Englisia: Journal of Language, Education, and Humanities
November 2023. Vol.11, No.1, 1-18

The development of syntactic complexity in university students' essays

Agus Ferani Istiharoh*, Dwiyanto Djoko Pranowo

Universitas Negeri Yogyakarta, Indonesia

Manuscript received July 11, 2023, revised August 29, 2023, accepted September 18, 2023, and published online November 7, 2023.

Recommended APA Citation
Istiharoh, A. F., & Pranowo, D. D. (2023). The development of syntactic complexity in university students' essays. *Englisia: Journal of Language, Education, and Humanities*, *11*(1), 1-18. https://doi.org/10.22373/ej.v11i1.19129

ABSTRACT

This research aimed to describe quantitatively the development of students' syntactic complexity during the Essay Writing Course measured with the L2SCA tool by Lu. The research basically implemented a descriptive-quantitative research design. The primary data were collected from 20 students majoring in Agribusiness and taking an essay writing course in Yogyakarta Muhammadiyah University (UMY) during one semester from February to July 2022 in a researcher class. They wrote three essays gradually from one session to the third session of the class. The essays were analyzed quantitatively using the L2SCA tool by following the indices of the tools as the reference to see the students' syntactic complexity. In addition, the participants wrote three essays divided into three times writing tests. Overall, the analysis method was conducted by following a systematic procedure; classifying, scoring, displaying, describing, interpreting, discussing, and concluding. As a result, this study reveals that syntactic development of students' essays during a semester session of writing course fluctuated from the first writing until the third writing. Although along with the writing supervision, they had consultation sessions, the result did not show gradual and constant progress. It implied that the students' acquisition of syntactic complexity represented in writing did not improve progressively in all syntactical indices. However, among the three essays the participants wrote, they excel most in the second essay or argumentative essay. Further, it is necessary to relate the essay's genre to the result of students' writing for each of the essays has a special characteristic and level of difficulty.

* ***Corresponding Author:***

Agus Ferani Istiharoh
Universitas Negeri Yogyakarta
Jl. Colombo Yogyakarta No.1, Karang Malang, Kec. Depok, Kabupaten Sleman, DIY 55281, Indonesia
Email: agusferani@gmail.com

Keywords: *Second language writing; Syntactic complexity; Essay; L2SCA*

1. Introduction

Syntactic complexity of second language writing has been considered as an interesting discussion in second language proficiency (Lu, 2010). As the inseparable section of writing, this discussion contributes to the enhancement of text meaning and it further impacts the reading comprehension. It is mentioned in Gustin (2019) that in the past, the stereotype of advanced students was that the more complex the sentences are, the more advanced the language proficiency they have (Gustin, 2019). However, complex sentences cannot be considered as the only consideration for assessing the quality/level of writing. Mostly, the students are trapped by the perception that the more complex the sentence they write, the more proficient they are. As a productive skill, writing quality is the representation of students' understanding of syntax, grammar, and another linguistic basis. Generally, a complex sentence is equal to good quality of writing. However, the study revealed that the writing quality is based on its syntactic complexity, which is not always complex in the sentence form.

There are some reasons why writing becomes so important in SLA development. First, being considered as a productive skill, it is important for learners to write correctly. Some linguistics aspects as mentioned above are required to master in order the messages and ideas delivered accurately. Therefore, the writings can be easily read and understood by the readers. Secondly, according to the survey of Human Development Index (HDI) reported on OECD (2018), Indonesia's reading performance mean score is 371 out of 487 (which means still in low-level). Reading performance score here does not only measure the ability and quantity of readings of learners, but also understanding, using and reflecting the written texts as means to have goals achieved, to gain the insights and potentials, and to be involved in society (OECD, 2018). As the reading performance is related to writing skill, the data above reveals that Indonesia fell back to its 2001 (371) level after a peak in 2009 (402). In addition, according to International Science Ranking in Scimago Journal and Country Ranking, Indonesia is on 45th place among 240 countries and on the 10th place among other Asiatic regions (China, Japan, India, South Korea, Taiwan, Hongkong, Malaysia, Singapore, Thailand, and Indonesia) (scimagojr.com, 1996-2020). Those statistical facts show that in terms of ranks among other countries, the quality and quantity of Indonesian learners/researchers writing need serious concern.

As an inseparable part of Second Language Acquisition (SLA), writing is considered as one of skills that is to measure the development or ability of learners in acquiring a second or foreign language. It relates to the concept that when learners have the ability to understand L2, they can easily express and deliver their thoughts and ideas in the form of written texts. As mentioned by Homstad and Thorson, "Writing has commonly been viewed as a support skill, used to reinforce the acquisition of grammar, as in the grammar-translation method, or to support the memorization of CO1TeCt

language structures, as in the audio-lingual method" (Homstad & Thorson, 1994, p.1). In order to do so, they also need to be aware of syntactic rules and accuracy. Those two aspects are considered as two of three fundamental dimensions characterizing the L2 usage (Michel, 2017). To date back on the early tradition of measuring L2 performance, there were two primary strands; 1) The mean length of utterance (MLU) that is used as established index of development and 2) The classroom-based research characterizing the language use in term of Complexity, Accuracy and Fluency (CAF). Then, in recent years, CAF is used as the measurement of global proficiency. In relation to language proficiency, students mastering good writing skills obviously have many advantages. In academic fields, it helps them to write the scientific paper, article, or thesis. In other fields, writing skill could help them to communicate effectively in the form of written text.

Norris and Ortega (2009) mentioned that the primary reason for measuring learners' L2 writing skill by using CAF is to consider how and why language skills develop in specific learners and target languages in response to specific tasks, learning, and other stimuli, with respect to the details of rate of development, pathways, and final outcomes. In order to have good writing skill, learners need to pay attention to the triad of complexity, accuracy and fluency (CAF). Complexity refers to the size, redundancy, abundance, and variety of L2 performance (Housen & Kuiken, 2009). Accuracy is a measure of the language's intentional and error-free use (Wolfe-Quintero et al., 1998). Fluency in Michel (2017) refers to fluent, light, and eloquent speech with a limited number of pauses, delays, or paraphrases (Lennon, 1990; Chambers, 1997; Freed, 2000; Koponen & Riggenbach, 2000). When learners use many ranges of and more 'complex' grammatical structures and vocabularies, are able to produce the accurate utterance and less error, and are able to speak or write fluently, they can be considered as proficient L2 learners.

However, students at the college level still face difficulties in writing. It is due to the lack of knowledge of vocabulary or grammar. The lack of vocabulary could impact their writing's syntactic complexity, while the lack of understanding of grammar is related to their accuracy. According to Skehan (2009) there is correlation between complexity and accuracy and L2 knowledge systems (Michel, 2017). Similarly, as mentioned by Norris and Ortega (2009) that the evaluation on complexity and accuracy are considered to understand how L2 learners' language skills develop. In long term goals, Complexity and Accuracy are thought to be able to characterize different levels of L2 performance (Wolfe-Quintero et al., 1998).

In general, linguistic complexity is closely related to syntactic complexity as the dimension of SLA research (Ellis, 2009; Bulte & Housen, 2012; De Clercq & Housen, 2017). Syntactic complexity is not merely defined as the long and complex sentences or phrases. However, it is more than just using of the variety and diversity of grammatical structures, considering the length of the sentences, clauses, or phrases, and the degree of excellence of linguistics resources (Inigaki & Kim, Quitero, 1998; Quintero, Inigaki, & Kim, 1998; Ortega, 2003; Gustin, 2019). A syntactic complexity can be measured with

some measures. Most research investigated the essay with five common Syntactic Complexity Measures (SCMs): Length of Production Unit, Sentence Complexity, Subordination, Coordination, and Overall Sentence Complexity. These are still focusing on quantitative readability. Meanwhile, Frantz (2015) proposes a different idea that the quality of writing should also emphasize on more specific aspects; level of vocabulary, language features, organizational structures and the cognitive aspects. According to Dale-Chall and Fresch-Kincaid, the syntactic complexity which examines the sentence's length belongs to traditional readability measures. By other means, it is still weak on cohesion aspect (Frantz et al., 2015).

Development of L2 learners' writing relates to the second language learning especially on writing skill. Moreover, the majority of the writing in our students' personal and professional lives will be online (Godwin-Jones, 2018). Writing is an important communication skill and plays an important role in a second language (Chastain, 1988, Simin & Tavangar, 2009, as cited in Javadi-Safa, 2018). For most students, writing was seen as a difficult and challenging task. It is because L2 learners' writing involves cognitive and emotional factors. According to McLeod (1987), affective factors influence all the writing process (McLeod, 1987, Yavuz-Erkan & İflazoğlu-Saban, 2011, Javadi-Safa, 2018). Shortly, writing is a productive skill that involves emotional and linguistic factors.

Time-consuming and copious commenting on writing are obviously ineffective (Cotos, 2014). Therefore, AWE was developed and seen as the silver bullet for language and literacy development. Based on some previous studies, it has been proven effective to evaluate the writing quality. AWE can be used for summative and formative assessment (Hockly, 2019). The use of AWE for summative assessment is controversial. Meanwhile, Chapella, Warschauer, and Grimes (2008) mentioned that formative assessment can encourage learners' motivation in revising their writing. Besides providing the automatic grading toward the learners' writing, it is also cheaper and faster in process. It is beneficial to use this tool to evaluate as it is becoming more powerful and readily available (Li et al., 2015). Moreover, it also offers a grammatical and spelling check that can guide learners to correct their sentences (Wang et al., 2013).

Lu (2010) has explored fourteen different measures of syntactic complexity on second language writing. She found that the system developed with those fourteen measures reaches very high reliability from the corpus (the result is similar to the result of human annotation). The source of writing is from advanced second language learners. Then it is just right that the systems and the background of the learners are similarly in the same level (Lu, 2010). In relation with the background of learners, a study exploring the syntactic complexity of college-level writers with the different first language (L1) backgrounds shows that backgrounds play significantly on the syntactic complexity measures. From fourteen measures employed, there were three that show significantly different emergence (when backgrounds are ignored). However, when the writing is

grouped based on the same background of the first language, the patterns of second language writing are more varied (Lu & Ai, 2015).

To identify the problems, the researcher did the initial observations related to syntactic complexity on second language acquisition development. It was found that students were mostly confused about expressing their thoughts in a second language (L2). Secondly, there are many inaccuracies in using vocabulary in their writing. In addition, there are some students who prefer to write sentences with simple patterns to avoid inaccuracy or error. Shortly, to write some ideas, they write their ideas in separated simple sentences. Whereas, according to CAF, the proficiency on L2 is based on their L2 complexity, accuracy and fluency. Therefore, if the case is so, the process of assessing their development in L2 writing development further needs to consider the aspect of complexity and accuracy. This is not only to characterize their L2 writing development, but also to help evaluate the writing at what level their writing proficiency are, based on the essay's syntactic complexity and grammar accuracy.

In addition, the writing evaluation which involves hundreds of writings is so time-consuming. Therefore, further, the recommendation towards the use of AWE tools is necessary to be explored in order to find the most appropriate and helpful tool that goes in line with learners' objectives. This is in order to reveal that though there are many tools that can be used for measuring the writing quality and development in terms of their syntactic complexity, there is still no obvious research underlying a more standardized tool for measuring the triad CAF as mentioned above. Considering all of the backgrounds provided above, the researcher focused on:

1. How is the student's syntactical structure developed during one semester?
2. How is the students' syntactic complexity development measured with the L2SCA tool?

2. Literature review

2.1. Second language writing

Second Language Writing (SLW) has been considered an interdisciplinary field rather than a disciplinary field for about sixty years. However, recently it has become the meta-disciplinary field in Applied Linguistics and Second Language Research (Matsuda et al., 2003, Fujieda, 2006; as cited in Javadi-Safa, 2018). What is more interesting is that SLW can be a topic that relates to some other fields, such as TESOL in general (Matsuda et al., 2003, as cited in Javadi-Safa, 2018), curriculum design, reading-writing connections, technology-assisted writing, material design, etc (Fujieda, 2006, as cited in Javadi-Safa, 2018). In other words, SLW can collaborate with diverse topics.

The old ways of annotating the writing quality by humans mostly takes time. Besides, it was not efficient in terms of procedure. As the development of the discipline, several computational systems for analyzing automatically have been created. Coh-Metrix is one of online toolkits developed by Graesser, et al (2004) for measuring the text coherence (Lu, 2010). There is a D-level Analyzer which was developed by Lu (2009) to

analyze the deep syntactic parsing. Unfortunately, those systems are mostly proposed and employed in first language acquisition and psycholinguistic research. Considering the wide interest in second language development, there is a need to fulfill this gap, especially on the syntactic complexity of second language writing systematically.

Automatic Writing Evaluation (AWE) has been increasingly used in learners' writing. According to Shermis, Burstein, and Bursky (2013), it is defined as the system which is web-based that can provide the feedback and scores for learners' writing (Zhai & Ma, 2021). Liao (2016) added that to get the feedback and scores, the learners just need to upload the writing in AWE systems. Then, directly, feedback and scores appear right after it is uploaded. Wilson and Andrada (2016) also mentioned that the writing is assessed by the automatic rating engines that was developed by computational linguistics (Miranty & Widiati, 2021). Similarly, Marie defines AWE as the computer-generated system that can provide scoring and feedback (Stevenson & Phakiti, 2014). To sum up, with the advance technology and availability of the internet, the evaluation toward writing tasks can be done by applying AWE.

AWE was first known as Page Essay Grade (PEG) in the 1960s when Page Ellis developed Project Essay Grade (PEG). Shermis, Mzumara, Olson, and Harrington (2001) stated that it provided the multiple regression analysis to measure the quality of writing (Miranty & Widiati, 2021). Then, as it was widely used, many language testing institutions collaborated to use this (e.g ETS). In 1966, there was found Automated Essays Scoring (AES) where it became the sophisticated language processing technology that enabled scoring or e-rater (Li et al., 2015). Nowadays, many schools and college classrooms in the United States use AWE as the helping tool for evaluation (Stevenson & Phakiti, 2014).

2.2. Syntactic complexity development

Principally, Syntactic Complexity can be defined based on two scopes (Szmrecsanyi, 2004). The first definition is based on the scope of pure length, duration, and size of a unit. The second is based on the way it appeals to the notions independent of pure length, duration, and size. Cited from the article entitled "An Operationalizing Syntactic Complexity", the following examples would best explain the definition above.

(1) I was not there because I had to fill out all this.

(2) I did not do it, and the reason for this was that.

Those two examples have the same number of words; 11. It fits definition 1. However, when it comes to the comparison of the main, adverbial, or compound clauses of each example, it meets definition 2. Both definitions are more like completing one another. On one hand, it defines the number of words or phrases in a clause, and on the other hand, it defines How such a clause is enhanced (Szmrecsanyi, 2004).

In addition, syntactic complexity is defined as the manifestation in second language writing in which it explored how varied and sophisticated the grammar and units are

produced (Foster & Skehan 1996, Wolfe-Quintero et al. 1998, Ortega 2003, Lu, 2010). Further, Lu added that it is the syntactic variation and degree of sophistication that becomes the proof in second language writing (Lu, 2011). It is considered as the inseparable part of overall development in target language learning. There are many measures that have been proposed to quantify the length of units etc. What becomes of the notions of this is that the measures are different from the measures used in the first language development studies.

Bates and Goodman (1999) stated that learners acquire the basic syntactic structures before they reach the age of four. However, to master academic writing proficiency, it can take years of learning (Beers & Nagy, 2011). Therefore, syntactic complexity and grammar accuracy are two among some other tools to measure the learners' writing proficiency.

Indeed, it is necessary to use the best measure to get the valid and reliable measures. Meanwhile, there is a question that emerges on what ways these measures are valid and reliable as the indices of second language learners' developmental level and global proficiency in target language. Fortunately, there have been several studies conducted to answer and to prove validity and reliability of the measures. Cross-sectional studies and longitudinal research have been conducted (Lu, 2010). Lu suggested that it should consider a full range comparison to get the best syntactic measures.

Syntactic complexity is one of the rising studies nowadays. As the development of a computational tool to annotate the syntax, it offers effectiveness for evaluating the writing. It leaves the old ways that the writing was annotated manually one by one. It took time and energy. However, with the existence of Syntactic Complexity Analyzer, its annotation is like human annotation is more effective and efficient (Gustin, 2019).

2.3. L2SCA Tool by Lu

In relation to SCMs, L2SCA is the tool that can analyze some measures that are valid and reliable for measuring the syntactic complexity. It is a corpus tool that was developed by Prof. Lu in 2010. The result of human annotation and this tool are similar, so there is no need to doubt its analysis result. This tool can count the syntactic complexity features such as, the production unit length, coordination amount, subordination, phrasal of sophistication level, and the whole sentence complexity (Gustin, 2019). The further details (Lu & Ai, 2015) are presented in the following table:

Table 1
Syntactic complexity measures in L2SCA.

Type	Measure	Definition
Overall sentence complexity	1. Sentence complexity ratio	#of clauses / # of sentences
Length of production unit	2. Mean length of clause	#of words / # of clauses
	3. Mean length of sentence	#of words / # of sentences
	4. Mean length of T-unit	# of words / # of T-units
Amount of subordination	5. T-unit complexity ratio	# of clauses / # of T-units
	6. Complex T-unit ratio	# of complex T-units / # of clauses
	7. Dependent clauses ratio	# of dependent clauses / # of clauses
	8. Dependent clauses per T-unit	# of dependent clauses / # of T-units
Amount of coordination	9. Coordinate phrases per clause	# of coordinate phrases / # of clauses
	10. Coordinate phrases per T-unit	# of coordinate phrases / # of T-units
	11. Sentence coordination ratio	# of T-units / # of sentences
Degree of phrasal Sophistication	12. Complex nominals per clause	# of complex nominals / # of clauses
	13. Complex nominals per T-unit	# of complex nominals / # of T-units
	14. Verb phrases per T-unit	# of verb phrases / # of T-units

According to Lu, the analysis by using the tool L2SCA have similar results with the human annotation (Gustin, 2019). Hence, it is reported that they found the error analysis regarding the errors made by learners in their writing. The issues that are found mostly are such as the use of collocation, determiners, or agreements. Besides, it is common because of the punctuation. Those errors actually do not relate to the syntactic complexity analysis that is being discussed in this research.

3. Method

Based on the identification of research problems and objectives, the researcher considered implementing the quantitative research method. The development of students' writing is recognized and measured with numerical data. Although the main data is the essays, the final score of the essays is in the form of numerical data. As asserted by Bhandari (2020), this method involves the process of collecting and analysing numerical data. In addition, one of the quantitative methods is descriptive research where it describes the overall summary of the variables. This research was conducted in Muhammadiyah Yogyakarta University especially in the class that takes the Essay Writing course taught by the researcher herself. The essays collection was conducted from February 2022 to July 2022 from the Agribusiness Class consisting 20 qualified students.

The data collection followed some procedures since the students had to write three kinds of essays; Cause and Effect, Argumentative and Reaction Essay. In a pre-planned task, the students were given the explanation toward the essays' writing like the organization of writing, language focus and any other features related to it. Following this, the students were then assigned to write the essays. The essays' written in a range of 300-800 words per essay. Next, the essays were checked for their similarity rate in Turnitin. Then, after they passed the minimum similarity rate as much as 20%, they had to have consultation to get the direct feedback from the lecturer. Then, the final step is for students to submit. Here, the students submitted the essays as the final essays that were documented for research analysis. The three essays written by each student were noted on its development. Thus, the total essays that are assessed and analyzed are 60 essays.

There were some steps taken in processing data for syntactic complexity data collection. The first, the data was collected after the students did their submission by email. Fortunately, since the software for measuring syntactic complexity required the file in typed-form, then there is no need for re-adjustment than if the essays are hand-written. Then, the files were renamed with the format: essays type, students' name followed by the major. After that, it was put into a different folder based on their major and types of essays. Next, the data were processed one by one into the SCA through the website https://aihaiyang.com/software/l2sca/. However, before processing it, the files were changed into the plain text format first (.txt file). Finally, the software will process and perform the number of syntactic complexities that later can be interpreted and explained.

4. Findings and discussion

4.1. Syntactic complexity structure

The development of syntactic structures is represented in 9 indices. It is measured with the tool called L2SCA. Additionally, L2SCA is a web-based tool that counts the total number of indices. The result of that measure is notably similar to human annotation, so that it is unquestionable for its validity and reliability.

The data and result presented in the following tables is the result of the tool calculation. The indices belonged to syntactic structures are the number of words (W), total sentences (S), verb phrase (VP), clause (C), T-unit (T), Dependent Clause (DC), Complex T-Unit (CT), Coordinate Phrase (CP), and Complex Nominal (CN). To see the development of syntactic structures of each class, the descriptive statistic of each index is presented as follows:

Table 2
Number of syntactic structures of agribusiness students' essays.

Syntactical Structure Indices	∑ of Essay 1	∑ of Essay 2	∑ of Essay 3
W	6639	7517	6924
S	366	388	343
VP	855	997	892
C	655	684	700
T	379	414	368
DC	279	276	327
CT	195	198	206
CP	223	225	213
CN	797	900	900

The table above presents the result of Cause and Effect, Argumentative and Reaction Essays' syntactical structure written by the twenty Agribusiness Students (N=20). Cause and Effect essay was set to be the first essay written among three other essays because structurally and technically, this type of essay is easier than the other essays. The cause-and-effect essays are the type of essays that typically highlight the cause or effect of certain issues. The students were free to choose the issues of the essays and were just instructed to write it based on the structure, language features and technique of the essays that have been explained previously by the lecturer. Secondly, the students wrote the argumentative essays which was noted as the middle level difficulty of essay, and reaction as the hardest level of essay (among three types of essays instructed in the present research). Those essays were written in three months which follow the longitudinal study design as framed in the research framework design.

Based on the table above, the agribusiness students' essays produce 6.639 words in total. They produce 7.517 words in total on the second essay. There is an increasing number from essay 1 to essay 2. However, on essay 3 which is notably as the hardest one among the two others showed the decreasing number of total Words (W), Number of Sentences (S), Verb Phrase (VP), T-Unit (T), and Coordinate Phrases (CP). These indices showed the up-to-down development.

On the other hand, there are increasing numbers of clause (C) and complex T-Unit (CT) production from Essay 1, 2, and 3 as provided in the table above. In the first essay,

the students produce 655 clauses in total, 684 clauses on the second essays and 700 clauses on the third essays. Meanwhile, on the Complex T-Unit production, it showed 195, 198, and 206 T-Units. Although that is not a significantly increasing number, it can be assumed that they tried to make their writing more varied in terms of those two indices. Shortly, these two indices showed up-to-up development.

Apparently, there is a down-to-up result on the number of Dependent Clause (DC) productions that is 279 on the first essays, 276 on the second essays and 327 on the third essays. In addition, the result on Complex Nominal (CN) increases; from 797 to 900, 900. To sum up, the development of syntactic structures of Agribusiness students' essays is presented in the following chart.

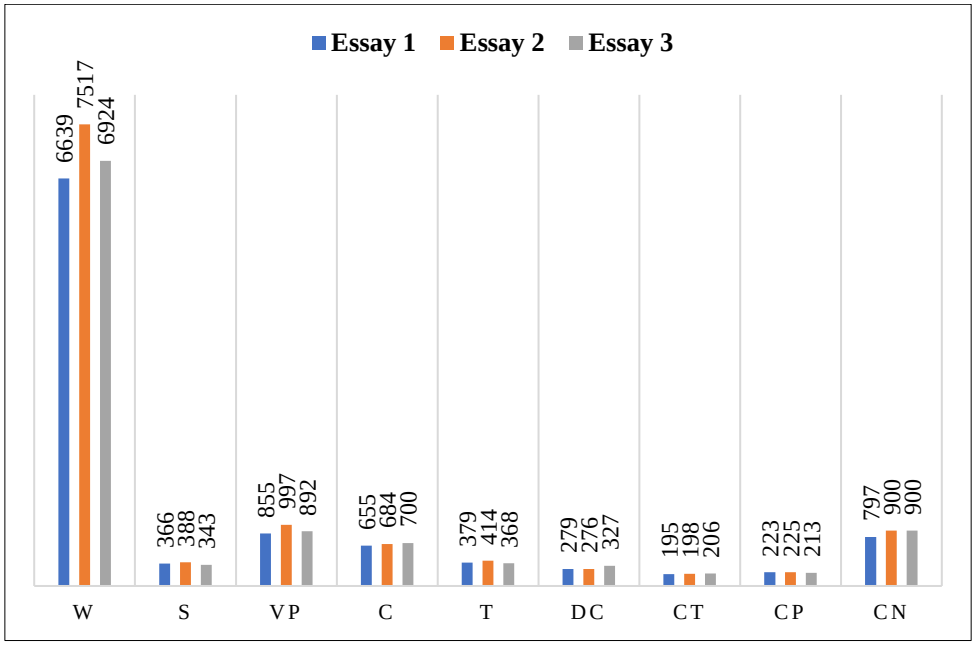

Figure 1. Syntactic structures development of agribusiness students' essays

The chart above shows the development of three essays that were written in a row from month to month of research data collection. The summary result above reveals the development of every single index measured to attain the syntactic structure of the essays. As seen above, most of the students of Agribusiness tend to be progressive on the second essays, argumentative essays. In contrast, they tend to show the low production of each index in essay 3, reaction essay. Based on the chart and summary above, it can be implied that Agribusiness students tend to produce more varied syntactic structures on argumentative essays, followed by cause-and-effect essays and reaction essays.

4.2. Syntactic complexity on SCMs

Sixty essays of the students were analyzed using the L2SCA tool to measure the syntactic complexity. The word development here refers to the timing of the writing where it is divided into three phases. The first phase is the writing of the first type of essay called Cause and Effect, which was written in March. The second one is the

Argumentative Essay, which was written in May, and the last is the Reaction Essay, which was written in June. The writing of the essays which were divided into three phases was proposed to see their development from time to time. Here, month to month of the writing process is considered as the ideal timing for this case of study, reconsidering that the leveling of the students' English class is determined and measured in one semester.

Therefore, considering those all, the development of the students' syntactic complexity was conducted by following the five types of measures that are known as overall sentence complexity, the length of production unit, the amount of subordination, amount of coordination, and degree of phrasal sophistication. Each of the measure types at least has one and or some ratio and means that define the measures itself. The following is the result from the L2SCA tool. It is displayed in five tables referring to the type of measures in syntactic complexity measurement.

Table 3

The syntactic complexity of agribusiness' essays.

Measure	**Code**	**Agribusiness**			**Progress**
		Essay 1	**Essay 2**	**Essay 3**	
Type 1. Overall Sentence Complexity					
Sentence Complexity Ratio	**C/S**	18,132	17,922	20,302	Down-to-up
Type 2. Length of Production Unit					
Mean length of clause	**MLC**	103,752	112,472	100,045	Up-to-down
Mean length of sentence	**MLS**	186,600	200,390	201,826	Up-to-Up
Mean length of T-unit	**MLT**	179,843	188,080	188,065	Up-to-down
Type 3. Amount of subordination					
T-unit complexity ratio	**C/T**	17,453	16,818	18,849	Down-to-up
Complex T-unit ratio	**CT/T**	0,515	0,496	0,559	Down-to-up
Dependent clauses ratio	**DC/C**	0,412	0,403	0,467	Down-to-up
Dependent clauses per T-unit	**DC/T**	2,952	3,261	5,198	Down-to-up
Type 4. Amount of coordination					
Coordinate phrases per clause	**CP/C**	0,348	0,348	0,315	Down-to-down
Coordinate phrases per T-unit	**CP/T**	2,219	1,549	1,078	Down-to-down
Sentence coordination ratio	**T/S**	8,784	9,508	10,319	Up-to-up
Type 5. Degree of phrasal sophistication					

Complex nominals per clause	**CN/C**	10,590	12,733	12,140	Up-to-down
Complex nominals per T-unit	**CN/T**	21,788	22,426	24,555	Up-to-up
Verb phrases per T-unit	**VP/T**	22,751	24,618	24,142	Up-to-down

First, the Overall Sentence Complexity Ratio counts the complexity of the total sentence in each essay. It is mentioned as the better discriminator to see the syntactic complexity in students' writing. The present study displays the result of sentence complexity ratio into two main divisions: agribusiness and medical school departments. It also includes the three essays that they wrote, respectively. In addition, since it consists of 60 essays per class, then it is summarized in the form of mean.

The table 2 above presents the result of syntactic complexity on the first type. It measures the sentence complexity ratio and codes with C/S. From the first to the third essay, there is no significant progress. As seen in agribusiness class, the result of C/S fluctuates from 18,132, 17,992 and 20,302. It does not show the rising progress from time to time. This indicates 'not so good' result because as the students wrote the more essays, their writing should be increasing from essay to essay as the writing has been supervised and commented on during the consultation session of the writing process. Besides, as there is no progressive result of the production, it implies that the more they write essays does not mean that it will increase the production of sentence complexity ratio as well.

Length of production is the second type of measure. It includes the measure of Mean Length of Clause (MLC), Mean Length of Sentence (MLS), and Mean Length of T-Unit (MLT). MLC result indicates the number of how many productions unit of clause length, MLS represents the mean of sentence length production and MLT represents the mean of T-Unit length production. The overall result showed that from essay 1 to 3, the mean is 103,752, 112,472, and 100,045. There is slightly little progress from essay 1 to 2 9,320. Next, the second measure on Length of Production Unit category is MLS. Yazdani (2018) mentions that it is one of the valid indicators to examine the syntactic complexity (Nur & Sulistyani, 2019). The result shows that the agribusiness class are progressive on MLS production. As in the first essay until the third essay, it increases from 186,600, 200,390, and 201,826. This indicates that the agribusiness class gets more understanding time by time in terms of producing the sentence with the lengthy and more varied structure. The last measure of the second category is the result of MLT. Nakamura (2019) states that MLT is the most extensive and valid measurement for measuring the syntactic complexity (Nur & Sulistyani, 2019). Based on the L2SCA tool, the result shows that the agribusiness class has improved from essay 1 to essay 2 (179,843-188,080). On the third essay, they have a decrease score from essay 2 to 3 (188,080-188,065), which is not such a significant decrease, or just about 0,015 and it is still

considered progressive. This indicates that an agribusiness class could produce the T-Unit length progressively from essay by essay.

Thirdly, the third category of syntactic complexity measure is the amount of subordination. It consists of four measures, such as T-Unit Complexity Ratio (C/T), Complex T-Unit Ration (CT/T), Dependent Clause Ratio (DC/C), and Dependent Clause per T-Unit (DC/T). Although the previous result of syntactical structure has been discussing the C/T, DC, etc., the following data is slightly different since this part examines its ratio instead of its syntactical structure. The first ratio is the T-Unit Complexity Ratio or C/T. In the amount of subordination, the agribusiness class development tends to be fluctuating from essay to essay. It can be seen from the scores in the table where the score shows down and up statistics. The next ratio is dependent clauses ratio or DC/C. Statistically the result above shows that the results tend to be not progressive from essay to essay (as seen in the table where the score shows the down-to-up flow). The agribusiness class scored 0,403 and 0,467 on the second and third essay. The gaps on the first and the third essay are 0,015 and 0,058, which are also not so significant. The last ratio on the amount of subordination category is the dependent clauses per T-unit or DC/T. The development of DC/T shows that the agribusiness class has increasing progress from essay 1 until essay 3. In addition, the Agribusiness class scores highly on the second and the third essay. In the third essay, the agribusiness class scored 5,198.

Next, the amount of coordination is the less syntactic structure predictor (Sulistyani, 2019). It covers the score of coordinate phrases per clause (CP/C), coordinate phrases per T-unit (CP/T), and Sentence coordination ratio (T/S). The first measure of the amount of coordination is the coordinate phrases per clause or CP/C. On this measure, the agribusiness class shows the fluctuating result where there is no increasing score from essay 1 to essay 2, but it falls on the third essay as much as 0,033. The next measure of the amount of coordination category is coordinate phrases per T-Unit or CP/T. Just like in the development of CP/C on the three essays, the result of CP/T also does not show the increasing score or progressive result. As seen in the table above, the scores of agribusiness class tend to fall from 2,219, 1,549, and 1,078, respectively on essay 1 to 3. The last measure of the fourth category is the sentence coordination ratio or T/S. The agribusiness class shows the progressive result on three essays, that in a row it increases from essay to essay. The increasing score is not too significant, yet it is progressive. In a more detailed overview, the agribusiness class achieved the highest scores in the first and the third essay.

The last category of the measurement for the syntactic complexity is the degree of phrasal sophistication. It is important to note that the highly proficient writers mostly do not produce essays that are cohesive, but instead produce writing that are linguistically sophisticated (Crossley & McNamara, 2012). This is an important finding to predict the L2 learners' proficiency. This degree measures the complex nominal per clause (CN/C), complex nominal per T-Unit (CN/T) and Verb Phrases per T-Unit (VP/T). The first

measure is the complex nominal per clause or CN/C. The agribusiness class has fluctuating score which indicates that they are not positively progressive in term of CN/C. The agribusiness' scores from essay 1 to 3 in a row are, 10,590, 12,733, and 12,140. Statistically, it shows the up-to-down result. This can be implied that the agribusiness students who are the 4th semester students understand and familiar with the use of CN/C. Secondly, the result of complex nominal per T-unit or CN/T shows that agribusiness class has a positive progress, or by other means, it has the rising result from essay 1 to 3; 21,788, 22,426 and 24,555. This indicates that the students get more familiar from essay to essay in using the complex nominal on every T-unit they wrote. To sum up, the agribusiness class achieved a higher score on the 2nd and 3rd essay. The last measure of the phrasal sophistication degree is the verb phrases per T-unit or VP/T. The development of their VP/T shows that Agribusiness class has a positive progress from essay 1 to 2 (22,751 to 24,618) but it decreases not so significantly from essay 2 to 3 (24,618 – 24,142). Briefly, it shows the up-to-down progress. This can be inferred that agribusiness classes are more sophisticated in terms of the phrasal degree.

The L2 complexity in terms of syntactic development is very quantitative. It is not an abstract definition which is derived from the relative sense of the individual. That's why the Bulte and Housen (2014) have tried to provide as much as possible definition for the term complexity itself. It is necessary to emphasize that the L2 complexity of students' writing is the concrete instance of L2 proficiency because it performs a persons' competence and ability (Thomas, 1994, p.330). The L2 development relates to the L2 proficiency of the L2 student's production over time and at different times. Essay is one of the writing samples in the case of writing production to find out the syntactic development as the part of recognizing and measuring the learners' proficiency in L2 (Bulté & Housen, 2014). The manifestation of syntactic development of the present study is the analysis on the length production scores of linguistics units on syntactic level such as phrase, clause, sentence and T-unit. The increase of the linguistics unit production from time to time of essays' writing is the indication of the syntactic development itself. Crossley and McNamara (2014) found that by the end of the English program, the learners tend to produce more complex and longer phrases and subordinate clauses (Crossley & McNamara, 2012). This corresponds to the Norris and Ortega (2009) arguments arguing that at the early stages of SLA, the learners tend to use more clausal coordination. On the intermediate stage, the use of subordination is more dominant than the coordination production, or the use of coordination diminishes (Norris & Ortega, 2009).

Other researcher also argues that the subordination is the indicative of syntactic complexity researchers have traditionally examined subordination as indicative of syntactic complexity in L2 writing (Ellis & Yuan, 2004; Kuiken, et.al, 2005; Kuiken & Vedder, 2008; Adams, et.al, 2015; Frear & Bitchener, 2015; Ruiz-Funes, 2015; Johnson, 2017). Then, at the advanced stages of L2 development, the syntactic complexification is no longer established through subordination but through clausal and phrasal

elaboration and or extended nominalization (Norris & Ortega, 2009). It implies that clausal and phrasal elaboration can suggest the level of L2 development at the more proficient level. Meanwhile, another study mentioned that among syntactic elaboration measures, MLS, CN/C, and C/T are generally three valid and reliable indicators of the writing proficiency of beginner and intermediate EFL learners (Bi & Jiang, 2020).

Interestingly, the present study found that the scores are not significantly increasing or decreasing. The scores are not so significant from essay 1 to 3, from time to time. Nevertheless, it indicates the progress no matter how small the progress is. Another surprising finding is that during the writing time which took more or less 6 months, the students wrote the three essays gradually with the three different essay types. Based on the study conducted by Byrnes (2014), L2 writing development performance also has association to the text genre/types (Johnson, 2017). As he/she studied, the subordination on narrative text is higher than on argumentative essay which mostly has higher phrasal-level features. When it is varied by genre, Beers and Nagy (2009) found that they do not find the differences on argumentative and narrative genres on words per clause. However, there are significant differences in words per clause between descriptive text and in argumentative text. Therefore, the variation by genre actually does not really correspond to the writing's syntactic complexity since it is hard to determine whether the scores appeared on the writing is the result from the text genre or grade level (Jagaiah et al., 2020).

Supporting Jagaiah et al, Abdi and Wang proposed a similar argument that topic familiarity, or specific prior knowledge about a particular topic, is correlated with the cognitive aspiration level of a task or task complexity (Robinson, 2011) or task difficulty (Skehan, 2014). This affects language performance (Abdi & Wang, 2022). In contrast, while other researchers mentioned some external factors other than linguistic factors influence the output of writing, Kang and Lee argued that the syntactic complexity measure results showed no significant effect for either plan type or task complexity (Kang & Lee, 2019). Following other researchers, Biber, Gray, and Poonpon (2011) sharing the similar arguments, it is agreed that the output of the syntactic complexity output should also be based on mode, genre, and communicative demand (Mazgutova & Kormos, 2015). Besides, it is also important to note that the ultimate goal of measuring the syntactic complexity is to recognize the students' writing ability as well as their proficiency in a second language. As also said by Hughes (1989) that the **"**The best way to test people's writing ability is to get them to write" (Qian et al., 2021).

5. Conclusion

This research aimed to describe the development of syntactic complexity of students' essays measured with the automated writing evaluation (AWE) tools; L2SCA. Based on the research formulations, the researchers concluded that agribusiness class mostly has fluctuating scores from essay 1 to essay 3. It does not show the gradual and constant increasing scores. It implies that the increasing scores or the progress is not only

based on the students' syntactic knowledge, but also on the familiarity toward the essays' type such as whether it is argumentative, reaction, or cause-effect essay. Each type of essay has a different level of critical thinking. Therefore, in the future, it needs more investigation on how big the gaps among the essay types are comparatively seen in the students' essays.

References

Abdi, M., & Wang, Y. (2022). Assessing linguistic complexity features in L2 writing: Understanding effects of topic familiarity and strategic planning within the realm of task readiness. *Assessing Writing*, *52*, 1–14. https://doi.org/10.1016/j.asw.2022.100605

Bi, P., & Jiang, J. (2020). Syntactic complexity in assessing young adolescent EFL learners' writings: Syntactic elaboration and diversity. *System*, *91*, 1–10. https://doi.org/10.1016/j.system.2020.102248

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

Cotos, E. (2014). Genre-based automated writing evaluation for l2 research writing.

Crossley, S. A., & Mcnamara, D. S. (2012). Predicting second language writing proficiency: the roles of cohesion and linguistic sophistication. *Journal of research in reading*, *35*(2), 115–135. Https://doi.org/10.1111/j.1467-9817.2010.01449.x

De Clercq, B., & Housen, A. (2017). A cross-linguistic perspective on syntactic complexity in l2 development: syntactic elaboration and diversity. *Modern language journal*, *101*(2), 315–334. Https://doi.org/10.1111/modl.12396

Frantz, R. S., Starr, L. E., & Bailey, A. L. (2015). Syntactic complexity as an aspect of text complexity. In *educational researcher* (vol. 44, issue 7, pp. 387–393). Sage publications inc. Https://doi.org/10.3102/0013189x15603980

Godwin-Jones, R. (2018). Second language writing online: An update. *Language Learning & Technology*, *22*(1), 1–15. https://doi.org/https://dx.doi.org/10125/44574

Gustin, S. (2019). Differences in syntactic complexity in the writing of el1 and ell civil engineering students differences in syntactic complexity in the writing of el1 and ell civil engineering students. (Issue June) [Portland State University]. https://pdxscholar.library.pdx.edu/open_access_etds

Hockly, N. (2019). Automated writing evaluation. *ElT Journal*, *73*(1), 82–88. Https://doi.org/10.1093/elt/ccy044

Jagaiah, T., Olinghouse, N. G., & Kearns, d. M. (2020). Syntactic complexity measures: variation by genre, grade-level, students' writing abilities, and writing quality. In *reading and writing* (vol. 33, issue 10). Springer Netherlands. Https://doi.org/10.1007/s11145-020-10057-x

Johnson, M. D. (2017). Cognitive task complexity and l2 written syntactic complexity, accuracy, lexical complexity, and fluency: a research synthesis and meta-

analysis. *Journal of second language writing*, *37*, 13–38. Https://doi.org/10.1016/j.jslw.2017.06.001

Kang, S., & Lee, J. H. (2019). Are two heads always better than one? The effects of collaborative planning on l2 writing in relation to task complexity. *Journal of second language writing*, *45*, 61–72. Https://doi.org/10.1016/j.jslw.2019.08.001

Li, J., Link, s., & Hegelheimer, V. (2015). Rethinking the role of automated writing evaluation (awe) feedback in ESL writing instruction. *Journal of second language writing*, *27*, 1–18. Https://doi.org/10.1016/j.jslw.2014.10.004

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International journal of corpus linguistics*, *15*(4), 474–496. Https://doi.org/10.1075/ijcl.15.4.02lu

Lu, X., & Ai, H. (2015). Syntactic complexity in college-level English writing: differences among writers with diverse l1 backgrounds. *Journal of second language writing*, *29*, 16–27. Https://doi.org/10.1016/j.jslw.2015.06.003

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

Miranty, D., & Widiati, U. (2021). Automated writing evaluation (awe) in higher education: Indonesian EFL students' perceptions about grammarly use across student cohorts. *Pegem egitim ve ogretim dergisi*, *11*(4), 126–137. Https://doi.org/10.47750/pegegog.11.04.12

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied linguistics*, *30*(4), 555–578. Https://doi.org/10.1093/applin/amp044

Qian, L., Yang, Y., & Zhao, Y. (2021). Syntactic complexity revisited: sensitivity of China's aes-generated scores to syntactic measures, effects of discourse-mode and topic. *Reading and writing*, *34*(3), 681–704. Https://doi.org/10.1007/s11145-020-10087-5

Stevenson, M., & Phakiti, A. (2014). The effects of computer-generated feedback on the quality of writing. *Assessing writing*, *19*, 51–65. Https://doi.org/10.1016/j.asw.2013.11.007

Sulistyani, U. N. L. (2019). *The analysis of syntactic complexity and grammatical accuracy in unisbank students' writing*. Diponegoro University.

Wang, Y. J., Shang, H. F., & Briody, P. (2013). Exploring the impact of using automated writing evaluation in English as a foreign language university students' writing. *Computer assisted language learning*, *26*(3), 234–257. Https://doi.org/10.1080/09588221.2012.655300

Zhai, N., & Ma, X. (2021). Automated writing evaluation (AWE) feedback: a systematic investigation of college students' acceptance. *Computer Assisted Language Learning*, *35*(9), 2817–2842. https://doi.org/10.1080/09588221.2021.1897019