

Ten years of the *International Journal of Learner Corpus Research*

Sandra Götz and Magali Paquot

Philipps University Marburg and FNRS - UCLouvain

The tenth volume of the *International Journal of Learner Corpus Research* (IJLCR) marks a decade of significant contributions to the field of learner corpus research, showcasing the evolution, innovation, and impact of our collective work. Over the past ten years, IJLCR has provided a window into current learner corpus research, thereby increasing the field's visibility and attracting greater attention to the diverse research conducted by the learner corpus community. With a growing number of studies exploring how learner corpora can advance our understanding of Second Language Acquisition (SLA) and linguistic theory, the journal has become a key platform for significant research into learners' grammar, lexis, phraseology and discourse. This research aims to describe L2 performance, assess L2 proficiency, and trace L2 development. Additionally, IJLCR not only features research that informs foreign and second language teaching, but also presents studies that report on tasks such as automatic grammatical annotation and proficiency classification. The journal thus effectively demonstrates the multidisciplinary nature of learner corpus research, highlighting the applicability of learner corpora across various domains and purposes.

The variety of studies published in IJLCR represent the current level of research in the field. The target languages covered are diverse, although English remains by far the most frequently investigated language, accounting for 46 out of the 67 empirical studies that have been published in the journal over the last decade (69%). While other languages are less

represented, their range has expanded to include L2 German, Spanish, French, Finnish, Korean and Dutch – a welcome development that broadens the scope and relevance of learner corpus research. Despite repeated calls for more research on L2 speech, the majority of the studies (63%) have focused on L2 writing, a trend already noted in previous reviews in the field of applied linguistics (e.g. Paquot & Plonsky, 2017).

While several studies use well-established learner corpora such as the *International Corpus of Learner English* (ICLE) (Granger et al., 2020), the *ICNALE* (Ishikawa, 2023), or the *LINDSEI* (Gilquin et al., 2010), others have used new resources such as the *Trinity Lancaster Corpus* (Gablasova et al., 2019) as well as more localized and diverse learner corpora. The increasing variety of learner corpora being compiled is also well documented in the corpus reports that IJLCR began publishing three years ago. Importantly, all the learner corpora featured in these reports are made publicly available to the research community. In terms of methods, one notable development since Paquot & Plonsky's (2017) review of the field is the significant increase in cross-sectional studies that compare learner language across proficiency levels. By contrast, the percentage of longitudinal studies has remained stable (just above 10%). Another important development is the growing number of studies that move beyond analyzing corpus frequencies to take into consideration the individual learner texts, which highlights the field's increasing interest in accounting for the heterogeneity in learner corpus data. This trend is also reflected in the decreased use of measures such as the X^2 test (dropping from 23% to 12%).

IJLCR is not only a window into current learner corpus research but also offers insights into ongoing developments and the direction in which the field is evolving. Emerging during a time of unparalleled advancements in research practices, the journal has consistently committed itself to staying at the forefront of scholarly discourse by embracing innovative practices. We have actively promoted principles such as open access, transparency, and

reusability. We encourage authors to consider uploading their materials for data collection and analysis to open access repositories such as the IRIS database (<https://www.iris-database.org/>). IRIS is an online repository for data collection materials used for second language research. This includes data elicitation instruments such as interview and observation schedules, language tests and stimuli, pictures, questionnaires, software scripts, word lists, teaching intervention activities, amongst many other types of materials used to elicit data (<http://www.iris-database.org> for more information). In practice, IJLCR authors have also started using other alternatives such as OSF to share their datasets and codes. The journal has also become part of the community of journals that award Open Science Badges. If authors make their data and/or materials available, their articles will be awarded recognition from the Centre for Open Science via an 'Open Science' badge displayed next to their article (<https://benjamins.com/content/authors/opensciencebadges> for more information). To promote transparency and documentation of research methodology, IJLCR provides a platform for scholars to disseminate and receive recognition for their methodological advancements. Since 2020, we have started to publish new article types and welcome submissions detailing corpus development, compilation, and annotation, as well as reports on research materials, software tools, and shared tasks (Paquot & Callies, 2020). These contributions are essential for advancing the field by documenting research methodologies and promoting reproducibility. By publishing replication studies, review articles and position papers, the journal also aims at advancing the field's conceptual, theoretical, methodological, and empirical foundations.

However, we should not rest on our laurels. There is still much work to be done, particularly in promoting FAIR data practices and ensuring reproducibility. To start with, we plan to encourage further data sharing by implementing data availability statements for all

publications soon, which will enhance transparency and facilitate the reuse of research data and contribute to cumulative knowledge building.

In the meantime, there are also reasons to rejoice. After a decade of publication, the IJLCR has established a robust presence within the academic landscape. The journal is now indexed in the *Web of Science* (Emerging Sources Citation Index) and other databases, with an impact factor of 1.1, a journal citation indicator of 1.09, and a *CiteScore* of 3.4. While IJLCR's impact factor is lower compared to well-established neighboring journals, this is typical for newer and more specialized journals, which often have lower impact factors as they build their reputation. By contrast, IJLCR's journal citation indicator is relatively close to the average for applied linguistics journals, reflecting a respectable level of citation impact. Similarly, IJLCR's *CiteScore* is competitive with established journals, indicating that its articles are cited frequently relative to the average for the field.

We believe that special issues have been particularly instrumental in shaping IJLCR into a leading journal in applied linguistics. By offering focused thematic collections, they have attracted high-quality research and established the journal's position within the field. We wish to extend our sincere gratitude to the guest editors of the following special issues for their invaluable contributions in establishing the journal's identity and scholarly impact:

- 2016: Linguistic Innovations: Rethinking linguistic creativity in non-native Englishes
(Sandra Deshors, Sandra Götz & Samantha Laporte)
- 2017: Segmental, prosodic and fluency features in phonetic learner corpora
(Jürgen Trouvain, Frank Zimmerer, Bernd Möbius, Mária Gósy & Anne Bonneau)
- 2018: Tense and aspect in Second Language Acquisition and learner corpus research
(Robert Fuchs & Valentin Werner)

- 2019: Corpus-based approaches to spoken L2 Production: Evidence from the *Trinity Lancaster Corpus*
(Vaclav Brezina, Dana Gablasova & Tony McEnery)
- 2021: Natural language processing for learner corpus research
(Kristopher Kyle)
- 2023: Learner translation corpus research
(Sylviane Granger & Marie-Aude Lefer)
- 2024: Learner corpus research for pedagogical purposes
(Sandra Götz & Sylviane Granger)

The journal has also published a number of articles that have attracted particular attention from our readership. The five most frequently cited articles from the IJLCR are listed in Table 1.

Table 1. IJLCR most frequently cited articles as of September 2024

Author(s)	Year	Title	Citations
Sylviane Granger	2015	Contrastive interlanguage analysis: A reappraisal	371
Magali Paquot and Luke Plonsky	2017	Quantitative research methods and study quality in learner corpus research	126
Nicole Tracy-Ventura and Florence Myles	2015	The importance of task variability in the design of learner corpora for SLA research	96
Dana Gablasova,	2019	The <i>Trinity Lancaster Corpus</i> : Development,	85

Vaclav Brezina and Tony McEnery		description and application	
Stefan Th. Gries and Sandra Deshors	2015	EFL and/vs. ESL?: A multi-level regression modeling perspective on bridging the paradigm gap	64

With the recent adoption of the Online First publication model by Benjamins, we anticipate that the journal will experience a significant increase in visibility and impact, facilitated by the accelerated dissemination of research.

In this anniversary issue, we are honored to feature an invited commentary chapter by Sylviane Granger, a pioneering scholar whose work has laid the foundation for current advancements in the field. In this special contribution, she shares her thorough insights and perspectives on the significant developments from first-generation LCR (from the earliest stages in the late 1980s/early 1990s until 2000) to second-generation LCR (covering the early 2000s to the present). Additionally, she outlines future directions she envisions for the field.

The three research articles that follow each exemplify different facets of our vision for the field. Magali Paquot, Alexander König, Egon W. Stemle and Jennifer-Carmen Frey discuss metadata as an essential element in the learner corpus research process, from corpus selection or compilation to study design and interpretation of the results. They present the *Core Metadata Schema for Learner Corpora* and call for more collaborative efforts to improve the robustness and comprehensiveness of metadata in learner corpus research. Sophia Minnillo, Claudia Sánchez-Gutiérrez, Ana Ruiz-Alonso-Bartol, Emily Morgan and Carmen González Gómez investigate the various factors influencing learners' accuracy in using past tense-aspect in L2 Spanish. The authors have been awarded an Open Data badge for their work in recognition of

their commitment to transparency, as they have shared all their data via GitHub. Vildan Özkan Miller and Tove Larsson explore the effect of both linguistic and extralinguistic features on adverb placement in English as a Foreign Language. Similar to Minnillo et al.'s work, their study adopts a multifactorial approach, highlighting the need to consider multiple factors simultaneously when analyzing learner language. Their study is also notable as one of the very few replication studies published in IJLCR.

Book reviews have also always figured prominently in the journal to give our readers informed opinions on the latest publications from our field. Naturally, our anniversary issue also features three book reviews.

As we celebrate this 10-year-milestone, we extend our deepest gratitude to our past and present authors, reviewers, book reviewers, guest editors, former editors, editorial board members, editorial assistants, and readers. Their unwavering support, rigorous scholarship, and enthusiastic engagement have been instrumental in IJLCR's success. We are proud of the vibrant and diverse community that has grown around our journal, and we look forward to continuing this journey together.

We also extend our heartfelt thanks to John Benjamins Publishing Company and especially to Kees Vaes, Susan Hendriks and Heleen Groesbeek for their enthusiasm and dedication in creating and always supporting this journal and helping out with all kinds of issues we run into. Their commitment has been vital to our growth and success.

This introduction is not just a reflection of past achievements but a call to future exploration. Learner corpus research is an ever-evolving field, and we are excited about the new horizons and challenges that lie ahead. We hope this anniversary issue will inspire

continued research, foster collaboration, and drive the next wave of discoveries in learner corpus research. As IJLCR editors, we are extremely enthusiastic and are looking forward to witnessing the continued growth of the field, engaging with emerging trends, and supporting groundbreaking studies that will shape the future of learner corpus research.

Thank you for being a part of our journey. Here is to the next ten years of innovation and excellence in learner corpus research.

References

- Gablasova, D., Vaclav, B., & McEnery, T. (2019). *The Trinity Lancaster Corpus: Development, description and application. International Journal of Learner Corpus Research*, 5(2), 126-158.
- Gilquin, G., De Cock, S., & Granger, S. (Eds.) (2010). *The Louvain International Database of Spoken English Interlanguage. Handbook and CD-ROM*. Presses universitaires de Louvain.
- Granger, S., M. Dupont, F. Meunier, H. Naets, & Paquot, M. (Eds.) (2020). *The International Corpus of Learner English (version 3)*. Presses universitaires de Louvain.
- Ishikawa, S. (2023). *The ICNALE guide: An introduction to a learner corpus study on Asian learners' L2 English*. Routledge.
- Paquot, M., & Callies, M. (2020). Promoting methodological expertise, transparency, replication, and cumulative learning: Introducing new manuscript types in the *International Journal of Learner Corpus Research*. *International Journal of Learner Corpus Research*, 6(2): 121-124.

Paquot, M., & Plonsky, L. (2017). Quantitative research methods and study quality in learner corpus research. *International Journal of Learner Corpus Research*, 3(1), 61-94.