John Benjamins Publishing Company

This is a contribution from IJLCR 9:2
© 2023. John Benjamins Publishing Company

This electronic file may not be altered in any way. The author(s) of this material is/are permitted to use this PDF file to generate printed copies to be used by way of offprints for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible only to members (students and faculty) of the author's institute. It is not permitted to post this PDF on the internet, or to share it on sites such as Mendeley, ResearchGate, Academia.edu.

Please see our rights policy at **https://benjamins.com/content/customers/rights**
For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: **www.copyright.com**).

For further information, please contact **rights@benjamins.nl** or consult our website: **www.benjamins.com**

William J. Crawford (Ed.), *Multiple Perspectives on Learner Interaction. The Corpus of Collaborative Oral Tasks.* Berlin/Boston: de Gruyter, 2021. viii + 278 pp. ISBN 9781501517372 . [Studies on Language Acquisition 60].

Reviewed by Sylvie De Cock (UCLouvain)

The main aim of the edited volume entitled "Multiple Perspectives on Learner Interaction. The Corpus of Collaborative Oral Tasks" is twofold. Not only does it introduce the *Corpus of Collaborative Oral Tasks* (CCOT) to the learner corpus and *Second Language Acquisition and Teaching* (SLAT) research communities, highlighing the unique make-up and characteristics of the corpus, but, as the title aplty suggests, it also brings together seven papers that showcase "a variety of ways that the CCOT can inform our understanding of peer interaction and task variation through the use of learner corpora" (p. 4). As underlined by the editor, William Crawford, in his introduction, the volume illustrates the potential of a shared corpus like the CCOT (which can be investigated by a range of scholars from different research traditions) to "promote the application of corpus linguistic methods to describing, interpreting, and theorizing about second language performance" (p. 1), thereby fostering more interactions between the fields of *Learner Corpus Research* (LCR) and SLAT (p. 4).

The opening chapter, "Introduction to the Corpus of Collaborative Tasks", written by Crawford and McDonough, is devoted to the design of the CCOT with a focus on the instructional context, the tasks, the learner variables and data formatting (e.g. text format, speaker identification, header information for each file). The corpus clearly fills a gap in the spoken learner corpus family where computerized collections of learner peer interactions are few and far between. The CCOT comprises 775 timed paired oral assessments collected within the framework of the *Program in Intensive English* (PIE) at Northern Arizona University between 2009 and 2012. The corpus totals 268,324 words produced by 576 PIE students (between 18 and 26 years old). As explained by the authors, the corpus is not balanced for gender and L1 background as there is a strong bias towards male learners (78%) and Arabic and (Mandarin) Chinese speakers (over 95%), reflecting the PIE student population at the time of collection. While no information is available about prior amount and type of exposure to English or other known L2s, information about the proficiency level of each of the students is available as the learners were assigned one of three general (intermediate) proficiency bands determined

https://doi.org/10.1075/ijlcr.00037.dec
International Journal of Learner Corpus Research 9:2 (2023), pp. 282–287. ISSN 2215-1478 | E-ISSN 2215-1486
© 2023 John Benjamins Publishing Company

by an in-house test whose results have been mapped onto TOEFL iBT scores (ranging from 32 to 69 on the TOEFL iBT). In addition, each paired test was rated by PIE instructors using an analytic rubric consisting of the following three categories: collaboration, task completion and style. The 24 tasks completed by the students are clearly given centre stage as very detailed information about these tasks is provided both in the chapter and in a 73-page Appendix at the end of the volume (e.g. communicative purpose of task, including persuasion, decision making or selecting from alternatives, level of students, number of words, mean words per text, planning time and type, task completion time, task directions, rubic(s) used to score the task). Surprisingly, the chapter and the Appendix do not include information about the actual transcription scheme used, which is listed among Paquot et al.'s (2023) core types of obligatory spoken learner corpus design metadata. Another regrettable omission concerns practical information about how the CCOT can actually be accessed, as it is labeled as "freely available" (p. 15).

In the second chapter, Alquraishi and Crawford seek to characterize L2 collaborative oral task language in the whole CCOT using Multi-Dimensional Analysis (MDA). Five dimensions emerge from an analysis of a selection of 43 linguistic features: expression of personal thoughts and judgements, interactive vs individual presentation, reaching agreement vs informational discourse, narration of human experiences vs evaluations of concepts and objects, and expression of preferences. The positive and negative sides of the dimensions are interpreted and illustrated thanks to paired tests from the corpus, which incidentally also makes it possible for the authors to highlight learners' use of formulaic language when performing the tasks. According to Alquraishi and Crawford, the fact that the five dimensions identified explain only 32% of the variance in the corpus can be explained by the limited number of features included and possibly also by the formative assessment context, which may well lead to fairly homogeneous linguistic performance across tasks and proficiency levels.

Nekrasova-Beker's paper examines CCOT learners' use of a particular phraseological unit, namely phrase-frames or p-frames, which are discontinuous sequences with a variable slot (e.g. *i * with you, we can * the*). The main focus is on the variability and on the functional characteristics of well-dispersed and topic-independent 4-word p-frames across three intermediate proficiency sublevels in the CCOT (from low-intermediate to high-intermediate). The analysis shows, among others, that the p-frames used by the high-intermediate level participants exhibit the most variability and that the majority of the p-frames are used by the learners across all three proficiency groups to express stance, which reflects the highly interactive nature of the tests in the corpus. Nekrasova-Beker's study illustrates some of the challenges of conducting research on subsamples of unequal size, which made it difficult for her to fully explore the effect of learners'

© 2023. John Benjamins Publishing Company
All rights reserved

L1 (Arabic vs Chinese) on the students' use of p-frames across the three sublevels. Given the rather limited number of contributions by Arabic students at sublevel 3 (which was used as a baseline for the number of contributions included in the other sets of data analysed in the L1 investigation), the author regards the whole sample examined as not large enough to be representative.

Chapter 4, by McDonough and Uludag, addresses one of the recurrent challenges of assessing peer-to-peer interactions: should paired oral tests be evaluated individually or jointly? The audio-recorded shared and individual performances of a subset of students taking the paired test in the CCOT were scored using a (particularly well-described) 4-category rating scheme (discourse management, collaborative communication, content development, language accuracy and complexity) specifically developed for the study. The analysis reveals few differences between the shared and individual ratings. An examination of a limited set of lexico-grammatical features generally associated with collaboration (*wh*-questions, 1st and 2nd person pronouns, subordinate conjunctions and discourse markers) points to a positive relationship between shared ratings and the joint production of the lexico-grammatical features investigated. The study also uncovers a tendency for the student who speaks first to take on greater responsibility to elicit information and use more of the lexico-grammatical features investigated, which would need to be confirmed by data collected using within-groups design.

In Chapter 5, Becker seeks to determine the extent to which two sets of information-exchange paired oral tasks from the CCOT that were rated as highly collaborative actually exhibit similar profiles of co-occurring linguistic features. To this end, Becker conducts a bottom-up cluster analysis of 29 features that were shown by Crawford et al. (2019) to be associated with high-collaboration interactions. Six main profiles are identified (three per set of data), which suggests that collaborative peer interaction is characterized by considerable variation: there is not just one combination of features that is associated with highly collaborative paired interaction. The chapter ends with a discussion of some of the possible implications of the findings for the fine-tuning of rating scales and for rater training in connection with paired speaking tasks, such as using interaction samples exhibiting different profiles.

The final three chapters report on studies that go beyond using the CCOT as "mute spoken data" or simply transcripts of spoken discourse (Ballier & Martin, 2015: 107) in that they involve going back to the audio files of the paired tasks under investigation to code and examine a series of pronunciation and fluency features using Praat and ELAN for example. In the chapter entitled "Exploring the Impact of Situational Characteristics on the Linguistic Features of Spoken Oral Assessment Tasks", Staples takes a register-functional approach to the CCOT. Her

© 2023. John Benjamins Publishing Company
All rights reserved

focus in on two CCOT assessment tasks that vary in terms of situational characteristics (one primarily informational task and one primarily argumentative task) as she sets out to investigate the possible impact of these situational characteristics on a series of lexico-grammatical (e.g. stance adverbials), interactional (e.g. discourse markers), and pronunciation and fluency features (e.g. silent pauses). The analysis points to a major impact of task type on the lexico-grammatical features used by the speakers, with features of involvement (e.g 1st and 2nd person pronouns, determiners and conditional clauses) and stance features (e.g. necessity modals) used to a greater extent in the argumentative task, and informational features (e.g. nouns and attributive adjectives) more prevalent in the informational task. The argumentative task also seems to be less cognitively demanding as the participants' contributions exhibit higher speech rate and mean length of run and fewer and shorter silent pauses than the contributions in the informationally oriented task. As rightly pointed out by Staples, the findings have clear implications for the creation and evaluation of oral tests in that test developers need to carefully consider (the situational characteristics of) the tasks that are included in tests to ensure that these tasks are directly relevant to the test takers in their specific assessment context.

Choi, McAndrews and Kang investigate the effects of participants' gender and of speaking tasks with different communicative purposes (making a decision, persuasion, selecting from multiple alternatives) on utterance fluency (operationalized as speaking rate and articulation rate) as well as the exent to which the relationship between gender and fluency is mediated by lexical and syntactic alignment. The study uncovers that task factors and learner variables do have an impact on spoken fluency: while speaking rate is shown to be higher in tasks that involve selecting from multiple alternatives, speech that is produced by same-gender interlocutors tends to be more syntactically (rather than lexically) aligned and it exhibits higher articulation rate.

In the final chapter, Ghanem focuses on a selection of informative and opinion-based tasks from the CCOT to investigate the effect of task type on learners' use of pronunication features including six fluency features (speech rate, average run duration, number and length of silent pauses, number of filled pauses, and speaking time difference) and one prosodic feature (mean pitch range). The analysis also explores the extent to which the pronunication features under study predict the overall and collaboration scores PIE instructors gave to the paired tests. The findings indicate, among others things, that, whereas average run duration (i.e. "average uninterrupted speaking time between two pauses", p. 183) is positively correlated with the various scores across the two task types, speaking time difference is significantly correlated with the scores for the opinion-based tasks only.

© 2023. John Benjamins Publishing Company
All rights reserved

Welcoming a new spoken learner corpus to the learner corpus family is always a cause for celebration given their relative scarcity compared with written learner corpora. Peer interaction occupies a significant place in SLAT research and it lies at the heart of the CCOT. Crawford therefore hopes that the new corpus (as well as research that centres on it) will help bring the LCR and SLAT communities closer together and will help "find more spaces where LCR and SLAT can 'meet'" (p. 4). The CCOT appears to have been designed to answer some of the pleas from SLAT researchers echoed in Myles (2021): more spoken data and a wider range of tasks given the impact of different production modes and tasks on results. The carefully documented CCOT elicitation tasks included in the Appendix arguably also go some way towards answering Myles's call for LCR and SLAT to share not only corpora but also elicitation tasks. Although most of the studies included in the volume are essentially exploratory in nature (and the words *explore* and *exploratory* are used recurrently throughout the chapters), they help foreground various aspects of the unique composition of the corpus. They also illustrate a whole series of rich CCOT-centered research agendas that make use of advanced quantitative procedures and statistical analyses (e.g. MDA, multiple regression, bootstrapping, cluster analysis). More generally, the volume is strong on methodology with detailed information about data selection, coding, scoring, interrater reliability, statistical analyses and systematic discussions of limitations. The challenges of using a corpus like the CCOT are not swept under the carpet (e.g. issues relating to the size of subsets of data) and Nekrasova-Beker calls for the compilation of a comparable corpus of peer interaction data produced by native speakers performing the same tasks. As shown by previous research (e.g., Gráf, 2019), examining both learner and native speaker oral productions collected from participants completing the same tasks can prove invaluable in identifying tasks that are cognitively very demanding irrespective of learner or native speaker status.

In view of its focus on task-elicited learner peer interaction and on its diverse (quantitative) approaches to the CCOT, "Multiple Perspectives on Learner Interaction. The Corpus of Collaborative Oral Tasks" will be of interest to a large group of scholars and language practitioners including LCR and SLAT researchers, language examiners and test developers, EFL/ESL teachers, and materials developers.

References

Ballier, N., & Martin, P. (2015). Speech annotation of learner corpora. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge Handbook of Learner Corpus Research* (pp. 107–134). Cambridge University Press.

© 2023. John Benjamins Publishing Company
All rights reserved

Crawford, W., McDonough, K., & Brun-Mercer, N. (2019). Identifying linguistic markers of collaboration in second language peer interaction. A lexico-grammatical approach. *TESOL Quarterly*, *53*(1), 180–207.

Gráf, T. (2019). Speech rate revisited: The effect of task design on speech rate. In S. Götz & J. Mukherjee (Eds.), *Learner corpora and language teaching* (pp. 175–190). John Benjamins.

Myles, F. (2021). Commentary: An SLA perspective on learner corpus research. In B. Le Bruyn & M. Paquot (Eds.), *Learner corpus research meets second language acquisition* (pp. 258–273). Cambridge University Press.

Paquot, M., König, A., Stemle, E., & Frey, J-C. (2023). *Core metadata schema for learner corpora*. Université catholique de Louvain.

Address for correspondence

Sylvie De Cock
Université catholique de Louvain
ILC
Place Cardinal Mercier 31/L3.03.33
1348 Louvain-la-Neuve
Belgium
sylvie.decock@uclouvain.be

Publication history

Date received: 23 May 2022
Date accepted: 3 September 2022

© 2023. John Benjamins Publishing Company
All rights reserved