RiCL Research in Corpus Linguistics

Review of Leńko-Szymańska, Agnieszka and Sandra Götz eds. 2022. *Complexity, Accuracy and Fluency in Learner Corpus Research.* Amsterdam: John Benjamins. ISBN: 978-9-027-21258-0. DOI: https://doi.org/10.1075/scl.104

Paweł Szudarski
University of Nottingham / United Kingdom

It is uncontroversial to say Learner Corpus Research (henceforth, LCR) has been on the rise in recent years, as shown by the increasing number of publications on the topic.[1] In this sense, *Complexity, Accuracy and Fluency in Learner Corpus Research* by Leńko-Szymańska and Götz (2022) is a welcome addition to the literature, contributing to a growing body of work and showcasing LCR studies conducted from the perspective of complexity, accuracy and fluency (henceforth, CAF). Focusing on the CAF triad both theoretically and methodologically, the book consists of 12 chapters that report state-of-the-art findings and novel methodologies that tap into lexis, grammar, phraseology and other aspects and dimensions of second language (L2) use as represented, operationalised and analysed by means of learner corpora.

The opening chapter by Leńko-Szymańska and Götz sets the scene for the book as a whole, usefully introducing the volume and its goals. Not only do the authors outline the structure of the book but also present CAF as a research strand of growing importance in the field, with examples of current topics such as the identification of the most suitable measures of CAF constructs, the use of increasingly sophisticated and refined methods and statistical procedures in CAF research and, finally, the application of the CAF triad as a starting point for corpus analysis. As regards the latter, the authors explain how CAF constructs and measures lend themselves well to the principles of LCR, particularly in the

[1] See Granger *et al.* 2015 for a comprehensive account of the field of LCR. For further examples, see also the *International Journal of Learner Corpus Research*.

form of the contrastive interlanguage analysis (Granger 2015), the focus on L2 learners' interlanguage and its juxtaposition with first language (L1) usage. That said, Leńko-Szymańska and Götz point to many questions that still remain unanswered in this strand of work, including, for instance, the developmental path and criterial features of L2 learners' production (in both writing and speech) at various levels of proficiency or inconclusive findings in terms of which measures best capture the dimensions of CAF. The authors also hint at the multiplicity of likely interactions between the three CAF constructs, while most extant research has only studied them in isolation, clearly showing the potential of LCR studies to investigate different aspects of L2 use in relation to each other.

Gaillat's chapter focuses on three selected aspects of complexity (lexical diversity, readability and syntactic complexity), explored in the context of the relevance of corpus-based measures for assessing L2 learners and distinguishing L2 performance at different proficiency levels. Specifically, recognising the challenge of working with corpus-based metrics of complexity, the chapter proposes a model of evaluation of such measures as a way of facilitating meaningful interpretations of L2 learner data. The model is built around the notion of linguistic scopes understood as links between a given metric's mathematical formula and its surface (textual) manifestation at the level of word, phrase, clause, sentence or even text. Using such a scopes-based approach as the textual delineation of CAF, Gaillat's analysis investigates 84 complexity measures and reveals some degree of homogeneity (in-cluster consistency). Findings suggest that in terms of the usability of complexity metrics, the diversity, repetition and size of the word and text scopes are particularly effective at discriminating between L2 production at different proficiency levels. On a practical level, these results mean that the scope approach can aid the design of fine-grained feedback messages aimed at L2 learners, responding to their current proficiency level and specific problematic areas.

The next chapter is by Kisselev, Klimov and Kopotev, who examine syntactic complexity measures as indicators of proficiency level in learner language. Using the *Russian Error-Annotated English Learner Corpus* (RULEC), which includes longitudinal, classroom-based, written data from learners of Russian at intermediate and advanced levels[2] and a list of 12 syntactic complexity indices, the authors test the feasibility of such measures as markers or indicators of L2 proficiency in Russian. Results show that differences in the numeric values for these indices point to learners' overall syntactic improvement as they

[2] http://www.web-corpora.net/RLC/rulec

grew in proficiency. Further, such complexity indices are also able to reveal differences between learners of Russian as a foreign vs. heritage language, such as, for instance, that the relative proportion of coordinate clauses is lower for the latter. That said, findings also point to the non-linearity and multi-dimensionality of L2 writing development. Overall, then, not only does this study confirm that corpus-based measures of syntactic complexity can be effectively used to track linguistic development in the L2, but it also demonstrates that such complexity measures can be employed in the analysis of languages other than English.

Also focusing on syntactic complexity, Dirdal's chapter reports on a study into the development of L2 writing complexity as dependent on clause types, L1 influence and individual differences. The study follows five L1-Norwegian learners of English over four school years, tracking their development and use of subordinate clauses at both the clausal and phrasal level. Results point to different developmental trajectories for individual clause types (e.g., clauses with a nominal function are the most frequent ones at the beginning of this period, while adnominal clauses are less frequent), with learners improving in syntactic diversity even when there is little evidence of change in syntactic subordination. Interestingly from the perspective of L2 learning theory, across the five learners included in the study, there is more variation and fluctuation in the frequencies of syntactic features in the earlier vs. later school years, potentially explaining why more advanced levels of proficiency are characterised by less individual variation. Further, L1 effects are also demonstrated, as the lack of specific clause types in the L1 Norwegian (e.g., *-ing* clauses) led only to few occurrences of this feature in the learner English data, suggesting difficulty and late development in the L2. Finally, given that the data analysis involves comparisons of individual learners, the study also touches upon the discussion of individual-level variation and the key role of individual differences in the process of L2 learning.

The chapter by Paquot, Gablasova, Brezina and Naets represents the growing body of corpus-based work into the use and learning of L2 phraseology, moving beyond the analysis of written language and usefully focusing on oral performance. Specifically, the authors examine phraseological complexity in L2 English learners' spoken production across different proficiency levels (B1 to C2 of the *Common European Framework of Reference for Languages*; CEFR)[3] as demonstrated by texts from the *Trinity Lancaster*

[3] https://www.coe.int/en/web/common-european-framework-reference-languages

Corpus (transcribed interactions between examiners of Graded Examinations in Spoken English and L2 candidates).[4] Approaching complexity through the lens of phraseological diversity (root type-token ratios) and sophistication (median mutual information scores, MI), the analysis deals with the use of verb-noun collocations (e.g., *dance tango*), a feature of L2 learning that has received a great deal of attention in corpus-based research,[5] but so far has not been studied much in relation to spoken learner data. Results suggest that, while overall phraseological diversity in L2 oral performance increases with proficiency, statistical significance is only found between B2 and C1 levels, that is, between learners who are at intermediate and advanced levels, respectively. At the same time, however, such findings need not necessarily be taken to mean that increased proficiency results also in similar upward trends in the construct of phraseological sophistication. As it turns out, MI scores decrease significantly from B1 to B2. Further, a follow-up qualitative analysis of the learner data seems to show that learners at the B2 level and above use more specific verbs and less idiomatic collocations, while lower-level students stick to a limited number of highly associated combinations. This suggests that relying on quantitative findings only might run the risk of hiding some important aspects of a qualitative change in L2 learners' development of phraseological complexity. In sum, by focusing on L2 speech, the study is an important step in extending LCR findings to the oral domains of L2 use and, therefore, responds to frequent calls within the corpus community to pay more attention to spoken corpora. Methodologically, the study also shows that MI scores, particularly used in measures of central tendency such as medians, may not be the most appropriate indicator or marker of phraseological development in L2 speech.

The focus of Graf and Huang's chapter is on persistent errors in the spoken language of L2 learners of English at different proficiency levels. Situated in the broader discussion of grammatical accuracy, the study seeks to provide empirical evidence for the ways in which L2 development surfaces at the B2 and C1 levels of the CEFR. In the analysis, data are sampled from the error-tagged Czech and Taiwanese components of the *Louvain International Database of Spoken English Interlanguage* corpus (LINDSEI),[6] with learners' global proficiency and five specific competencies (namely, range, accuracy, fluency, phonological control and coherence) assessed by two professional raters. In terms of learner errors, they are classified and counted with the help of the *Louvain Error Tagging*

[4] https://cass.lancs.ac.uk/trinity-lancaster-corpus/
[5] See Szudarski (2023) for an overview of corpus-based analysis of L2 collocations.
[6] https://uclouvain.be/en/research-institutes/ilc/cecl/lindsei.html

Manual.[7] The analysis reveals a clear difference between the two levels under study, with the vast majority of errors committed by B2 speakers (84.4%) compared to C1 speakers (15.6%). Graf and Huang take this finding as evidence of a threefold increase in grammatical accuracy. Juxtaposed with previous LCR focused on written data (see Le Bruyn and Paquot 2021 for recent examples), it is perhaps unsurprising to see that accuracy in L2 speech and writing develop at different rates. From the perspective of L2 learning and teaching, however, the study usefully points to errors in the use of articles and grammatical tenses as particularly problematic and persistent in learner language. Even though such errors decrease in their overall frequency at the higher proficiency level (C1), they are still present in the learner output, which the authors argue singles them out as potential criterial features for distinguishing learners at different levels of grammatical accuracy.

Similarly to Gaillat, Hoffmann's chapter is methodological in nature and revolves around challenges with the measurement and description of lexical accuracy by means of learner corpora. Specifically, Hoffmann discusses error annotation schemes applied in CAF research and focuses on their effectiveness and accuracy in terms of identifying relevant features of learner language (e.g., types of errors identified or potential overlaps between specific categories). With lexical accuracy in written L2 English as the focal point, the author examines the taxonomies and tag sets of errors employed in three major LCR projects: 1) the *International Corpus of Learner English*,[8] 2) the *Cambridge Learner Corpus*[9] and 3) the *Teaching Resource Extraction from an Annotated Corpus of Learner English Project*,[10] using them as the basis for his own analysis of data from the *Marburg Corpus of Intermediate Learner English* (MILE; Kreyer 2015). By referring to specific examples of overlaps in error categories between these taxonomies, Hoffmann convincingly argues for the presence of hierarchical structure in the organisation of error tags, the application of clear annotation guidelines, and more transparency and open science practices in research reports (e.g., annotation guidelines being available not only to annotators but also any interested researcher). The chapter concludes with a discussion of how these recommendations should increase the usability of error tags, as well as greatly benefit the comparability of findings across various LCR studies.

---

[7] https://repository.uantwerpen.be/docman/irua/102b7d/granger_et_al__error_tagging_manual_v2_0_2022.pdf
[8] https://corpora.uclouvain.be/cecl/icle/home
[9] https://www.sketchengine.eu/cambridge-learner-corpus/
[10] http://www.treacle.es/

Concerned with the area of L2 phraseological development, Spina's chapter offers a novel and comprehensive account of the effects of time and various dimensions of collocability on phraseological accuracy. Specifically, with the help of longitudinal data from beginner and pre-intermediate L1-Chinese learners of Italian, the study is a multi-layered analysis of the accuracy of two types of collocations (noun + adjective/adjective + noun combinations and verb + noun combinations) in L2 Italian writing as dependent on time (that is, learner essays written at the beginning of a six-month language programme vs. essays written at the end) and specific dimension of collocational relationship (namely, collocation frequency, association measure, exclusivity of collocational relationship and directionality of collocational relationship). Results indicate that L2 collocational accuracy varies differently over time and across the three types of combinations, with noun-adjective combinations decreasing in accuracy after six months of studying Italian. This is unlike the adjective-noun collocations, for which the number of errors drop. From the perspective of SLA, it is also worth adding that as the learners in the study represent different proficiency levels (beginner vs. intermediate), Spina is also able to show the effects of L2 proficiency on phraseological accuracy. But the effect of time does not vary significantly across proficiency levels, suggesting a non-linear developmental path for L2 collocations. As regards the effects of different dimensions of collocational relationship, only the exclusivity of combinations (i.e., how strong the association is between collocating words) positively affects the accuracy of learners' production, showing that frequency cannot be regarded as the sole defining feature of phraseological units. The study is also commendable from the methodological standpoint, combining a longitudinal, corpus-based design with the use of multifactorial mixed-effects statistics.

Continuing the line of research into developmental changes in learner language, Thewissen and Anishchanka examine the interaction between grammatical accuracy and syntactic complexity at different proficiency levels. Focusing on intermediate and advanced students of L2 English (third- and fourth-year university students from the L1-French, Spanish and German components of the *International Corpus of Learner English*), the authors submit these data to the automatic L2 *Syntactic Complexity Analyzer* (Lu 2010), with a view to discovering evidence of 'interactional dynamics' between the two constructs. Their analysis reveals some interesting patterns of findings, such as for instance a competitive relationship between grammatical accuracy and syntactic complexity at B1 and B2 CEFR levels, with learners' grammatical accuracy displaying marked improvement as they

grow in L2 proficiency. Further, while comparisons between B2 and C1 levels show only subtle developmental shifts, the juxtaposition of C1 and C2 levels offers more supportive evidence for 'interactional dynamics' between the two constructs under study. This is evidence of improvement in both syntactic complexification and grammatical accuracy, although the latter fails to reach statistical significance. On a methodological level, the study convincingly shows how this type of corpus-based research into processes such as L2 development, while necessarily needing to rely on statistical comparisons, may also benefit from engaging in greater detail with seemingly random non-significant results.

In the following chapter, Lyashevskaya, Vinogradova and Scherbakova zoom in on the relationship between syntactic complexity and accuracy as revealed by their analysis of the impact of task types on written data produced by L1-Russian learners of L2 English. Drawn from the RULEC corpus (over 5,000 examination papers written in response to two different task types of description and opinion essay), this learner data is used to operationalise the two constructs under study: 1) syntactic complexity (20 indices) and 2) syntactic accuracy (frequency of syntactic errors). And indeed, statistical analyses, perhaps unsurprisingly, point to a significant link between the two, also showing clear task effects as another factor that mediates learners' performance. While three syntactic parameters are significantly related to learners' accuracy in the description task, six different parameters are found for the opinion task; the only two syntactic complexity metrics that significantly predict accuracy in both tasks are the number of sentences and adverbial clauses. Overall, then, the study demonstrates how corpus-derived indices of syntactic complexity can assist with the assessment of L2 written production, helping to quantify and categorise the most common types of syntactic errors committed by L2 learners. By considering the task effects, the results of the study are also relevant pedagogically, showing how automated tracing of syntactic features can inform the delivery of more bespoke error correction and L2 instruction more broadly.

Encouragingly, the final two chapters demonstrate the usefulness of learner corpora for investigating L2 fluency, a construct that has so far received less research attention than the other two elements of the CAF framework. Respectively, Götz, Wolk and Jäschke examine the development of fluency as dependent on such key SLA variables as L1 transfer, the length of instruction or the role of interlocutors' communicative behaviour, while Aas and Rorvik address individual variation in learners' L2 fluency by comparing their speaking styles in both the L1 and L2. Focusing on four indicators of fluency (filled pauses, unfilled

pauses, discourse markers and repeats) in data from the LINDSEI corpus, the findings of Götz, Wolk and Jäschke reveal clear L1 effects, as well as a positive impact of study abroad and years of instruction on learner spoken English. Their analysis also points to the importance of confluence, that is, the convergence of all interlocutors in the completion of specific tasks, including the role of the interviewer in shaping L2 learners' output (both its fluency and amount). In turn, Aas and Rorvik's study focuses on the frequency, types and position of repeats (reiterations of certain groups of sounds) in interview data taken from the Norwegian component of the LINDSEI corpus (both L1 and L2 data). Their results suggest that while repeats occur more frequently in the learner data, such repetitions or 'disfluencies' appear in the L1 data as well, serving different discourse functions, contributing to the structure of conversations and constituting an important feature of one's idiolect. Thus, rather than regarding repeats as an undesired feature of L2 speech, there might need to be more pedagogical focus on such fluency enhancement strategies, raising L2 learners' awareness of these features.

By way of closing, it is without a doubt that LCR has been on the increase via different research avenues, and this includes CAF studies as attested by the chapters in *Complexity, Accuracy and Fluency in Learner Corpus Research*. For anybody interested in corpus-based analysis, and particularly the affordances of learner corpora, Leńko-Szymańska and Götz's volume will be a wealth of insights, both theoretically and methodologically. The volume is also likely to inspire future corpus-based studies in the area of CAF and SLA more broadly. Such research is encouraged, particularly in the light of the recognised distance (and limited dialogue so far) between LCR and SLA (Granger 2021; Myles 2021). As Granger (2021) aptly points out, the mutual benefits of a reproachment between the two fields are substantial, and there is a great deal of potential to be realised in future corpus-based work, particularly in relation to spoken language and the construct of fluency. As already signalled, the dominance of studies focused on written language is notable in the volume.

That comment notwithstanding, Leńko-Szymańska and Götz's volume is testament to how CAF research has capitalised on learner data and corpus-based methods of analysis. As promised by the editors in their introduction, the book covers a wide range of topics and research designs, benefiting from interdisciplinary approaches and conceptual novelty. Another asset I would personally highlight is the methodological innovation and diversity demonstrated in the reported research, both of which transpire from the individual

chapters and provide ample examples of designs and perspectives that can be employed in corpus-based research concerned with CAF. As such, I view this volume as a timely and valuable contribution to the field, likely to become a useful reference work for individuals working in the area of LCR and beyond.

References

Granger, Sylvianne, Gaëtanelle Gilquin and Fanny Meunier eds. 2015. *The Cambridge Handbook of Learner Corpus Research.* Cambridge: Cambridge University Press.

Granger, Sylvianne. 2021. Commentary: Have learner corpus research and second language acquisition finally met? In Bert Le Bruyn and Magali Paquot eds. *Learner Corpus Research Meets Second Language Acquisition*. Cambridge: Cambridge University Press, 243–257.

Kreyer, Rolf. 2015. The *Marburg Corpus of Intermediate Learner English* (MILE). In Marcus Callies and Sandra Götz ed. *Learner Corpora in Language Testing and Assessment*. Amsterdam: John Benjamins.

Le Bruyn, Bert and Magali Paquot M. 2021. *Learner Corpus Research Meets Second Language Acquisition.* Cambridge: Cambridge University Press.

Lu, Xiaofei. 2010. Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics* 15/4: 36–62.

Myles, Florence. 2021. Commentary: An SLA perspective on learner corpus research. In Bert Le Bruyn and Magali Paquot eds. *Learner Corpus Research Meets Second Language Acquisition*. Cambridge: Cambridge University Press, 258–273.

Szudarski, Paweł. 2023. *Collocations, Corpora and Language Learning*. Cambridge: Cambridge University Press.

Reviewed by

Paweł Szudarski
University of Nottingham
School of English
Room A84 Trent Building
University Park
NG7 2RD
Nottingham
United Kingdom
E-mail: pawel.szudarski@nottingham.ac.uk