

Exploring linguistic complexity in two varieties of learner English: ESP and applied languages

Linguistic complexity (Bulte and Housen 2012) is one of the main dimensions used for the characterisation of an L2. It is a multidimensional construct that can be exploited in different tasks such as L2 proficiency level identification (Kyle 2016) or proficiency progress (McAllister and Belan 2014), L1 identification (Xiaofei and Haiyang 2015) or error detection. The task of L2 variety identification can also benefit from the use of the construct. By operationalising the construct in terms of specific metrics it is possible to provide points of comparisons that can inform on possible differences or similarities of two types of L2.

Many tools exist to compute various indices of complexity, and they have been widely used in recent years in the research on criterial features (Hawkins and Buttery 2010). However, little research has been conducted in the domain of L2 varieties linked to linguistic complexity metrics. It is assumed that *applied languages* learners and LSP learners rely on a common functional language culture due to the professionalisation aspect of their field. It is therefore important to provide empirical results regarding the vicinity of these two varieties. Our research question consists in analysing the difference between the two varieties in terms of linguistic complexity.

Our proposal is to compare two corpora from the two L2 varieties, i.e. a corpus of 50 texts written by LEA learners of English and a corpus of 282 texts written by LSP learners in several domains (Gaillat and Ballier 2019). The B2 proficiency level of all the texts is used as a control variable. After computing various syntactic complexity and readability metrics with a pipeline program composed of CoreNLP (Manning et al. 2014), L2SCA (Lu 2010) and Quandeta (Benoit et al. 2018), pairwise comparisons are computed and significance tests are conducted. The results inform the researcher on the complexity indices that do or do not show significant differences. They contribute to the debate on the difference between the two types of L2.

References

- Benoit, Kenneth, Kohei Watanabe, Haiyan Wang, Paul Nulty, Adam Obeng, Stefan Müller, and Akitaka Matsuo. 2018. "Quanteda: An R Package for the Quantitative Analysis of Textual Data." *Journal of Open Source Software* 3 (30): 774.
<https://doi.org/10.21105/joss.00774>.
- Bulte, Bram, and Alex Housen. 2012. *Defining and Operationalising L2 Complexity*. John Benjamins Publishing Company.
- Gaillat, Thomas, and Nicolas Ballier. 2019. "Expérimentation de Feedback Visuel Des Productions Écrites d'apprenants Francophones de l'anglais Sous MOODLE." In *Actes de La Conférence EIAH2019*. Paris, France: Association des Technologies de l'Information pour l'Éducation et la Formation.
- Hawkins, John A., and Paula Buttery. 2010. "Criterial Features in Learner Corpora: Theory and

Thomas Gaillat Université Rennes 2
Sophie Belan Université de Nantes

- Illustrations." *English Profile Journal* 1 (01).
<https://doi.org/10.1017/S2041536210000103>.
- Kyle, Kristopher. 2016. "Measuring Syntactic Development in L2 Writing: Fine Grained Indices of Syntactic Complexity and Usage-Based Indices of Syntactic Sophistication." Georgia State University. https://scholarworks.gsu.edu/alesl_diss/35.
- Lu, Xiaofei. 2010. "Automatic Analysis of Syntactic Complexity in Second Language Writing." *International Journal of Corpus Linguistics* 15 (4): 474–496.
- Manning, Christopher D., Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. 2014. "The Stanford CoreNLP Natural Language Processing Toolkit." In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 55–60. <http://acl2014.org/acl2014/>.
- McAllister, Julie, and Sophie Belan. 2014. "L'anglais de spécialité en LEA à la croisée des domaines : étude de l'acquisition du lexique spécialisé." *ASp. la revue du GERAS*, no. 66 (November): 41–59. <https://doi.org/10.4000/asp.4564>.
- Xiaofei, Lu, and Ai Haiyang. 2015. "Syntactic Complexity in College-Level English Writing: Differences among Writers with Diverse L1 Backgrounds." *Journal of Second Language Writing*, New developments in the study of L2 writing complexity, 29 (September): 16–27. <https://doi.org/10.1016/j.jslw.2015.06.003>.