

Fostering learning with the EcoLexicon corpus in the ESP classroom

Maria Rudneva¹

Abstract. This pilot study provides a preliminary account on students' attitudes toward using specialized corpora in English for specific purposes (ESP) classes. Learners (N=39) were introduced to the EcoLexicon corpus and trained to use its basic query tools. The rationale behind this activity was to introduce learners to contextualization patterns and genre-specific features of the professional target language, which in its turn would ensure acceptability and appropriateness of their linguistic choices. The learners were offered a series of guided and independent tasks on terminology disambiguation and corpus-assisted speech production. At the end of the semester a survey was administered to the students to assess their perception of hands-on corpus experience. Descriptive statistics show preliminary evidence that corpus tools provide illuminating data, foster understanding of nuances within synonymous groups of words, and increase overall language awareness. However, hands-on Data Driven Learning (DDL) experience presented a few challenges which, however, may be remedied by careful design of teaching materials and assignments.

Keywords: DDL, specialized corpora, ESP.

1. Introduction

Introducing corpora in language learning draws on the DDL (Johns & King, 1991). Various tools and methodologies are widely used in DDL, this list includes but is not limited to language for specific purposes, frequency lists and learner corpora, error correction and contrastive analysis, corpus use in syllabus design, etc. (Boulton, 2017). According to meta-analysis of quantitative DDL studies by Boulton and Cobb (2017) this approach offers numerous benefits on

1. RUDN University, Moscow, Russia; rudneva-ma@rudn.ru; <https://orcid.org/0000-0003-1161-9100>

How to cite: Rudneva, M. (2020). Fostering learning with the EcoLexicon corpus in the ESP classroom. In K.-M. Frederiksen, S. Larsen, L. Bradley & S. Thoušný (Eds), *CALL for widening participation: short papers from EUROCALL 2020* (pp. 304-308). Research-publishing.net. <https://doi.org/10.14705/rpnet.2020.48.1205>

the instructor's side, fosters efficient language acquisition and develops students' analytical and problem-solving skills as well as learner autonomy (Vyatkina & Boulton, 2017).

Potential limitations of DDL are associated with the complex interface of many available corpora, as they were designed by linguists for linguists; therefore, introduction of corpus query tools to non-linguist students requires preliminary training. Apart from that, numerous examples derived from large corpora might be misleading for learners. On the instructor's side corpus-based pedagogic design requires considerable preparation time. Having said that, from the learners' perspective the level of language proficiency might be an obstacle to direct implementation of corpus data, as the query output needs to be carefully tailored and softened for novice learners.

Today DDL research falls into three major categories:

- learner corpora research (analysis of learners' oral and written production);
- corpus-based pedagogic material design; and
- inductive learning (hands-on experience of learners with existing or specifically designed corpora) (Boulton, 2017).

The intention of this project was to investigate if specialized language corpora belong in an ESP classroom, what their perceptions are of the hands-on DDL experience and to discuss potential limitations as well as the ways to address them.

2. Method

This study collected preliminary data regarding learners' (N=39) attitudes toward using a freely available corpus as a lexicographic reference tool in their ESP/translation studies classes. All participants were offered preliminary training on Sketch Engine basic query functions. This was followed by a series of guided search activities aimed at key terminology disambiguation. At the final stage the search activities were assigned as weekly homework with on-site follow-up discussion. At the end of the semester the students were requested to respond to an anonymous questionnaire grading their experience, with open answer options. Microsoft Excel was employed to analyze data using descriptive statistics.

2.1. Participants

39 RUDN university students (median age=19) took part in this project; their English proficiency is B2-C1 CEFR, according to the results of Cambridge English exams. All participants are environmental sciences majors. All students are enrolled in double diploma programs and minor in specialized translation.

2.2. Instruments and procedure

During the spring semester 2018-2019, three groups of students aged 18-20 were offered to use EcoLexicon online tool as a lexicographic reference source in cases when bilingual and monolingual dictionaries failed to provide clear understanding of meaning or usage differences between near-synonymic words.

EcoLexicon is a corpus of contemporary environmental texts, the size is 23 million words and it is an extensive terminological knowledge base on the environment (León-Araúz, Martín, & Reimerink, 2018). It is available for access and query in the corpus query system Sketch Engine.

The students were introduced to the basic features of Sketch Engine analysis tool and pre-taught to use it. The project lasted 16 weeks: 4 weeks of introduction and guided practice, 10 weeks of independent practice with on-site follow-up, and 2 weeks of evaluation.

3. Results and discussion

3.1. Corpora for ESP

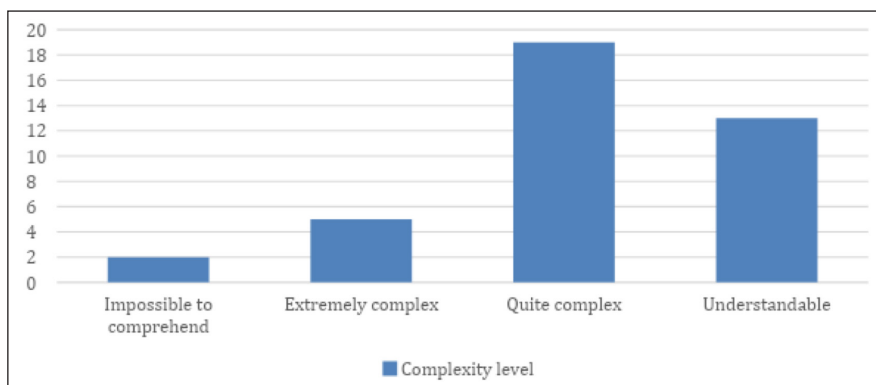
In an ESP class a specialized corpus can serve as a unique tool for overcoming existing asymmetries in terminological systems of source and target languages. The existence of such asymmetries often draws on extra-linguistic factors, e.g. numerous nature conservation technologies are not yet implemented in Russia which is directly reflected in learners' source language. Prospective specialized translators need to patch numerous lexical gaps by creating new terms in the L1. In this sense it is essential for language for specific purposes instructors to provide novice specialists with reliable tools and means that would facilitate creation of precise, accurate, and non-ambiguous terms. Therefore, LSP learners need specific corpus tool training to be able to make informed linguistic choices in future.

3.2. Learners' perception of DDL experience

Descriptive statistics provided preliminary results on the perception of corpus tools by the learners. Figure 1 is a histogram of the distribution of learners' perception of the complexity level of corpus-based assignments. The perception survey was administered to gather informal feedback on the project at its preliminary stage to explore principal possibility of corpus-based activities for non-linguist students. The results are provided here to illustrate the outcome of the project; however, the author intends to address the perception issue in more detail in future research.

The majority of respondents (N=19) considered lexicographic assignments quite complex, the second biggest category (N=13) considered the assignments understandable, few learners considered corpus tools extremely complex (N=5) or impossible to comprehend (N=2).

Figure 1. Complexity level



However, the majority of the participants (85%) acknowledged that corpus tools were helpful for terminology disambiguation. Among their comments were: “truly illuminating”, “like a linguistic detective”, “seems reliable reference source”, and “sometimes might be useful”. The remaining 15% were overall reluctant to master corpus tools commenting as “why do we need to do this at all”, and even “holy mother of god, get me out of this”.

3.3. Limitations and possible solutions

The survey also asked *Do you see any challenges in using a specialized corpus?* The answers can be subdivided into three respective categories. First of all,

insufficient language proficiency might be a considerable pitfall; exposure to authentic professional contexts can be discouraging for lower level students. The solution here might be to design instructor-guided activities, simplify and tailor tasks. Secondly, non-linguist students in general demonstrate less interest in lexicographic discoveries. Therefore, it might be a good idea to introduce corpus tools gradually and only when other reference sources are of no help. Thirdly, however user-friendly corpus query interface is, learners find it challenging. To address this issue the instructor needs to pre-teach and guide search activities.

4. Conclusions

Corpus tools have immense potential for providing precise, accurate, and non-ambiguous data on specific terminology in professional contexts. Increasing availability of specialized corpora holds great promise of new advances for ESP learners, shifting the pedagogic focus from prescribed vocabulary lists to inductive learning and learner autonomy. Overall, the learners demonstrated positive attitude toward hands-on corpus-based experience. Potential limitations of the approach, such as insufficient language proficiency, low motivation, and complexity of user interface can be remedied by thoughtful pedagogic design. It might be of interest for further research to develop a systematic approach to overcoming terminological asymmetries of source and target professional language by means of corpus tools.

References

- Boulton, A. (2017). Corpora in language teaching and learning. *Language Teaching*, 50(4), 483-506. <https://doi.org/10.1017/S0261444817000167>
- Boulton, A., & Cobb, T. (2017). Corpus use in language learning: a meta-analysis. *Language Learning*, 67(2), 348-393.
- Johns, T., & King, P. (1991). Classroom concordancing. *English Language Research Journal*, 4, 1-16.
- León-Araúz, P., Martin, A. S., & Reimerink, A. (2018). The EcoLexicon English corpus as an open corpus in Sketch Engine. *arXiv preprint arXiv:1807.05797*.
- Vyatkina, N., & Boulton, A. (2017). Corpora in language learning and teaching. *Language Learning & Technology*, 21(3), 1-8. <https://doi.org/10125/44750>



Published by Research-publishing.net, a not-for-profit association
Contact: info@research-publishing.net

© 2020 by Editors (collective work)
© 2020 by Authors (individual work)

CALL for widening participation: short papers from EUROCALL 2020
Edited by Karen-Margrete Frederiksen, Sanne Larsen, Linda Bradley, and Sylvie Thouéšny

Publication date: 2020/12/14

Rights: the whole volume is published under the Attribution-NonCommercial-NoDerivatives International (CC BY-NC-ND) licence; **individual articles may have a different licence.** Under the CC BY-NC-ND licence, the volume is freely available online (<https://doi.org/10.14705/rpnet.2020.48.9782490057818>) for anybody to read, download, copy, and redistribute provided that the author(s), editorial team, and publisher are properly cited. Commercial use and derivative works are, however, not permitted.

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.

Typeset by Research-publishing.net

Cover theme by © 2020 Marie Flensburg (frw831@hum.ku.dk), based on illustration from [freepik.com](https://www.freepik.com)
Cover layout by © 2020 Raphaël Savina (raphael@savina.net)

ISBN13: 978-2-490057-81-8 (Ebook, PDF, colour)

British Library Cataloguing-in-Publication Data.

A cataloguing record for this book is available from the British Library.

Legal deposit, France: Bibliothèque Nationale de France - Dépôt légal: décembre 2020.