https://doi.org/10.53656/for21.45corp

INTRODUCTION OF ONLINE LANGUAGE CORPORA INTO FUTURE TRANSLATORS' AND INTERPRETERS' TEACHING

Dr. Elena Zelenina, Assoc. Prof.
Federal State Budgetary Educational Institution of Higher Education "Tula State University" (Russia)

Abstract. This paper outlines some benefits from online language corpora use in translators' and interpreters' teaching, especially its advantages for developing target language competences, translation competences and scientific research competences of future translators and interpreters. The author shares the results of her five year theoretical and empirical studies in training students to apply online language corpora in doing their language research projects, as well as offers rationalizing contributions to existing pedagogical language corpora applications in university teaching practice. The main resultant outcome of the author's scientific pedagogical studies has become the Research-Oriented Methodology of Gradual Online Language Corpora Introduction.

Keywords: online language corpus; future translators and interpreters; higher education, competences, research.

Introduction

Language corpora, large textual databases (Leńko-Szymańska & Boulton 2015, 1), have been gradually becoming an important part of language researchers' and language teachers' work. Many modern online corpora tools provide researchers with an instant access to the information on a word functioning in different contexts and collocations, being a part of wordlists, grammatical patterns and stylistic profiles. It can be safely claimed that since their introduction, corpora have been changing not only research practices of linguistic analyses that now, due to corpora data, rely on attested examples of language use and of account for frequency phenomena, but also have been having a considerable influence on language and translation pedagogy.

Developing Competences through Online Language Corpora

The conceptual basis for understanding the nature and components of competences of future translators and interpreters is laid by the Russian Federal State Educational Standard of Higher Education for Degree 45.03.02 "Linguistics"[1]. According to this Standard, in the process of the university study a student acquires various competences

that comprise target language competences, translation competences, scientific research competences, cultural and ethical competences, information competences and others.

These competences, provided with descriptions, serve as clear fundamental educational goals both for students and university teachers.

However, considering the number and complex character of the aforementioned competences acquired by university students, further the article will concentrate mostly on target language competences, translation competences and scientific research competences of future translators and interpreters.

Developing Target Language Competences

Target language competences comprise competences which, after being considered thoroughly, seem to be connected not only with purely linguistic knowledge and skills such as lexical, phonological, grammatical, syntactical, but also can represent the ability of a person to produce texts in the target language, characterized by the correct use of collocations and idioms and by the level of authenticity that appeals to native speakers of the target language.

Writers on translation and interpretation pedagogy often tend to emphasize the effect of language corpora use on learning the target language that reveals in enriching the students' repertoires of phraseological items, rules of collocations as well as grammatical patterns of the language (Aston, 2015; Chujo, Oghigian & Akasegawa 2015).

As Thomas states, full texts provided by corpora afford language learners "many opportunities to observe a wide range of linguistic features whose typicality they can ascertain through corpus searches" (Thomas 2015, 85).

These and other examples prove how online language corpora can effectively help learners to improve their foreign language knowledge.

Developing Translation Competences

Over the last few decades there were a lot of publications on using language corpora for teaching translation and interpreting.

For example, Molés-Cases and Oster propose corpus-based webquests: this learning task format involves students in the investigation of open-ended questions (Molés-Cases & Oster 2015).

Marco and van Lawick describe how the use of comparable corpora can enhance translator trainees' awareness of source text interference in their translations (Marco and van Lawick 2015).

The main pedagogical trend, observed in various studies, seems to be the development of data-driven tasks aimed at enhancing general language awareness of the students and helping them to develop their professional expertise by analyzing multiple examples of different translation phenomena in online language corpora.

Developing Scientific Research Competences

From the perspective of a wide range of language corpora application for scientific purposes, O'Keefe, McCarthy and Carter outline several spheres (O'Keefe, McCarthy & Carter 2007): Lexicography, Grammar, Stylistics, Translation,

Forensic Linguistics, Sociolinguistics. According to the authors, language corpora can provide information that helps to examine word frequency, grammar patterning and semantics, present literary contexts that surround a particular word or word combination, concordancing extracts from original and translated texts and other relevant to a researcher data.

Káňa presents a collection of ready-made models for conducting student research based on the German language corpora (Káňa 2014) as well as the examples of research projects with explanations of logics of conducting such a project.

Summing it up it seems possible to conclude that researchers and university teachers consider ways in which online language corpora can be used in training research skills.

The Characteristics of the Author's Research-Oriented Methodology of Gradual Online Language Corpora Introduction

Working at the problem of rationalizing existing pedagogical language corpora applications the author developed the Research-Oriented Methodology of Gradual Online Language Corpora Introduction into university teaching of future translators and interpreters (herein also termed "the Methodology").

By the Research-Oriented Methodology of Gradual Online Language Corpora Introduction is meant a balanced complex of pedagogical approaches, educational principles, instructional techniques, practical teaching procedures, aimed at organizing and structuring students' initial research practices based on online language corpora use so as they could act independently and effectively to accomplish their current and future research goals.

The Methodology suggests using online language corpora for classroom activities and for training students to apply language corpora in doing research projects and is designed to build a bridge between real life professional contexts, real life linguistic research contexts and teaching translation and interpreting environments.

The Methodology

– creates situations in which students explore online language corpora data themselves directly. This contributes to enhancing of the development of target language competences due to the "discovery effect", when a student translator/interpreter "discovers" examples of the target language functioning as grammar patterns, idiomatic combinations in original authentic contexts;

– suggests the development and implementation of data-driven tasks in translation analysis aimed at promoting general language awareness of students and helping them to develop their professional expertise by analyzing multiple examples of different translation phenomena in online language corpora;

– suggests that students explore language corpora data for research purposes and create their course research projects with their help. By introducing the corpus analysis to students doing their scientific research work, the Methodology positive-

ly contributes to enhancing students' prospects as future researchers and scientists. It provides them with reliable empirically tested models of collecting and analyzing linguistic data and gives the experience of interaction with language corpora, thus purposefully assisting to the development of scientific research competences.

The main pedagogical principles of the Research-Oriented Methodology of Gradual Online Language Corpora Introduction:

1. Research-oriented teaching approach

This approach reflects Humboldt's ideal of education in balance with scientific research, supported by modern researchers (Thiel & Böttcher 2014). This means that the Research-Oriented Methodology of Gradual Online Language Corpora Introduction into university teaching of future translators and interpreters intends to encourage students to carry out scientific research acting as independent researchers, respecting their scientific work as a process and a result, thoroughly developing the research methodology, carefully interpreting their results and openly discussing their outcomes.

2. The combinatorial character

The Methodology supposes both direct and indirect uses of online corpora following a distinction supported by Leech (Leech, 1997), Römer (Römer 2011). This means that the Methodology is designed to create situations, in which students are taught with the help of specially developed corpus-based teaching materials and tasks presuming previous interaction between a teacher and a corpus that serve models for students' work with corpora (indirect application), and situations, in which students explore corpus data themselves for various educational and research purposes as a part of interaction between a student and a corpus (direct application).

3. The process-oriented approach

Adopting a process-oriented approach here supposes that the teacher should focus on the training and procedural aspects of the online language corpora use by the students, not only on the end products. Along with the creation of data-driven corpus-based tasks, methodological guidance should be provided as how to achieve quality in performing these tasks, how to select the optimal corpus, how to plan and implement strategies and tactics of work with the corpus, how to identify the reasons for errors, etc.

4. The reliance on the data visualization

Students are likely to learn to use online corpora mechanisms and tools faster if these are presented to them visually, not only verbally. The inclusion of various types of infographics into the first introductory lessons that contain methodological guidance on online language corpora intensifies the teaching effect on student translators and interpreters and better illustrates the significance of corpus-based practices.

5. The progressive character

Gradual learning by small measured steps, supplemented by the teacher's instructing, commenting, analyzing and correcting of errors serves, in particular at the initial stages of university training, as a basement for more intensive teaching and the acquisition of more extensive and precise linguistic information by the students later on the course.

6. The multidisciplinary approach

In order to ensure the progress of the students' acquisition of skills in online language corpora use it seems to be logical to introduce corpora-related tasks and projects into teaching different disciplines from the very beginning of the educational process, and not only into teaching theoretical disciplines connected with the linguistic theory and research, but also into practical disciplines, like the Practical Course of the English Language, the Practical Course of the German Language, the Practical Course of Interpreting and Translating.

7. The student-centered approach

The idea of this approach is based on active encouragement of students in developing and implementing their own strategies and tactics of work with online language corpora, in order to help them become more creative in solving any corpus-related research problem.

The Research-Oriented Methodology of Gradual Online Language Corpora Introduction involves various types of educational activities developed on the basis of the general principles of organizing tasks for English language teaching (Ur 2012).

The set of curricular and extracurricular activities for university teachers includes:

– preparing presentations to introduce basic online language corpora tools into everyday university teaching practice;

– the design of classroom tasks for students that can show and clarify the principles of work with corpora data and the ways these data could be processed, analyzed and presented;

– ongoing consultation on the online language corpora use and collaboration with the students to help them to produce better results than students could have done on their own;

– active stimulating students' initiatives to create their own ideas of corpora-based assignments for their classmates in response to the teacher's tasks by engaging students into game-like activities;

– developing and giving to the students creative assignments and projects based on an extensive online language corpora use that can provide students with knowledge of extra opportunities for linguistic studies outside the classroom and foster students' ability to work on their own as autonomous researchers.

The main types of curricular and extracurricular activities for students include:

– intense perception and analytical activity in the process of introduction of the new information about basic online language corpora tools so as to be able to think about linguistic categories, terms, connections etc.;

– interpersonal communication and interaction in the process of doing new types of data-driven tasks presented by the teacher so as to be able to share, analyze and discuss ideas, put forward hypotheses;

– ongoing collaboration with the teacher who can consult on the online language corpora use to produce a better result than students could have achieved working without the teacher's guidance;

– generating and implementing their own ideas of corpora-based assignments in response to the tasks given by the teacher;

– completing creative assignments and projects based on an extensive online language corpora use that can provide students with knowledge of extra opportunities for language studies outside the classroom and let them progress independently of the university teacher in the future.

Examples of the Author's Corpora-Oriented Teaching Materials

The first example is the author's Algorithm of an Online Language Corpus Selection (herein also termed "the Algorithm"). It was developed with the aim to help students in selecting the optimal online language corpus for their linguistic research as a component of course works consultation session for the disciplines "Basic Theory of the German Language" and "Basic Theory of Linguistics" (Figure 1). In terms of corpora types and definitions the Algorithm relies upon the typologies presented in the works by O'Keeffe, McCarthy and Carter (O'Keeffe, McCarthy & Carter 2007) and Káňa (Káňa 2014).

The implementation of this Algorithm into university teaching practice have resulted in the increase of students' understanding of online language corpora functions, more organized work with language corpora during doing their research projects and in higher performance and motivation in the disciplines in which language corpora use was introduced with the help of this Algorithm.

Another example presents the author's scheme "Ten Steps to Successful Language Corpora Using" intended for students who want to enrich their linguistic research projects with authentic language patterns.

Step 1. Study the rules of the corpora use in the manual on its website.

Step 2. Try various combinations and possibilities of search to study the corpus' interface better.

Step 3. Outline characteristics of the language units being the object of your research project that can be illustrated with the corpus' search results. This information can be very supportive for designing and arranging the empirical part of the project.

Step 4. State what kinds of linguistic contexts are provided in the corpus, what literary, newspaper or Internet sources are used there, what registers of speech are illustrated. This information can be often supportive for specifying the topic of the research and achieving a more logical structure of the empirical part of the research.

Step 5. Choose the criteria of selecting examples from the corpus based on the general concept of the research project, its theoretical grounds and characteristics of the language units provided by the corpus' search mechanisms.

Step 6. Compile the language material provided by the corpus that is necessary to confirm the hypothesis of the research project using the previously chosen criteria.

Step 7. Analyze the obtained language material for further possible classification considering the theoretical aspects studied in the research project.

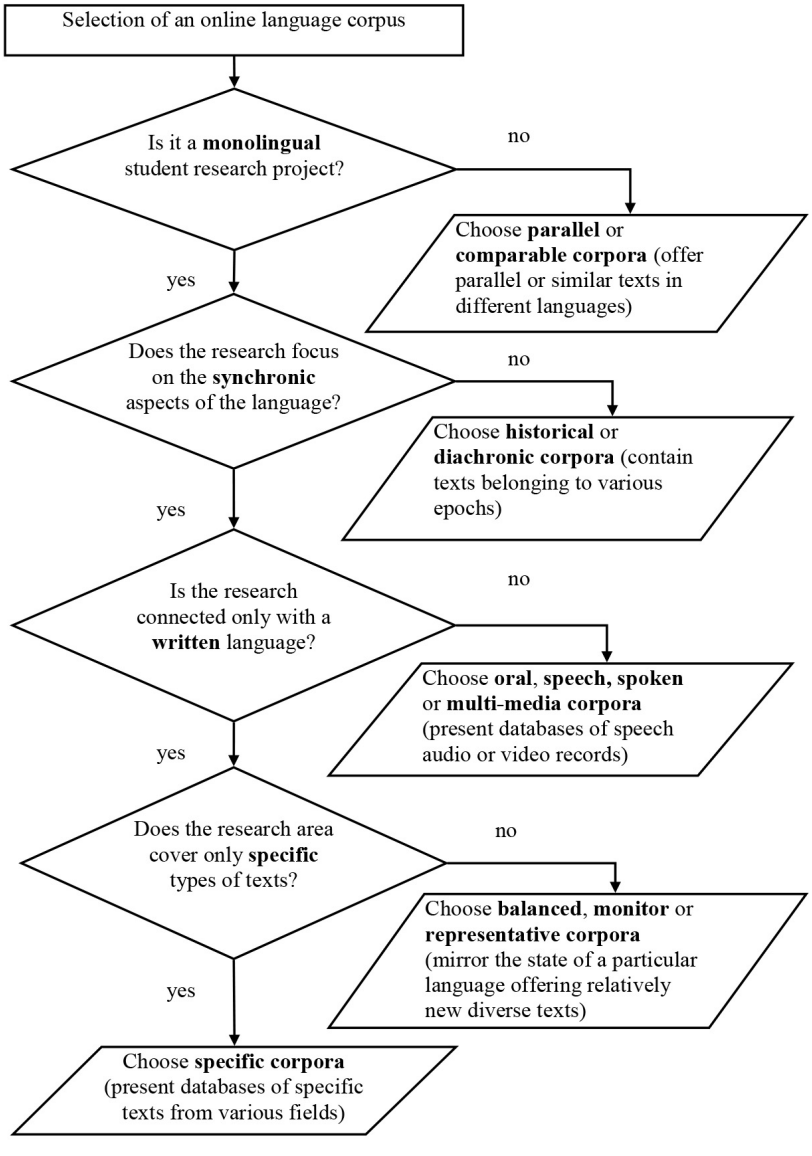

Figure 1. The Algorithm of an Online Language Corpus Selection

Step 8. Study the supplementary information on the collected language units provided by the corpus, especially various statistical data, that can make your research more thorough.

Step 9. Study the examples (if any) of the language units obtained with the help of the corpus that may somehow contradict to the expected results of the research and try to deduce whether they may serve as a ground for a constructive academic discussion or become a material for a future separate in-depth research.

Step 10. Include the most illustrative examples of the language units obtained with the help of the corpus (with references to the source), as well as respective scientific descriptions made by you on the basis of the actual linguistic theories in the empirical part of the research project.

From the experience of the application of this scheme, it seems possible to state that the scheme can be used as a part of Power Point presentation for initial instruction of the students, who are to prepare their first research projects, or as a variant of a printed checklist: students may check stages of their research and tick the actions they have already done.

Describing the experience of online language corpora introduction to student translators and interpreters and provision of methodological support for corpus-based projects it is important to illustrate not only carefully developed and practically attested examples of the corpora-related teacher's instructing materials but also examples of the corpus-related assignments developed by students.

For example, students can be offered to prepare mini research projects (individual, in pairs or in groups) as homework united by one theme ("Idioms", "Translator's False Friends", "Polysemantic Lexical Units", etc.) according to the following plan:

1) make the list of the language units studied (for example, the list of idioms, the list of translator's false friends);

2) choose the language corpus that can provide the necessary information and linguistic contexts;

3) collect examples of linguistic contexts containing the necessary language units from the chosen corpus;

4) develop a package of corpus-related tasks from the language material collected from the corpus and present it in the form of Power Point presentation in class.

When interpreting and translating exercises (that make up the core of the curriculum) are systematically accompanied by supplementary corpus-related creative assignments developed by the teacher and students, this contributes to maintaining a more creative and at the same time more comfortable atmosphere in the classroom, allowing the students possibility to feel more confident as language explorers and more responsible as target text producers.

Conclusion

The results of the implementation of the author's Research-Oriented Methodology of Gradual Online Language Corpora Introduction into university teaching of future translators and interpreters suggest that it effectively, with the improvement

in academic performance, contributes to the development of target language competences, translation competences and scientific research competences, as well as actively fosters a wide range of related skills and abilities:

– target language enhancement in general;

– in-depth understanding of the essence of theoretical phenomena and their practical representation;

– better understanding and adopting translation strategies and tactics;

– accelerated development of the skills directly and indirectly linked to the translation process: analytical skills, decision-making skills, editing skills, etc.;

– recognition of the necessity of intensive scientific research work;

– development of effective strategies and tactics of scientific research work based on the use of online language corpora.

The application of the author's Research-Oriented Methodology of Gradual Online Language Corpora Introduction also motivates students for further scientific research work of a more advanced level such as writing graduation works, scientific articles and work in scientific communities.

NOTES

1. Ministry of Education and Science of the Russian Federation, 2014. *Russian Federal State Educational Standard of Higher Education for Degree 45.03.02 "Linguistics"*. Retrieved from: http://fgosvo.ru/uploadfiles/fgosvob/450302_Lingvistika.pdf.

REFERENCES

Aston, G., 2015. Learning Phraseology from Speech Corpora. In: Leńko-Szymańska, A. & Boulton, A. (Eds.) *Multiple Affordances of Language Corpora for Data-driven Learning,* 65-84. Amsterdam/Philadelphia: Benjamins.

Chujo, K., Oghigian, K & Akasegawa, S., 2015. A Corpus and Grammatical Browsing System for Remedial EFL Learners. In: Leńko-Szymańska, A. & Boulton, A. (Eds.) *Multiple Affordances of Language Corpora for Data-driven Learning,* 109 – 128. Amsterdam/Philadelphia: Benjamins.

Káňa, T., 2014. *Sprachkorpora in Unterricht und Forschung DaF/DaZ*. Brno: Masaryk-Universität.

Leech, G., 1997. Teaching and language corpora: A convergence. In: Wichmann, A., Fligelstone, S., McEnery, T. & Knowles, G. (Eds.) *Teaching and Language Corpora,* 11 – 23. London/New York: Longman.

Leńko-Szymańska, A. & Boulton, A., 2015. Introduction. Data-driven Learning in Language Pedagogy. In: Leńko-Szymańska, A. & Boulton,

A. (Eds.) *Multiple Affordances of Language Corpora for Data-driven Learning,* 1 – 14. Amsterdam/Philadelphia: Benjamins.

Marco, J. & van Lawick, H., 2015. Enhancing Translator Trainees' Awareness of Source Text Interference through Use of Comparable Corpora. In: Leńko-Szymańska, A. & Boulton, A. (Eds.) *Multiple Affordances of Language Corpora for Data-driven Learning*, 225 – 244. Amsterdam/Philadelphia: Benjamins.

Molés-Cases, T. and Oster, U., 2015. Webquests in Translator Training Introducing Corpus-based Tasks. In: Leńko-Szymańska, A. & Boulton, A. (Eds.) *Multiple Affordances of Language Corpora for Data-driven Learning,* 201 – 224. Amsterdam/Philadelphia: Benjamins.

O'Keeffe, A., McCarthy, M. & Carter, R., 2007. *From Corpus to Classroom: language use and language teaching*. Cambridge: Cambridge University Press.

Römer, U., 2011. Corpus Research Applications in Second Language Teaching. *Annual Review of Applied Linguistics* 31, 205 – 225.

Thiel, F. & Böttcher, F., 2014. Modellierung fächerübergreifender Forschungskompetenzen. In: Berendt, B., Szczyrba, B., Voss, H. P. & Wildt, J. (Eds.) *Neues Handbuch Hochschullehre,* 109 – 124. Berlin: Raabe.

Thomas, J., 2015. Stealing a March on Collocation: Deriving Extended Collocations from Full Text for Student Analysis and Synthesis. In: Leńko-Szymańska, A. & Boulton, A. (Eds.) *Multiple Affordances of Language Corpora for Data-driven Learning,* 85 – 108. Amsterdam/Philadelphia: Benjamins.

Ur, P., 2012. *A Course in English Language Teaching*. Cambridge: Cambridge University Press.

✉ **Dr. Elena Zelenina, Assoc. Prof.**
https://orcid.org/0000-0001-8905-5881
eLibrary Science Index SPIN-code: 5689-4225
https://www.elibrary.ru/author_profile.asp?authorid=732296
Department of Linguistics and Translation
Institute of Humanities and Social Sciences
Federal State Budgetary Educational Institution of Higher Education "Tula State University"
92, Prospect Lenina
300012 Tula, Russia
E-mail: ZeleninaE@yandex.ru