ICAME Journal • 48(1) • 2024 • 113–116
DOI: 10.2478/icame-2024-0007

sciendo

ICAME Journal

# Agnieszka Leńko-Szymańska and Sandra Götz (eds.). *Complexity, accuracy and fluency in learner corpus research*. Amsterdam: John Benjamins, 2022. vi. 327 pp. ISBN 9789027212580 (HB)

**Review**

Raffaella Bottini

*Lancaster University, United Kingdom*

Learner corpus research (LCR) on complexity-accuracy-fluency (CAF) can inform theories of second language acquisition (SLA) and has practical applications for language teaching and testing. This theoretical and practical orientation is demonstrated in the volume *Complexity, accuracy and fluency in learner corpus research* edited by Agnieszka Leńko-Szymańska and Sandra Götz. This book is a collection of eleven papers providing an overview of the latest trends in research on CAF using learner corpora. In addition to the individual papers, the introduction of the volume offers a clear contextualisation of CAF in LCR, including a historical overview, a categorisation of existing studies, examples, key references, and areas for future development. For comparison, similar collections include Granger, Gilquin and Meunier (2013), Castello, Ackerley and Coccetta (2015), Brezina and Flowerdew (2017), and Götz and Mukherjee (2019). This review reflects on the contribution of Leńko-Szymańska and Götz (2022) to LCR.

A quantitative summary of the studies in the volume is presented in Table 1, following Brezina and Bottini's (2019) example. This quantitative breakdown is also a proxy for identifying methodological trends in LCR on the CAF triad. Most of these studies focus on syntactic complexity. Among independent variables, speaker-related variables are the most researched, especially L2 proficiency, with a prevalence of research on the intermediate (B1–B2) and advanced (C1–C2) levels. In terms of target language, most of the papers analyse L2 English, in line with a more general trend in LCR (Paquot and Plonsky 2017). A wide range of L1 groups is analysed, favouring Spanish, German, and Norwegian learners. The corpora used represent either written or spoken L2 production. They are relatively small, ranging from 40k to 190k tokens, consisting of 12 to 5,331 texts, with two studies using longitudinal data. The overview of the methods shows a balanced number of qualitative and quantitative approaches, including relatively sophisticated inferential statistics and data visualization.

Among the four studies on complexity in this volume, one has a conceptual and methodological focus, and three papers analyse developmental trends across time and/or proficiency levels. These four chapters provide a notable contribution to the literature, discussing both the construct and operationalization of complexity (e.g. presenting new indices), and addressing areas which have not been explored fully in the field (e.g. the longitudinal development of complexity). **Gaillat**'s chapter addresses the call for a clear definition of the construct of linguistic complexity to ensure its operationalization is reliable and valid (Bulté and Housen 2012; Norris and Ortega 2009; Pallotti 2009). His paper investigates 84 metrics of lexical diversity, syntactic complexity, and readability. The metrics are first categorised depending on the variables in the mathematical formulas, their relationships, and the type of mathematical operation (e.g.

* *Corresponding author: Raffaella Bottini, E-mail: r.bottini@lancaster.ac.uk*

Open Access. © 2024 Raffaella Bottini, published by Sciendo.
This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 License.

**Table 1.** Quantitative overview of Leńko-Szymańska and Götz (2022).

| Area | Number of studies |
|---|---|
| Research focus within the CAF framework | Complexity (6); accuracy (5); fluency (2) |
| Linguistic variables | Complexity: lexical (1), syntactic (5), phraseological (1)<br>Accuracy: grammatical (2), syntactic (1), phraseological (1)<br>Fluency: filled and unfilled pauses (1), discourse markers (1), repeats (2) |
| Independent variables | Speaker-related e.g. proficiency, L1, age, gender, educational background (7); context-related e.g. time (2); task-related e.g. topic (4); language-related e.g. collocation type, association measures (2) |
| Target language | English (8); Russian (1); Italian (1) |
| L1 | Bulgarian (1); Chinese (1); Czech (1); French (1); German (2); Italian (1); Japanese (1); Norwegian (2); Russian (1); Spanish (3); Taiwanese (1); not reported (2) |
| L2 proficiency (CEFR) | Elementary A1–A2 (2); intermediate B1–B2 (7); advanced C1–C2 (7) |
| Learners' age | Teenagers (1); university students (7); various (1) |
| Corpus size | 40,695–193,111 tokens (not reported 6); 12–5,331 texts |
| Target genre | Written essays (6); spoken dialogues (4); spoken monologues (2) |
| Research design and statistical tests | Qualitative analysis (6); quantitative analysis: descriptive (5), inferential (8); statistical tests: cluster analysis e.g. Principal Component Analysis (2), differences between group means e.g. t-test and ANOVA (3), correlation (2), regression model (4); data visualization (10) |
| Implications | Theoretical, SLA (9); methodological (6); practical: language testing (3), language teaching (3) |

*Note*. CEFR = *Common European Framework of Reference* (Council of Europe, 2001).

mean, rate, ratio). Then, the construct validity of the metrics is examined applying hierarchical clustering to evaluate their homogeneity. Finally, a supervised machine learning approach is used to measure how accurately the metrics can predict L2 English writers' proficiency. Gaillat develops a taxonomy of indices which could help interpret their linguistic meaning to provide fine-grained automatic feedback in language assessment.

Two chapters focus on the longitudinal development of syntactic complexity in L2 writing. **Kisselev**, **Klimov** and **Kopotev** compare two groups of learners of Russian as a heritage language (HL) and as a foreign language (FL). The authors also evaluate the ability of syntactic complexity measures (including a new measure of phrasal depth) to differentiate L2 proficiency levels. The results show that syntactic complexity increases across proficiency levels in L2 Russian writing. HL learners produce longer sentences and a larger proportion of subordination than FL learners. **Dirdal** analyses how subordination develops in a longitudinal corpus containing writings produced by five Norwegian learners of L2 English over a four-year period. Cross-linguistic influence is explored comparing syntactic features in L2 English with 1) L1 Norwegian texts produced by the same learners, and 2) comparable L1 English texts. This study uses a fine-grained distinction of clause types, providing clear details of their categorisation and coding, which also allows the investigation of clause diversity in addition to frequency, thus addressing the call for more detailed measures of subordination (Lambert and Kormos 2014). **Paquot**, **Gablasova**, **Brezina** and **Naets** focus on the development of complexity across proficiency levels. They analyse phraseological complexity for the first time in L2 English speech using the Trinity Lancaster Corpus. The researchers fit a mixed effects regression model to investigate the effect of proficiency, task, and L1 on the phraseological diversity and sophistication of verb+object collocations. In addition to discussing the effect of topic and priming from the interlocutor, the chapter includes some inspiring methodological reflections on 1) measuring phraseological complexity in written vs. spoken data, 2) the suitability of measures based on central tendency (mean, median), 3) the analysis of combinations of words with negative MI scores, and 4) the use of metrics of phraseological complexity based on collocation types rather than tokens.

Accuracy is the focus of three chapters in the volume, including one methodological paper. These three studies complement each other and, together, offer a clear overview of some of the key aspects

involved in research on accuracy, e.g. coding errors, analysing how accuracy develops across proficiency levels, and suggesting practical implications for language teaching. **Hoffmann** discusses some of the main issues involved in error annotation. After comparing the error annotation schemes used in three learner English corpora (i.e. International Corpus of Learner English, Cambridge Learner Corpus, and Teaching Resource Extraction from an Annotated Corpus of Learner English Corpus), the author presents a series of examples of common practical issues in error annotation. Hoffmann suggests adopting a hierarchical structure of error tag sets which "put the annotator in control over the level of granularity that can confidently be considered appropriate for any given error" (p. 164). A similar taxonomy can address some of the typical issues related to subjectivity, consistency, and comparability of error annotation, while also allowing multi-layer annotation to include both descriptive and explanatory tags. **Graf** and **Huang** use the Czech and Taiwanese components of the Louvain International Database of Spoken English Interlanguage (LINDSEI) to compare the frequency of different types of errors (morphological, lexical, grammatical) across L2 speakers of English at the B2 and C1 level. Through the identification of frequent errors (e.g. articles and verb tenses), this study informs language teaching practices. **Spina** contributes to research on phraseological accuracy from a perspective that has not been explored in LCR before. Her study analyses how different dimensions of collocational relationship (e.g. repetition, strength and rarity, exclusivity, directionality) can predict phraseological accuracy in a subset of the Longitudinal Corpus of Chinese Learners of Italian (LoCCLI). Spina also analyses the effect of proficiency, topic, and word combination type (verb+object, adjective+noun, noun+adjective) on phraseological accuracy. Her findings provide evidence that L2 phraseological knowledge does not follow a linear pattern of development and varies as a function of time, combination type, and exclusivity of collocational relationships which is measured by logDice.

Two studies combine research on complexity and accuracy, thus setting an example for possible future developments in LCR on CAF. In **Thewissen** and **Anishchanka**'s contribution, a subset of the International Corpus of Learner English was manually annotated for grammatical accuracy, while syntactic complexity was computed automatically. The authors present inspiring reflections on the relationship between complexity and quality ("is more syntactically complex better?" p. 214), on teaching L2 syntactic complexity, and on the importance of reporting statistically non-significant results in LCR since these "remain a meaningful source of information on the process of language development as they point to the subtle changes that the L2 is going through" (p. 229). Their study also provides a comprehensive coding scheme of grammatical accuracy with examples based on the Louvain Error Manual. **Lyashevskaya**, **Vinogradova** and **Scherbakova** analyse the relationship between syntactic complexity and syntactic accuracy in L2 English texts written by Russian university students. They find that some indices of syntactic complexity can predict syntactic accuracy. They discuss how task type can influence syntactic complexity and, in turn, its accuracy.

Finally, two studies offer a valuable contribution to research on fluency. **Götz**, **Wolk** and **Jäschke** analyse four strategies that can enhance fluency, also called fluencemes (i.e. filled and unfilled pauses, discourse markers, repeats). They use a subset of the LINDSEI corpus to test the use of fluencemes by advanced learners of English across different tasks, also investigating the effect of L1 transfer, L2 educational background, learner's gender and age, and interlocutor's speech behaviour. Their findings are particularly revealing regarding the extent to which the interviewer's speech behaviour can influence learners' fluency. Based on this, they discuss methodological implications for future research and practical implications for language testing. Their chapter also provides detailed information on the coding of fluencemes and the disambiguation of problematic cases. **Aas** and **Rørvik** investigate interlanguage fluency variation comparing the use of repeats in L1 Norwegian and L2 English speech produced by the same learners. Their chapter presents a comprehensive overview of how repeats have been defined and identified in studies on interlanguage speech, as well as their functional features. They analyse the frequency, position, and type of repeats. Their findings highlight individual variation and a heterogeneous picture. The researchers reflect on the importance of including activities in language teaching to raise learners' awareness of fluency enhancement strategies and the contexts where they typically occur.

Overall, this volume offers a good overview of learner corpus research on the CAF triad. It includes theoretical implications that can inform theories of SLA, practical applications for language teaching and testing, and insightful reflections on methodological practice in the field. The volume also represents an opportunity to identify possible areas that have not received enough attention in LCR and could be researched further in the future. These include low proficiency levels (A1 and A2), the comparison of written and spoken texts produced by the same learners, across a wide range of genres, and multimedia data.

## References

Brezina, Vaclav and Raffaella Bottini. 2019. Review of Erik Castello, Katherine Ackerley and Francesca Coccetta (eds.). *Studies in learner corpus linguistics. Research and applications for foreign language teaching and assessment* (2015). *International Journal of Learner Corpus Research* 5 (1): 113–117.

Brezina, Vaclav and Lynne Flowerdew (eds.). 2017. *Learner corpus research: New perspectives and applications*. London: Bloomsbury.

Bulté, Bram and Alex Housen. 2012. Defining and operationalising L2 complexity. In A. Housen, F. Kuiken and I. Vedder (eds.). *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*, 21–46. Amsterdam: John Benjamins.

Castello, Erik, Katherine Ackerley and Francesca Coccetta (eds.). 2015. *Studies in learner corpus linguistics. Research and applications for foreign language teaching and assessment*. Berlin: Peter Lang.

Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* Cambridge: Cambridge University Press.

Götz, Sandra and Joybrato Mukherjee (eds.). 2019. *Learner corpora and language teaching*. Amsterdam: John Benjamins.

Granger, Sylviane, Gaëtanelle Gilquin and Fanny Meunier (eds.). 2013. *Twenty years of learner corpus research: Looking back, moving ahead. Proceedings of the First Learner Corpus Research Conference (LCR 2011).* Louvain: Presses Universitaires de Louvain.

Lambert, Craig and Judit Kormos. 2014. Complexity, accuracy, and fluency in task-based L2 research: Toward more developmentally based measures of second language acquisition. *Applied Linguistics* 35 (5): 607–614.

Norris, John and Lourdes Ortega. 2008. Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics* 30 (4): 555–578.

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

Paquot, Magali and Luke Plonsky. 2017. Quantitative research methods and study quality in learner corpus research. *International Journal of Learner Corpus Research* 3 (1): 61–94.