A Longitudinal Learner Corpus Study of Syntactic Complexity in L2 Writing

Nur Najah Radhiah Zainal Abidin (University of Malaya, Malaysia), Meng Huat Chau (University of Malaya, Malaysia) and Adam Ramadhan Saiful Bahri (Bank Simpanan Nasional, Malaysia)

The study of syntactic complexity has attracted considerable attention in second language (L2) writing research in recent years, revealing the range of grammatical resources employed by learners in their text construction and the role these resources play in L2 writing development (e.g., Ortega, 2003; Ai & Lu, 2013; Wood & Struc, 2013; Kim, 2014; Scontras et al., 2014; Lu & Ai, 2015; MacDonald, Montag & Gennari, 2015; Mazgutova & Kormos, 2015; Vyatkina, Hirschmann & Golcher, 2015; Biber & Gray, 2016). Most studies conducted in this area have been cross-sectional in design (cf. Wood & Struc, 2013; Vyatkina et al., 2015) and often involve comparisons with a target language norm. Several early studies, which focused on a number of specific syntactic measures, have provided some initial findings which offer useful longitudinal insights (see, e.g., Larsen-Freeman, 1983; Arnaud, 1992; Kern & Schultz,1992; Casanave, 1994; Ishikawa, 1995). More longitudinal research, however, is required to make sense of how changes in the employment of a wide range of grammatical resources might take place over time, which is essential to a fuller understanding of what development may mean in the study of L2 writing.

This paper reports on a longitudinal corpus study that considered 14 syntactic complexity measures in L2 texts to determine whether and to what extent changes may be observed in the writing of a group of 124 secondary school students at four points in time over a 24-month period. All the texts produced by the students were based on a common narrative task. The results of the 14 measures are generated using L2 Syntactic Complexity Analyzer (Lu, 2010), which examined such constructs as mean length of sentence (MLS), mean length of T-unit (MLT), mean length of clause (MLC), clauses per sentence (C/S), verb phrases per T-unit (VP/T), clauses per T-unit (C/T), dependent clauses per clause (DC/C), dependent clauses per T-unit (DC/T), T-units per sentence (T/S), complex T-unit ratio (CT/T), coordinate phrases per T-unit (CP/T), coordinate phrases per clause (CP/C), complex nominals per T- unit (CN/T), and complex nominals per clause (CN/C). The scores from the L2 Syntactic Complexity Analyzer were analyzed using repeated measures ANOVA to ascertain if there are statistical differences of the 14 indices as observed across four points in time in the writing of the 124 students.

The findings of the study indicate a statistically significant change in the use of syntactic structures based on 10 out of the 14 measures over time (i.e., MLS, MLT, DC/C, DC/T, CP/C, CN/C, C/S, VP/T, CN/T and C/T). More dependent clauses per clauses and per T-unit, complex nominal per clauses,

mean length of T-unit, clauses per sentences and complex nominal per T-unit were observed to be used over time in the student texts. The use of coordinate phrases per clauses, mean length of sentences, verb phrases per T-unit and clauses per T-unit, on the other hand, have significantly decreased over time. Interestingly, while a high employment of dependent clauses has been noted to be a feature of spoken language (e.g., Halliday, 1989; Biber et al, 2011), this category of clauses has been found to be used more frequently over time in this study (see also Pallotti & Ferrari (2008) for similar observations). A heavy use of subordinations also suggests high causative relations in writings (Ryshina-Pankova, 2015), indicating an increased emphasis on such relations in the narrative texts. Equally interesting, as time goes by, the use of complex nominals was found to have significantly increased, which opens up questions about the use of complex nominals in genres other than academic writing (cf. Biber & Gray, 2016). Furthermore, the findings show that while changes as measured in certain constructs can be observed within six months, changes in other syntactic features may only be observable over a full period of 24 months, indicating an organic, developmental process at play (e.g., Chau, 2015; Larsen-Freeman, 2015).

The paper concludes by offering some cautionary notes. While the observed increase in complexity may imply that there is evidence of more sophisticated language use over time in the writing of the students, it should be remembered that "more complex does not necessarily mean better" (Ortega, 2003). Development in L2 writing essentially entails a complex interplay between one's deployment of linguistic resources and what is socially valued as 'good' writing (see, e.g., Skehan & Foster, 2007; Ortega, 2015). As Pallotti (2009, p. 14) observes, the ability for learners to put language into use involves not only syntactic complexification; it "also entails the development of discourse and sociolinguistic repertoires that the language user can adapt appropriately to particular communication demands". Further implications for future research into syntactic complexity in L2 writing are considered.

References

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). Amsterdam/Philadelphia: John Benjamins.

Arnaud, P. J. (1992). Objective lexical and grammatical characteristics of L2 written compositions and the validity of separate-component tests. In *Vocabulary and applied linguistics* (pp. 133-145). Palgrave Macmillan UK.

Biber, D. (1986). Spoken and written textual dimensions in English: Resolving the contradictory findings. *Language, 62*(2), 384-414.

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.

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

Casanave, C. P. (1994). Language development in students' journals. *Journal of Second Language Writing*, *3*(3), 179-201.

Chau, M. H. (2015). *From language learners to dynamic meaning makers: A longitudinal investigation of Malaysian secondary school students' development of English from text and corpus perspectives*. Unpublished doctoral thesis, University of Birmingham, UK.

Ishikawa, S. (1995). Objective measurement of low-proficiency EFL narrative writing. *Journal of Second Language Writing*, *4*(1), 51-69.

Kern, R. G., & Schultz, J. (1992). The effects of composition instruction on intermediate level French students' writing performance: Some preliminary findings. *The Modern Language Journal*, *76*(1), 1-13.

Kim, J. Y. (2014). Predicting L2 writing proficiency using linguistic complexity measures: A corpus-based study. *English Teaching*, *69*(4), 27-51.

Larsen-Freeman, D. (1983). Assessing global second language proficiency. In H.W. Seliger, & M.H. Long (eds.), *Classroom oriented research* (pp. 287–304). Rowley, MA: Newbury House.

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

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.

MacDonald, M. C., Montag, J. L., & Gennari, S. P. (2015). Are There Really Syntactic Complexity Effects in Sentence Production? A Reply to Scontras et al. (2015). *Cognitive science, 40*, 513-518.

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.

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.

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

Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*, *30*(4), 590-601.

Pallotti, G., & Ferrari, S. (2008). La variabilità situazionale dell'interlingua: implicazioni per la ricerca acquisizionale e il testing linguistico. In G. Bernini, L. Spreafico, & A. Valentini (eds): *Competenze lessicali e discorsive nell'acquisizione di lingue seconde*. Perugia: Guerra.

Ryshina-Pankova, M. (2015). A meaning-based approach to the study of complexity in L2 writing: The case of grammatical metaphor. *Journal of Second Language Writing*, *29*, 51-63.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, *10*(3), 209-232.

Skehan, P., & Foster, P. (2007). Complexity, accuracy, fluency and lexis in task-based performance: A meta-analysis of the Ealing Research. In S. Van Daele, A. Housen, F. Kuiken, M. Pierrard, & I. Vedder (eds.), *Complexity, accuracy, and fluency in second language use, learning, and teaching* (pp. 207–226). Brussels: University of Brussels Press.

Scontras, G., Badecker, W., Shank, L., Lim, E., & Fedorenko, E. (2015). Syntactic Complexity Effects in Sentence Production. *Cognitive science, 39(3)*, 559-583.

Vyatkina, N., Hirschmann, H., & Golcher, F. (2015). Syntactic modification at early stages of L2 German writing development: A longitudinal learner corpus study. *Journal of Second Language Writing*, *29*, 28-50.

Wood, N., & Struc, N. (2013). A corpus-based, longitudinal study of syntactic complexity, fluency, sentence variety, and sentence development in L2 genre writing. *Reitaku University Jouurnal*, *96*, 1-44.