# Using the bilingual Corpus of Romanian Academic Genres (ROGER) platform to improve students' academic writing

Alexandru Oravițan[1], Mădălina Chitez[2],
Loredana Bercuci[3], and Roxana Rogobete[4]

**Abstract**. Learner corpora of written texts from academic writing assignments provide a practical resource for students, particularly in fostering academic writing skills. One such corpus is the newly available ROGER (Corpus of Romanian Academic Genres), a bilingual comparable corpus containing learner discipline-specific academic writing data in Romanian native language (L1) and English as a foreign language (L2). This paper aims to illustrate a series of academic writing teaching approaches supported by the ROGER platform (launched in May 2022) to be applied by tutors in an academic writing classroom setting. The results are structured according to Ädel's (2010) methodological model for fostering rhetorical functions and specific phraseology in academic writing, coupled with addressing metadiscourse markers to better assist in the enhancement of students' academic writing skills at the university undergraduate level.

**Keywords**: academic writing, bilingual corpus, corpus-based teaching, learner corpora.

## 1.    Introduction

While access to general and specialised corpora has improved significantly in recent years, corpus-based teaching in academic writing has yet to be exploited consistently in different student groups. One such case pertaining to teaching

---

1. West University of Timișoara, Timișoara, Romania; alexandru.oravitan@e-uvt.ro; https://orcid.org/0000-0001-5256-6880
2. West University of Timișoara, Timișoara, Romania; madalina.chitez@e-uvt.ro; https://orcid.org/0000-0001-9005-3429
3. West University of Timișoara, Timișoara, Romania; loredana.bercuci@e-uvt.ro; https://orcid.org/0000-0003-1124-6279
4. West University of Timișoara, Timișoara, Romania; roxana.rogobete@e-uvt.ro; https://orcid.org/0000-0002-4155-3918

---

315

English for academic purposes is employing learner corpora, which directly contribute to improving second language teaching and learning through classroom use (Granger, 2002). Learner academic corpora may equally constitute a precise reference point and a practical resource for students (Timmis, 2015), particularly in the context of fostering academic writing skills. However, learner corpora of academic writing are limited in availability: they are frequently not publicly accessible due to copyright laws, commercial reasons (Nesi, 2012), or emerging data protection regulations around the world (such as the General Data Protection Regulation in the EU).

Bilingual learner corpora of academic writing are even less common, even though their added pedagogical benefit is significant. Specifically, these corpora provide students with the opportunity to discover interlanguage interference in a corpus of texts that are at the same level as their own. Previous studies (Bercuci & Chitez, 2019; Chitez & Bercuci, 2019) have shown that, in a Romanian higher education context, such corpora can be utilised as support for self-directed learning and are a source of motivation for students. As far as Romanian L1 implications are concerned, using corpora for academic writing is a complete novelty. Instruments such as academic phrase banks or academic word lists, already developed for English, are non-existent (see Rogobete, Mureșan, & Chitez, in press) because Romanian students' writing skills in L1 are taken for granted. At the same time, a marked tendency towards lingering under the influence of stereotypical language structures has been observed and demonstrated (Tucan, Pop, Bercuci, & Chitez, 2020).
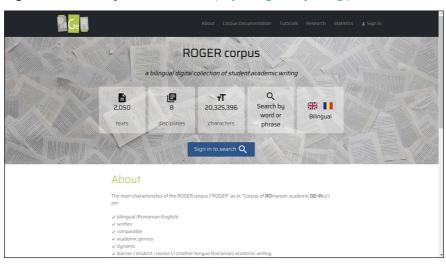
Figure 1.   ROGER platform interface (https://roger-corpus.org/)

Given the limited availability of bilingual learner corpora of academic writing, a research project was launched at the West University of Timișoara, Romania, in 2017, aiming to produce the first bilingual Romanian and English learner corpus of academic writing texts from both undergraduate and graduate students in a variety of disciplines. ROGER (Chitez et al., 2021), with a dedicated open-source query platform (Strilețchi, Chitez, & Csürös, 2022), was made available to the public in late May 2022 (Figure 1 above). This paper aims to illustrate a series of academic writing teaching approaches supported by the ROGER platform.

## 2. A methodological model

A useful methodological model of employing learner corpora to foster academic writing skills can be found in Ädel (2010), who directs attention toward rhetorical functions and specific phraseology as essential features to be mastered by novice academic writers. In what is dubbed as a 'hands-on' model, Ädel (2010) recommends addressing the following research questions to students using learner corpora in an academic writing classroom setting:

> "What do academic writers say when they…? (a) give an example, (b) refer to other texts or researchers, (c) introduce the topic, (d) start their Conclusion section?" (p. 43).

The investigation of these questions through a learner corpus reveals hitherto unnoticed specific phrases and patterns in rhetorical functions and academic phraseology. The replication of this 'hands-on' model may be undertaken by employing the ROGER platform, with specific results showcasing salient dimensions of Romanian L1 and English L2 learner academic writing. Also, aside from Ädel's model, students may discover metadiscourse markers that are erroneously used in L2 by extracting word and N-gram lists in both languages. In teaching activities, students may detect the overuse of specific metadiscourse markers by using an English L1 corpus (such as BNC, COCA, or MICUSP) as a reference point.

## 3. Results

(a) When inputting the Romanian 'de exemplu' (253 occurrences), one notices that the phrase may be encountered either in sentence-initial or mid-sentence positions,

with a wide variety of punctuation patterns: while it is usually preceded by a comma or a full stop, it is followed by an array of punctuation marks, which can be used to address punctuation issues in student writing.

(b) In the case of referring to other texts, materials, or researchers, the parallel querying of the corpus with the study level filter at the undergraduate level showcases significant differences which can be exploited pedagogically (Table 1).

Table 1. Examples of in-text referencing from the ROGER corpus

| RO | BA SY 1 | BA SY 2 | BA SY 3 | EN | BA SY 1 | BA SY 2 | BA SY 3 |
|---|---|---|---|---|---|---|---|
| potrivit lui | 3 | 0 | 2 | according to | 93 | 61 | 42 |
| conform lui | 8 | 0 | 4 | | | | |
| afirm* că | 34 | 4 | 40 | state* that | 27 | 20 | 28 |
| consideră că | 34 | 3 | 31 | considers that | 4 | 7 | 4 |
| prezentat* în | 8 | 1 | 18 | present* in | 23 | 38 | 35 |
| menționat* în | 2 | 1 | 5 | mention* in | 8 | 15 | 25 |
| preluat* din | 2 | 1 | 5 | quoted in | 0 | 1 | 2 |
| citat* în / din | 0 | 0 | 0 | cited in | 0 | 3 | 3 |

(c) The introduction to any academic paper highlights several phrases which can be elicited by students, particularly in connection to the thesis statement. Thus, students may be asked to predict patterns headed by 'this paper *' for English and 'această lucrare' for Romanian, and to compare and contrast their examples with concordance lines obtained when querying the corpus (Figure 2 below).

(d) As concerns the insights that students might gain from how academic writers start their conclusion section, the querying of the keywords 'concluzie' for Romanian (271 occurrences) and 'conclusion' for English (511 occurrences) illustrates that both of these structures are usually preceded by a preposition or preposition plus article combination ('în'/'ca o' for Romanian, 'in'/'as a' for English), followed by a comma and a summarising/concluding statement.

In the case of metadiscourse markers, a persistent structure used in Romanian academic writing is 'din punct de vedere'. In the ROGER corpus, the phrase appears 20 times both in Romanian and in English. However, in English L2, the phrase 'point of view' is often overused in the following configuration: 'from a * point of view' by Romanian native speakers.

Figure 2.   Search results for 'this paper' (EN) and 'această lucrare' (RO) in the
ROGER corpus





## 4.      Discussion and conclusion

The approaches presented in this short paper illustrate how the ROGER platform
becomes an important resource for showcasing key academic phraseology in
authentic bilingual contexts. Thus, valuable insights may be gained regarding

punctuation errors in (a) and (d), cases of cross-contamination (such as 'consideră că' – 'considers that' and 'prezentat* în' – 'present* in'), which may be studied separately, but also situations that could be approached in individual concordance lines to observe citation and acknowledgement practices (b), or frequent action verbs and patterns in use when writing thesis statements (c).

The results are structured according to Ädel's (2010) methodological model for fostering rhetorical functions and specific phraseology in academic writing and suggest adding another function addressing metadiscourse markers to better assist with enhancing undergraduate university students' academic writing skills. In sum, we recommend the use of bilingual learner corpora as a tool for improving and enhancing academic writing skills, thus building a legitimate contrastive pedagogical model.

## 5.    Acknowledgements

## References

Ädel, A. (2010). Using corpora to teach academic writing: challenges for the direct approach. In M. C. Campoy-Cubillo, B. Bellés Fortuño & M. L. Gea-Valor (Eds), *Corpus-based approaches to English Language Teaching* (pp. 39-56). Bloomsbury.

Bercuci, L, & Chitez. M. (2019). A corpus analysis of argumentative structures in ESP writing. *International Online Journal of Education and Teaching*, *6*(4), 733-747. http://iojet.org/index.php/IOJET/article/view/655

Chitez, M., & Bercuci, L. (2019). Data-driven learning in ESP university settings in Romania: multiple corpus consultation approaches for academic writing support. In F. Meunier, J. Van de Vyver, L. Bradley & S. Thouësny (Eds), *CALL and complexity – short papers from EUROCALL 2019* (pp. 75-81). Research-publishing.net. https://doi.org/10.14705/rpnet.2019.38.989

Chitez, M., Bercuci, L., Dincă, A., Rogobete, R., & Csürös, K. (2021). Corpus of Romanian Academic Genres (ROGER). https://roger-corpus.org/

Granger, S. (2002). A bird's eye view of learner corpus research. In S. Granger, J. Hung & S. Petch-Tyson (Eds), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 3-33). Benjamins. https://doi.org/10.1075/lllt.6.04gra

Nesi, H. (2012). ESP and corpus studies. In B. Paltridge & S. Starfield (Eds), *The handbook of English for specific purposes* (pp. 407-426). John Wiley & Sons. https://doi.org/10.1002/9781118339855.ch21

Rogobete, R., Mureşan, V., & Chitez, M. (in press). Academic word lists in English and Romanian: a corpus-based contrastive analysis. *Quaestiones Romanicae, X.* West University of Timişoara Press.

Strileţchi, C., Chitez, M., & Csürös, K. (2022). Building roger: technical challenges while developing a bilingual corpus management and query platform. *Proceedings of the 17th International Conference on Software Technologies (ICSOFT 2022)* (pp. 390-398). https://doi.org/10.5220/0011144500003266

Timmis, I. (2015). *Corpus linguistics for ELT. Research and Practice*. Routledge. https://doi.org/10.4324/9781315715537

Tucan, D., Pop, A.-M., Bercuci, L., & Chitez, M. (2020). Writing transition: a corpus-supported study of the freshmen's written argumentation competence building. *Conference Proceedings, 13th International Conference Innovation in Language Learning* (pp. 62-66). Filodritto Editore. https://doi.org/10.26352/EY12-2384-9509

**Intelligent CALL, granular systems and learner data: short papers from EUROCALL 2022**
**Edited by Birna Arnbjörnsdóttir, Branislav Bédi, Linda Bradley, Kolbrún Friðriksdóttir, Hólmfríður Garðarsdóttir, Sylvie Thouësny, and Matthew James Whelpton**

**Publication date: 2022/12/12**

**Disclaimer**: Research-publishing.net does not take any responsibility for the content of the pages written by the authors of this book. The authors have recognised that the work described was not published before, or that it was not under consideration for publication elsewhere. While the information in this book is believed to be true and accurate on the date of its going to press, neither the editorial team nor the publisher can accept any legal responsibility for any errors or omissions. The publisher makes no warranty, expressed or implied, with respect to the material contained herein. While Research-publishing.net is committed to publishing works of integrity, the words are the authors' alone.

**Trademark notice**: product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

**Copyrighted material**: every effort has been made by the editorial team to trace copyright holders and to obtain their permission for the use of copyrighted material in this book. In the event of errors or omissions, please notify the publisher of any corrections that will need to be incorporated in future editions of this book.