**Expanding the scope of complexity research in SLA: a phraseological perspective**

Magali Paquot

Université catholique de Louvain

Usage-based research in corpus linguistics, psycholinguistics and cognitive linguistics has provided convergent evidence that lexis and grammar are inextricably intertwined and that word combinations, be they framed in terms of phraseological units, formulaic sequences or constructions, play crucial roles in language acquisition, processing, fluency, idiomaticity and change (e.g. Ellis, 1996; Sinclair, 1991; Wray, 2002; Schmitt, 2004; Goldberg, 2006). Second language research has been relatively slow to follow suit but phraseology, formulaic language and constructions are now at the forefront of debates in foreign language learning and teaching (Meunier & Granger, 2008; Polio, 2012). Learner corpus studies have already provided unique insights into the links between word combinations and L2 proficiency and development (e.g. Paquot & Granger, 2012; Ellis et al., 2015; Ebeling & Hasselgård, 2015). Research, however, is still extremely fragmented and not all domains of L2 research have navigated the transition. Interlanguage complexity research, most particularly, has remained impervious to current theoretical developments related to how words combine together to form meaningful units: Complexity has traditionally been narrowed down to syntactic complexity with a strong focus on clause-related measures (e.g. the number of coordinate or dependent clauses per T-unit), and lexical complexity is still very much regarded as its poor relation (Ortega, 2012). This is particularly unfortunate since complexity is considered one of the “major research variables in applied linguistic research” (Housen & Kuiken, 2009): measures of linguistic complexity are widely used to describe L2 performance, assess L2 proficiency, and trace L2 development.

Today, interlanguage complexity research stands at a crossroads. Ortega (2012) described complexity as “a construct in search of theoretical renewal”. Measures of complexity have been repeatedly criticized for their lack of theoretical foundation and construct validity (i.e. how well they measure the construct that they are intended to measure) (e.g. Norris & Ortega, 2009; Biber et al. 2011; Pallotti, 2015). Leading researchers in the field have also called for an expanded view of complexity as a multifaceted and multidimensional construct that cannot be fully explored via just one of its dimensions (as is commonly done) but requires to be operationalized with a battery of measures (including new and more specific measures) tapping different properties of the construct in multivariate research designs (e.g. Ortega, 2012; Bulté & Housen, 2012). In Paquot (2019), I have argued that a successful renewal of the domain will also require a better appreciation of the phraseological dimension of language use.

In this presentation, I will report the first results of a 5-year FNRS research project (2016-2021) that aims to define and circumscribe the linguistic construct of phraseological complexity, i.e. “the range of phraseological units that surface in language production and the degree of sophistication of such phraseological units” (Paquot, 2019), and to theoretically and empirically demonstrate its relevance for L2 complexity research, and more generally for theories of L2 use and development. The project centres around four main objectives: (1) determine the dimensions of phraseological complexity, (2) establish the construct validity of phraseological complexity measures automatically calculated using natural language processing (NLP) techniques and corpus data, (3) chart the development of phraseological complexity in L2 writing and speech, and (4) identify the best set of complexity measures to adequately capture the dynamics of phraseological complexity development over time.

To achieve these objectives, I have started investigating the diversity and sophistication of word combinations in a variety of cross-sectional and longitudinal written and spoken EFL learner corpora (e.g. the Varieties of English for Specific Purposes dAtabase (VESPA) learner corpus, the Longitudinal Database of Learner English (LONGDALE), and the Trinity Lancaster Spoken Learner Corpus). In the presentation, I will briefly summarize some of the results and focus more particularly on the conceptual/theoretical and methodological issues faced.

I will round off my talk with a discussion of what I believe are the most important implications and most promising applications of this research programme.

**References**

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.

Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & I. Vedder, I. (Eds), *Dimensions of L2 performance and proficiency: complexity, accuracy and fluency* (pp. 21–46). Amsterdam: John Benjamins.

Ebeling, S. & Hasselgård, H. (2015). Learner corpora and phraseology, In Sylviane Granger; Gaëtanelle Gilquin & Fanny Meunier (ed.), *The Cambridge Handbook of Learner Corpus Research* (pp. 207–230). Cambridge: Cambridge University Press.

Ellis, N. C. (1996). Sequencing in SLA: Phonological Memory, Chunking and Points of Order. *Studies in Second Language Acquisition, 18*, 91–126.

Ellis, N., R. Simpson-Vlach, U. Römer, M. O’Donnell & S. Wulff (2015). Learner corpora and formulaic language in second language acquisition. In S. Granger, G. Gilquin & F. Meunier (Eds), *The Cambridge Handbook of Learner Corpus Research* (pp. 357–378). Cambridge: Cambridge University Press.

Goldberg, A. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Housen, A., & Kuiken, F. (2009). Complexity, Accuracy, and Fluency in Second Language Acquisition. *Applied Linguistics, 30*(4), 461–473.

Meunier, F. & Granger, S. (Eds.), (2008). *Phraseology in Foreign Language Learning and Teaching*. Amsterdam: John Benjamins.

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.

Ortega, L. (2012). Interlanguage complexity: A construct in search of theoretical renewal. In B. Kortmann & B. Szmrecsanyi (Eds.), *Linguistic Complexity: Second Language Acquisition, Indigenization, Contact* (pp. 127–155). Berlin, Boston: De Gruyter.

Pallotti, G. (2015). A simple view of linguistic complexity*. Second Language Research, 31*(1), 117–134.

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

Paquot, M. & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics* *32*, 130-149.

Polio, C. (Ed.), (2012). Special issue on formulaic language. *Annual Review of Applied Linguistics, 32*.

Schmitt, N. (Ed.), (2004). *Formulaic Sequences: Acquisition, Processing and Use*. Amsterdam: John Benjamins.

Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Wray A. (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.