 2014, p. 30)	
o. negation		
Analytic negation	The occurrence of the word not and contracted form n’t followed by an adverb (Nini, 2014).	Not only am I...”; “Unfortunately, I have not found...” (L2)
Synthetic negation	“no followed by any adjective and any noun or proper noun” (Nini, 2014)	“I have no idea...” (H4).



© 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) licence (<http://creativecommons.org/licenses/by/4.0/>).