Language Proficiency and Syntactic Complexity of Chinese EFL Writers: A Corpus-based Study

Zhiming Tang [1], Jingjie Cao [2*]

[1]International College, Fujian Agricultural and Forestry University, Fuzhou, Fujian, China

[2]Chengdu Technological University, Chengdu, Sichuan, China

*Corresponding Author.

Abstract:

This study examined the argumentative essays written by Chinese English major students studying at four different college levels. Statistical analysis of syntactic complexity indices showed that despite the effect size being minimum, significant development in majority of the measures was observed. In particular, at phrasal level as measured by complex nominals and verb phrases, there was a linear model of growth with learners at a higher proficiency outperforming the adjacent lower-level learners. These results seemed to confirm some of previous research showing a developmental trajectory of writing complexity moving from clausal to phrasal elaboration. Implications of these results for writing pedagogy were discussed.

Keywords: *syntactic complexity; Chinese EFL majors; argumentative writing; corpus-based study*

I. INTRODUCTION

Syntactic development has been an crucial constituent of second language acquisition and has extensively studied in recent decades. For applied linguists, the value of research into syntactic development is mainly threefold: to gauge performance, to describe proficiency, to benchmark development [1]. Insights into what and how linguistic features develop as learners' language instruction and exposure increases proves to be valuable for language education practitioners to adjust or improve pedagogy. Besides, the understanding on how proficiency and syntactic complexity are related provides basis for automated writing evaluation. Previous research has shown mixed results regarding relations between syntactic complexity indices and learner proficiency.

In this study, the focus of investigation is on the developmental pattern of various syntactic complexity measures as learners are exposed to more instruction and move to higher level of proficiency in Chinese EFL context. This study focuses on complexity at syntactic levels, trying to provide a profiling of the development of EFL college-level students in China as measured by a wide variety of indices which are selected based on previous research. The definition and operationalization of L2 complexity will be based on the framework established by Bulte and Housen [2]. And the computation of data will be conducted with the aid of two automated tools, Lexical Complexity Analyzer and Syntactic Complexity Analyzer by Lu[3-4]. The measures selected from these two automated tools will be aligned as much as possible with Bulte and Housen's framework.

II. DEFINITION AND OPERATIONALIZATION

Complexity is a multifaceted construct with several sub-constructs, dimensions, levels. In a comprehensive review, Bulté and Housen proposes a taxonomic model of L2 complexity making a distinction between absolute and relative complexity. Absolute complexity has to do with the inherent properties of language units and the system itself while relative complexity implies the difficulty level of processing or learning a language for individual learners. They made a further distinction between three kinds of absolute L2 complexity: propositional complexity, discourse-interactional complexity, and linguistic complexity. Of these three, linguistic complexity is what most studies have been focusing on.

Bulté and Housen specified two main subcomponents of linguistic complexity, namely the language system as a whole (system complexity) and individual linguistic features that make up such systems (structure complexity). The latter category of syntactic complexity can be observed at sentential, clausal, and phrasal level measured by computing the length and ratio of different structures in a discourse, which is probably the most popular approach to syntactic complexity study in the past years, though recently some scholars have paid attention to the role of multi-word units (i.e. phraseological) in the development of L2 writing proficiency[5-6].

III. L2 SYNTACTIC COMPLEXITY AND PROFICIENCY

Ortega argued that as the capacity to deploy the language resource mature overtime and with more instruction, syntactic complexity of writing increases[7]. Studies have consistently indicated that learners of higher proficiency are likely to produce longer syntactic structures as measured by mean length of T-unit [8-9].

In addition to length of production unit, clausal level complexity such as subordination and coordination as well as phrasal level complexity such as noun phrase elaboration show significant differences among learners of different proficiencies. Ai and Lu found discrepancies between non-native students of both low and high proficiency levels and that of native students in aspects such as length of production unit, amount of subordination and coordination, and degree of phrasal complexity[10]. Essays that score higher in complexity measures tend to be rated higher[11-12].

Clausal complexity has been considered an idiosyncratic feature of informal conversational register whereas phrasal complexification is typical of academic writing.[13-14]. Biber hypothesized that as L2 learners move to higher proficiency level, their writing would progress from a stage characterized by finite dependent clauses to one that is characterized by nonfinite dependent clauses and eventually to a phase that is characterized by dependent phrases[15]. Phrasal embedding is less explicit than clausal subordination but cognitively more complex[16]. Higher-rated writing samples seem to be correlated more with phrasal elaboration and nonfinite clauses than with clausal indices.

Norris and Ortega suggested a three-staged pattern of syntactic complexity development: at early stage of language acquisition, L2 writers achieve complexity through coordination of clauses and phrases, but as they advance to intermediate level, coordination diminishes, and complexity will be

mainly established through subordination. At advanced level, subordination subsides, and phrasal level complexification will be the pervasive means to achieve complexity[17].

But quite a few studies seem to suggest that the three-staged pattern of syntactic complexity development might apply better over a longer span in a wider and general L2 population than in the case of a particular cluster of L2 learners in a particular language program where the stages of development might not be as discrete as Norris and Ortega suggested.

Bulté and Housen found not a single of their various subordination measures changed significantly except a medium increase in sentence coordination and mean length of noun phrases[18]. On the contrary, in Crossley and MacNamara's study, L2 writers produced significantly fewer clauses including coordinated and subordinated clauses, smaller number verb phrases (which indicates fewer embedded clause), significantly greater phrasal modifications (longer noun phrases and increased number of words before the main verb)[19].

The above studies of intermediate L2 writers might suggest that even beyond intermediate L2 writers their syntactic complexification may still rely on coordination instead of subordination, and phrasal complexification does not necessarily have to occur after they reach an advanced level.

Findings from cross-sectional studies also seem to support that the three stages may partially overlap[20]. In Lu's study of EFL Chinese students' argumentative writing, writers of higher proficiency tended to produce significantly longer clauses and T-units, significantly higher ratio of complex nominals and coordinated phrases, but showed no significant difference in subordination measures. Similarly, Kim's study of EFL Korean college-level writers found that higher-proficiency writers wrote texts that are not only longer but also have more complex nominalizations[21].

The stable linear development observed in some study and the variation of development found in others may not be so surprising because it has been found that L2 complexity is modulated by various factors, such as L1 influence[22], input[23-24], native language use at work[25], task and genre [26-30] and the trade-off between lexicon and syntax[31]

And what's more, according to Complex Dynamic Systems theory, the non-linear developmental pattern in syntactic complexity is a normal part of development by virtue of the variation and instability which are essential features of dynamic systems[32]. The study by Verspoor et al. shows non-linear development in almost all specific linguistic constructions[33].

IV. RESEARCH QUESTIONS

The present study aims to investigate the following questions:

1) Do students of different proficiency differ significantly in terms of length of production unit, coordination, subordination, phrasal complexity?
2) What are the general developmental patterns of sentential, clausal, and phrasal complexity across different proficiency?

V. METHODOLOGY

5.1 Corpus

The data is extracted from Spoken and Written English Corpus of Chinese Learners (SWECCL 2.0) [34]. It has two sub-corpora: spoken and written. The present study choose to use the written section of the corpus. It has a total of 1.2 million tokens, consisting of 4950 essays on 26 argumentative topics written by either English-major students or non-English majors from a wide range of colleges in different cities and provinces of China. Students were free to select one of 26 topics to write about in timed or untimed condition.

With a sub-corpus generator, the corpus allows users to create sub-corpora according to different variables such as task types (timed and untimed writing), the year of entering university (2003, 2004, 2005, 2006, 2007), and Year of students when they were writing these essays (College Year 1, Year 2, Year 3, Year 4).

With the help of the sub-corpus generator, I extract all timed essays of all topics written by English majors of each Year. Table I provides a profile of the essays in each Year:

TABLE I. Descriptive statistics of the four sub-corpora generated from SWECCL 2.0

	NO.of ESSAYS	NO. of TOKENS				NO. of TYPES			
		Mean	Max.	Min.	SD	Mean	Max.	Min.	SD
Year 1	344	298	562	117	89	138	227	66	34
Year 2	1071	253	727	108	72	124	246	62	27
Year 3	272	298	597	105	92	145	257	60	36
Year 4	90	334	502	146	64	156	253	86	31

5.2 Complexity Measures

The present study uses the automated analysis tool, L2 Syntactic Complexity Analyzer developed by Lu [3], which provides 14 measures or indicators of syntactic complexity in five types: the length of production, sentence complexity, the number of subordination, the number of coordination, particular complex structures such as complex nominals. Studies have indicated high correlation coefficients (.834 to 1.000) between the statistic outcome generated by the automated analyzer and that by human annotators for all measures except T-units per sentence ($r = .74$)

I selected all those indices (Table II) in the analyzer. It should be noted that in these indices, some overlap with others, but all indices are chosen in my study because these indices reflect complexity at different levels. For example, mean length of sentence is a different level from the amount of subordination since the former may be due to use of coordination of sentences while the latter requires

the use of dependent clauses. By including all these indices, I hope to find out at what levels of complexity does their language develop.

TABLE II. The 14 indicators or measures of syntactic complexity from Lu (2010)

MEASURES	ABBREVIATION	MEANING
Mean length of clause	*MLC*	No. of words / No. of clauses
Mean length of sentence	*MLS*	No. of words / No. of sentences
Mean length of T-unit	*MLT*	No. of words / No. of T-units
Sentence complex ratio	*C/S*	No. of clauses / No. of sentences
T-unit complexity ratio	*C/T*	No. of clauses / No. of T-units
Complex T-unit ratio	*CT/T*	No. of complex T-units / No. of T-units
Dependent clause ratio	*DC/C*	No. of dependent clauses / No. of clauses
Dependent clauses per T-unit	*DC/T*	No. of dependent clauses / No. of T-units
Coordinate phrases per clause	CP/C	No. of coordinate phrases / No. of clauses
Coordinate phrases per T-unit	*CP/T*	No. of coordinate phrases / No. of T-units
Sentence coordination ratio	*T/S*	No. of T-units / No. of sentences
Complex nominals per clause	*CN/C*	No. of complex nominals / No. of clauses
Complex nominals per T-unit	*CN/T*	No. of complex nominals / No. of T-units
Verb phrases per T-unit	*VP/T*	No. of verb phrases / No. of T-units

5.3 Data Analysis

All the data were automatically retrieved with aid of L2 Syntactic Complexity Analyzer. Data were then entered into SPSS for further analysis. I used one-way ANOVA of the mean values of various indices. When statistically significant difference between four groups is registered, a post hoc test is conducted to find out specifically between which two groups such difference exists.

VI. RESULTS

Table III provides descriptive and ANOVA results, which shows that out of 14 measures that are employed in the present study, 13 showed significant differences between different Years. The only measure that did not show significant difference is sentence coordination ratio (T-units / sentences, T/S).

To my surprise, sentence complexity ratio is the only one that lower-Year level students (Year 1) outperformed significantly higher-level students (Year 2).

Table 4 summarizes the post hoc Tukey HSD tests results. It indicates that in terms of production length and ratio of complex nominals, significant differences were observed between the 3rd and 4th year students and those in 1st and 2nd year. MLS (mean length of sentence) and MLT(mean length of T-unit), for instance showed similar pattern in that Year 3 and 4 students outperformed significantly students in Year 2 and 1. MLC, CN/T, and CN/C showed almost linear progressive growth in which students in Year 3 and 4 outperformed in Year 2 which in turned outperformed Year 1.

It is interesting to note that Year 3 and Year 4 students did not show any significant difference in any measure. This indicates that Year 3 and Year 4 students remained at a stable level of syntactic complexity as measured by these 14 measures.

TABLE III. Descriptive and ANOVA Results

	Year 1 (n=344)		Year 2 (n=1071)		Year 3 (n=272)		Year 4 (n=90)		ANOVA		Partial η^2
	mean	*SD*	*mean*	*SD*	*mean*	*SD*	*mean*	*SD*	*F*	*Sig.*	
MLS	15.797	4.344	15.853	3.954	16.827	4.362	17.669	3.815	9.205	.000*	.015
MLT	14.038	3.093	14.304	2.872	15.204	3.882	16.144	3.494	16.832	.000*	.028
MLC	8.757	1.380	9.170	1.606	9.714	2.266	9.762	1.590	19.848	.000*	.032
C/S	1.815	.421	1.749	.412	1.747	.353	1.817	.326	3.028	.028*	.005
C/T	1.612	.273	1.575	.279	1.577	.274	1.658	.281	3.636	.012*	.006
DC/C	.341	.086	.330	.980	.336	.093	.363	.092	4.104	.007*	.007
DC/T	.568	.226	.542	.248	.552	.242	.624	.259	3.641	.012*	.006
T/S	*1.122*	*.142*	*1.107*	*.132*	*1.108*	*.111*	*1.098*	*.104*	*1.512*	*.209*	*.003*
CT/T	.437	.144	.416	.150	.418	.142	.461	.140	4.323	.005*	.007
CP/T	.320	.174	.337	.194	.369	.198	.399	.211	6.100	.000*	.010
CP/C	.202	.109	.218	.128	.237	.126	.241	.122	4.885	.002*	.008
CN/T	1.324	.401	1.489	.498	1.599	.514	1.728	.534	25.109	.000*	.041
CN/C	.827	.232	.953	.302	1.018	.293	1.043	.280	28.914	.000*	.047
VP/T	2.215	.404	2.152	.423	2.233	.415	2.389	.457	10.912	.000*	.018

p<0.05,

It is also interesting to note that at clausal level, such as clause per T-unit (C/T), dependent clause ratio (DC/C), dependent clause per T-unit (DC/T), and complex T-unit ratio (CT/T), there are only significant difference between Year 4 and Year 2. In these measures, Year 2 students had the lowest mean value of all Years.

In addition, despite the finding that students show significant improvement across a vast majority of 14 measures, the effect size for each of these measures is all below 0.1, which is considered to be quite a

small size. But the effect size in measures of production length and phrasal complexity was larger than sentential and clausal level.

VII. DISCUSSION

The study aims to find out whether there is significant difference among the four Years of students' timed argumentative writings and if any what the patterns of change are. The following is a summary of main findings:

Statistical analysis implies that Chinese college-level English majors showed development in almost every measure of syntactic complexity, though in minimal degrees. At phrasal level which is indicated by the measure of complex nominals and verb phrases, there was a linear model of growth with learners at a higher proficiency outperforming the adjacent lower-level learners.

In terms of clausal subordination, Year 2 students showed a dramatic reduction, and significant differences could only be seen between Year 4 and Year 2 students. Coordination of phrases were found to be significantly different between the higher and lower level of students, but sentence coordination showed no difference. The effect size was the smallest (almost negligible) in clausal complexification and phrase coordination. A surprising finding is the significantly lower level of sentence complexity ratio in Year 2 learners.

TABLE IV. A summary of post hoc tests

Measure	Post Hoc Tukey HSD Tests	Effect size (partial $\eta^{2)}$)
MLS (mean length of sentence)	Year 3, 4 > Year 1,2	0.015
MLT (mean length of T-unit)	Year 3,4 > Year 1, 2	0.028
MLC (Mean length of clause)	Year 3> Year 2> Year 1 Year 4>Year 2> Year 1	0.032
C/S	Year 1>Year 2	0.005
C/T	Year 4> Year 2	0.006
DC/C	Year 4> Year 2	0.007
DC/T	Year 4>Year 2	0.006
CT/T	Year 4>Year 2	0.007
CP/T	Year 4> Year 1, 2 Year 3> Year 1	0.001
CP/C	Year 3, 4 > Year 1	0.008
CN/T	Year 3> Year 2> Year 1 Year 4>Year 2> Year 1	0.041
CN/C	Year 3> Year 2> Year 1 Year 4>Year 2> Year 1	0.047
VP/T	Year 4> Year 1, 2, 3 Year 3> Year 2	0.018

* " >" indicates significantly higher than*

The growth in mean length of clause, T-unit, and sentences implies that as learners move to higher level of study, their writing becomes longer, containing more words in a sentence. The linear growth which is indicated by the numbers of complex nominals and coordinate phrases seems to suggest that longer and more complex writing of higher proficiency students is a result more of phrasal complexification and coordination than clausal subordination.

These finding echoes that of Lu who studied Chinese college level of EFL learners writing with an earlier version of the SWCCEL [35]. Of 14 measures, six measures were found to differ significantly across three Years. These are length of production units, amount of phrasal coordination, and complex nominals.

Different from Lu and Kim who found no significant difference in amount of clause-level subordination, in my study, subordination showed significant difference between Year 4 and Year 2 though such a difference is minimal. It should be noted that in Lu's study the fourth Yearrs' writings were excluded from due to smaller sample.

The findings of my study were more in close proximity to Kim's who used the same measures and the same analyzer but situated the study in the context of Korea and found that Korean college-level EFL learners' writing showed significant increase in 12 out of 14 syntactic measures across three levels of proficiency. These indices span all five categories of syntactic complexity. The indicators that showed the largest effect size were mean length of T-unit, complex nominals per T-unit and clause.

These findings of my study also echo some studies conducted in ESL context. Bulté and Housen's study found significant growth in mean length of noun phrase and mean length of sentence and T-unit, and in the amount of coordination but no significant difference in measures such as sub-clause ratio, complex sentence ratio and compound complex sentence ratio. This shows that ESL learners tend to complexify through coordinating and writing longer production units as they move along in language learning.

One difference, though, is that the effect size was medium with a *d*-value ranging from 0.3-0.5 while in my study the effect is quite small with all measures below 0.1. The small effect size is probably attributable to the cross-sectional nature of this research. The participants in my study are different student population in four college years. They are not instructed by the same group of teachers in the same program as is the case in the study of Bulté and Housen. This might have compromised the effect size.

According to Biber, as students move toward a higher level of proficiency, their writing complexity increases at phrasal level instead of clause level, and clausal complexity is a characteristic of informal conversations while phrasal complexity is of academic writing. The linear developmental trajectory in measures of complex nominals in my study seem to indicate such a trend, but more fine-grained measures are needed to confirm this and before that, I can only say that in the case of my study, students' writing showed a mixture of conversational and academic linguistic features since there was development at both phrasal and clausal level though the latter had a smaller magnitude of effect.

In the study by Bulté and Housen, Crossley, and McNamara, they found growth in coordination and phrasal elaboration but no subordination, which did not seem to correspond to the three-stage pattern of development of syntactic complexity proposed by Norris and Ortega. But results from my study seem to indicate that as students move to higher level of study, the amount of coordination and subordination did not "subside" but only slowed down and phrasal elaboration did increase but not necessarily "pervasive".

As Bulté and Housen suggested, the three-staged developmental pattern of complexity suggested by Norris and Ortega might be considered as recursive rather than a fixed developmental trajectory of L2 writing complexity. The development of L2 complexity may not rigidly follow coordination > subordination > clausal and phrasal elaboration but rather, at some points in time and under different task types, students show variability, especially when their L2 language system is still in development.

Another finding in my study is the significant decline in C/S in Year 2 where a significantly lower value was observed compared with in Year 1. This may be explained by a dynamic system perspective of L2 development. Just as there is an interaction and alternation of development between complexity, accuracy and fluency, the development of all syntactic features does not develop successively nor linearly. At some points, there might be temporary backsliding before progress picks up again.

Variation and fluctuation are essential characteristics of dynamic systems. Variability is high when the system is reconstructing while it will remain low if in a more stable system[36]. This may apply in the case of L2 syntactic complexity development. Until L2 learners reach an advanced level, their linguistic system is quite flexible, and syntactic features such as coordination, subordination and phrasal elaboration may progress in a non-successive, non-linear manner, with progress and regress alternating[37].

VIII. CONCLUSION AND IMPLICATION

The present study indicates that for students at different levels of language study, a writing teacher should adjust their criteria of evaluation, or shift their focus accordingly. Considering the public educational system which requires students to learn English as a school subject from primary 3 to high school graduation, students, upon entering university, have learned English for no less than 9 years. After four years of study, college English majors should be expected to produce texts that demonstrate complexity not just in terms of coordination and subordination, but also at phrasal level by phrase embedding, which is a preparation for their future academic study as international exchange students or graduate students home or abroad. This expectation requires teachers to play more facilitative roles in teaching and assessing writing and providing corrective feedback. We cannot simply expect learners to develop such as academic writing style automatically by self-guided reading or listening. The minimal magnitude of between-level effect in terms of all measures observed in the present study indicate that there is a lot more for teacher to do to raise students' awareness and capacity to write toward not just longer texts but also more academically featured essays.

Both linear and non-linear developmental trajectories found in the present study might suggest that teachers need to take a holistic view toward linguistic development in written production, especially in an EFL context such as China where preparation for examination or securing a certificate for future work is still a big part of the general purpose of L2 learning. On the one hand, students of higher language level are supposed to show more phrasal elaboration, but on the other they may still struggle at sentential level grammatical accuracy issues. Teachers are expected to strike a balance between these different needs rather than rigidly move toward cognitively more complex structures.

The assessment of writing both in classroom context and beyond should take into account these recent linguistic observations. At tertiary level, the assessment of writing should encourage a more academic style of writing but also differentiate writing evaluation criteria between different task types. An argumentative writing should not adopt the same scoring rubric as an expository writing. Development of fine-grained scoring rubrics based on understanding of these linguistic observations is more than desirable.

IX. LIMITATIONS AND FUTURE RESEARCH

Though the corpus used in the present study is a large-scale one, the distribution across four Years is not even, with essays by Year 2 students far outnumbering those by Year 4. The essays were written on 26 different topics and students were allowed to freely choose one out of 26 to write about. This has also caused an uneven distribution of essays by topics with those related to college education being favored by students and hence eliciting a fairly large number of essay in each Year while other topics related to laws or criminals unfamiliar to students eliciting only a few or even none in some Year. An ideal option would be to generate essays written on the same topic by four different Years of students, but that would lead to a situation where a very few or even no essay can be generated for a certain Year.

The present study only provides a general profiling of Chinese English majors' argumentative writing development in terms of syntactic complexity. Future study might examine certain particular language features such as the use of modifiers of nouns, the use of various types of clauses, and the usage of multiword units, which would provide more interesting information.

Further research can also focus on the predictive power of different syntactic measures on the learners' proficiency or human rating of writing quality, by adopting regression analysis, which might contribute to our understanding on what kind of role automated syntactic complexity analyzer can play in writing evaluation.

REFERENCES

[1] Ortega L (2012) Interlanguage complexity: A construct in search of theoretical renewal. In:Kortmann B and Szmrecsanyi B (eds) Linguistic complexity: Second language acquisition, indigenization, contact. Berlin: De Gruyter, pp. 127-55.

[2] Bulté , B & Housen, A. (2012). Defining and operationalizing L2 complexity, In A. Housen, F. Kuiken, & I. Vedder (Eds.) Dimensions of L2 performance and proficiency: Investigating complexity, accuracy, and fluency in SLA, Amsterdam: John Benjamins.

[3] Lu, Xiaofei. (2010). Automatic analysis of syntactic complexity in second language writing. International Journal of Corpus Linguistics, 15(4), 474-496.

[4] Lu, Xiaofei (2012). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. TESOL Quarterly, 45(1):36-62.

[5] Bestgen, Y., Granger, S. (2014). Quantifying the development of phraseological competence in L2 English writing: An automated approach, Journal of Second Language Writing, 26, 28-41.

[6] Paquot, M. (2019). The phraseological dimension in interlanguage complexity research. Second Language Research, 35(1), 121–145

[7] Ortega, L. (2015). Syntactic complexity in L2 writing: Progress and expansion. Journal of Second Language Writing, 29, 82-94.

[8] 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.

[9] Wolfe-Quintero, K., Inagaki, S., Kim, H. Y., Kim, H. Y., & Inagaki, S. (1998). Second language development in writing: Measures of fluency, accuracy, & complexity. Honolulu: University of Hawai'i Press.

[10] Ai, Haiyang & Lu, Xiaofei. (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. In Ana Díaz-Negrillo, Nicolas Ballier, and Paul Thompson (eds.), Automatic Treatment and Analysis of Learner Corpus Data, (pp. 249-264). Amsterdam/Philadelphia: John Benjamins.

[11] Martínez, A. C. L. (2018). Analysis of syntactic complexity in secondary education EFL writers at different proficiency levels. Assessing Writing, 35, 1-11.

[12] Kyle, K., & Crossley, S. A. (2018). Measuring Syntactic Complexity in L2 Writing Using Fine‑Grained Clausal and Phrasal Indices. The Modern Language Journal, 102(2), 333-349.

[13] Biber, D., Gray, B.,& Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? TESOL Quarterly, 45, 5–35.

[14] Gray, B. (2015). On the Complexity of Academic Writing. Disciplinary Variation and Structural Complexity. Corpus-Based Research in Applied Linguistics: Studies in Honor of Doug Biber, 49-77.

[15] Biber, D., Gray, B., & Staples, S. (2014). Predicting patterns of grammatical complexity across language exam task types and proficiency levels. Applied Linguistics, 37, 639–668.

[16] Biber, D., & Gray, B. (2016). Grammatical complexity in academic English: Linguistic change in writing. Cambridge University Press.

[17] Norris, J., Ortega, L., (2009). Towards an Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity, Applied Linguistics, 30, 4, 555–578

[18] Bulté, B., Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity, Journal of Second Language Writing, 26, 42-65

[19] Crossley, A., D., McNamara, S., D. (2014). Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. Journal of Second Language Writing, 26, 66-79

[20] Kuiken, F., Vedder, I. (2019). Syntactic complexity across proficiency and languages: L2 and L1 writing in Dutch, Italian and Spanish. International Journal of Applied Linguistics,1, 1-19, https://doi.org/10.1111/ijal.12256

[21] Kim, Ji-Young. (2014). Predicting L2 writing proficiency using linguistic complexity measures: A corpus-based study. English Teaching, 69(4), 27-51.

[22] Lu, Xiaofei & Ai, Haiyang. (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. Journal of Second Language Writing, 29, 16-27.

[23] Housen, A., & Kuiken, F. (2009). Complexity, accuracy and fluency in second language acquisition. Applied Linguistics, 30(4), 461–473.

[24] van Geert, P. (2009). A comprehensive dynamic systems theory of language development. In K.de Bot & R. W. Schrauf (Eds.) Language development over the lifespan. New York, NY: Routledge.

[25] Lahmann, C. Steinkrauss. R.,Schmid, S. M. (2016). Factors affecting grammatical and lexical complexity of long-term L2 speakers' oral proficiency. Language Learning, 66 (2), 354-385

[26] Adams, R., Aloesnit, N., Mohd Alwi, N., Newton, J. (2015). Task complexity effects on the complexity and accuracy of writing via text chat, Journal of Second Language Writing 29, 64–81

[27] Johnson, 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.

[28] Ong, J., Zhang, J., L. (2010). Effects of task complexity on the fluency and lexical complexity in EFL students' argumentative writing. Journal of Second Language Writing,19, 218-233.

[29] Polio, C., & Yoon, H. J. (2018). The reliability and validity of automated tools for examining variation in syntactic complexity across genres. International Journal of Applied Linguistics, 28(1), 165-188.

[30] Yoon, H. J., & Polio, C. (2017). The linguistic development of students of English as a second language in two written genres. TESOL Quarterly, 51(2), 275-301.

[31] Verspoor, M., Lowie, W., van Dijk, M. (2008). Variability in second language development from a dynamic systems perspective. The Modern Language Journal, 92, 214-231.

[32] van Geert, P. and M. van Dijk. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. Infant Behavior & Development, 25, 340–74.

[33] Verspoor, M., Schmid, M, Xu, X., (2012). A dynamic usage-based perspective on L2 writing. Journal of Second Language Writing, 21, 239-263.

[34] Wen, Liang, Yan, (2008). Spoken and Written English Corpus of Chinese Learner, Beijing: Foreign Language Teaching and Research Press.

[35] Wen, Liang, Yan, (2005). Spoken and Written English Corpus of Chinese Learner, Beijing: Foreign Language Teaching and Research Press.

[36] Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. Applied Linguistics 27(4), 590–619

[37] Spoelman, M., & Verspoor, M. (2010). Dynamic patterns in development of accuracy and complexity: A longitudinal case study in the acquisition of Finnish. Applied Linguistics, 31(4), 532-553.