VYTAUTO DIDŽIOJO
UNIVERSITETO
ŠVIETIMO
AKADEMIJA

Discourse-Annotated Corpus for Foreign Language Teaching/Learning: Prior Knowledge Meaning

Giedrė Valūnaitė Oleškevičienė[1], Guilhermina Lobato Miranda[2], Liudmila Mockienė[3], Dalia Gulbinskienė[4]

[1]   Mykolas Romeris University, Ateities g. 20, LT-08303 Vilnius, Lithuania, gvalunaite@mruni.eu
[2]   Lisbon University, University Alley, PRT-1649-013 Lisbon, Portugal, gmiranda@ie.ulisboa.pt
[3]   Mykolas Romeris University, Ateities g. 20, LT-08303 Vilnius, Lithuania, liudmila@mruni.eu
[4]   Vilnius Gediminas Technical University, Saulėtekio al. 11, LT-10223 Vilnius, Lithuania, dalia.gulbinskiene@ vilniustech.lt

Annotation. The process of teaching and learning is characterized by the usage of prior knowledge schemas in the procedure. The research aims to explore the meaning of initial knowledge in the process of teaching and learning a foreign language at more advanced levels while developing and using discourse-annotated corpora. The research reveals two essential dimensions-organizational and personal, which structurally represent the experience of the research participants.

Introduction

Discourse-annotated corpus is a collection of written texts on a particular subject, and the practice of adding interpretative, linguistic information to an electronic corpus of written language data. The idea advocated in this article is that discourse-annotated corpora and the development of such corpora can be used as a supplement in the teaching and learning of a language, especially for the more advanced learners when dealing with the textual cohesion and coherence. However, corpora development and annotation pose an intense challenge during the teaching and learning processes which requires initial linguistic knowledge of syntax, meta-language on types of connectives and coherence

relations. During the corpus annotation process, students also need to master a certain software program designed for corpus annotation, which is challenging as well. Since students need to apply information technologies during the process, the current study is based on the idea of the authors of the four-component instructional design (4C/ID) model (van Merriënboer et al., 2003; van Merriënboer & Kester, 2014; Frerejean et al., 2019) that teaching and learning results are based on the availability of rich and automated knowledge schemas which equip an experienced person with the necessary prior knowledge to process new information more easily. This research is exploratory and descriptive and focuses on the meaning of prior knowledge including the benefits and the challenges faced by the research participants, with the purpose of formulating certain conclusions and recommendations for the establishing of what should be considered concerning the initial knowledge in using discourse annotations to the corpus while teaching / learning foreign languages at advanced levels. The current study contributes to consideration of the teaching / learning practices

The object of the research is using discourse-annotated corpus for foreign language teaching / learning.

The aim of the research is to investigate the role of prior knowledge perception in applying corpus annotation while teaching a foreign language at more advanced levels.

The objectives of the research have been set as follows: to present a theoretical background of language teaching and corpora; to describe the lived experience of the research participants in a structural way; to disclose the meaning of prior knowledge that influences foreign language teaching / learning.

Theoretical Background

The development of large language databases known as corpora revealed the potential of language research using the corpus techniques focused on researching patterns of lexis, grammar, semantics, pragmatics and textual features. Many corpora are coded according to the parts of speech or analyzed for grammatical structure, or they are examined focusing on the pragmatic features.

Corpora development has enriched the knowledge concerning lexis, grammar, semantics, pragmatics, and textual features (Zanettin, 2011; Losey-Leon, 2015). Corpus linguistics is based on the theory that language varies according to the context related to space and time, which sustain the infinite potential for establishing new facts about language. If the same theoretical insight is applied to language teaching and learning practices, then the use of corpora in teaching and learning languages becomes very significant. Dictionaries and grammars do not have the capacity to fully describe the language. So, while applying corpora for teaching and learning languages, both teachers and learners can identify certain regularities and irregularities of the language based on the corpora data. According to Aston (2001) supported later by Koester (2010), another

benefit is that a corpus-based approach provides real data of real language used in certain contexts. The author also stresses the importance of the frequency information, which might be helpful while making teaching and learning choices. Scott and Tribble (2006) observe that at more advanced levels it is also important to acquire certain knowledge of genres and registers. The idea is well-supported by the learner corpus research (Granger, 2015) which reveals that language patterns used by relatively advanced language learners exemplify stylistic discrepancies rather than grammatical problems. The problematic areas for advanced language learners include coherence, cohesion, and textual rhetorical features. Cohesive devices and discourse markers attract researcher attention as the tools for ensuring textual and discourse management looking for answers as to what and how to teach at more advanced levels concerning the matters of textual coherence. Recent research suggestions have led to the idea of direct corpus use by language learners and teachers. The studies by Cobb and Boulton (2015) have shown that the application of such an innovative idea in teaching and learning has proven too effective and efficient. The authors have revealed that learners were better able to use linking adverbs by using corpus concordances rather than using bilingual dictionaries or grammars. The development of discourse-annotated corpora could present a substantial supplement in the surplus of teaching/learning materials, especially for more advanced learners in dealing with textual cohesion and coherence.

The four-component instructional design (4C/ID) model introduces four components essential for learning: learning tasks, supportive information, procedural information and part-task practice (van Merriënboer et al., 2003; van Merriënboer & Kirschner, 2013; van Merriënboer & Kester, 2014; van Merriënboer et al., 2019). In turn, each component is divided into subcomponents reflecting the systematic approach to learning. The authors of the 4C/ID model share the assumption of cognitive load theory (Sweller et al., 2011; Sweller et al., 2019) that human knowledge is stored in cognitive schemas, pointing out, that learning processes are related to the construction or reconstruction of schemas.

The second component of the 4C/ID model, which is titled "supportive information", deals with the necessity of sustaining information to enhance learning processes. It includes the subcomponent of prior knowledge. In fact, prior knowledge activation facilitates further development based on prior knowledge and enhances the process of the integration of new knowledge into the existing knowledge foundation. Many authors discuss the importance of prior knowledge and even suggest various approaches to activating it. For example, de Grave et al. (2001) suggest problem analysis for the activation of prior knowledge; Machiel-Bongaer et al. (1995) discuss mobilization; and Gurlitt et al. (2006) propose concept mapping.

Frerejean et al. (2019) analyze how teaching and learning outcomes are achieved and come to the conclusion that teaching and learning process establishes the availability of rich and automated knowledge schemas. The schemas organize and store knowledge, and a large number of new elements for a non-experienced person might be just one

element for an experienced person because the experienced person already has a cognitive schema available which incorporates other elements. As a result, new information is more easily processed by an experienced person with the prior knowledge than by a person without experience.

Methodology

We applied inductive qualitative methodological approach with applications of ground theory and performed semi-structured interviews (Ghiglione & Matalon, 2001; Smith et al., 2009) by asking the following questions:

- How do you perceive the process of applying corpus annotation while teaching a foreign language at more advanced levels? What is this process for you?
- What do you need in applying corpus annotation while teaching a foreign language at more advanced levels? What do you use?
- How do you perceive the role of prior knowledge in applying corpus annotation while teaching a foreign language at a more advanced level? How can it help you? How do you use it?
- How do you solve the challenges of using corpus annotation in teaching a foreign language at a more advanced level? What do you do?
- How could the teaching-learning process be improved by using corpus annotation while teaching a foreign language at more advanced levels?

The interview guides were devised, used, revised and newly devised as flexible tools containing open-ended questions that helped us to start each interview and achieve a natural and comfortable conversational flow throughout each interview, as Charmaz (2014) advises to "re-evaluate, revise, and add questions throughout the research process".

The current research is exploratory and descriptive, focusing on the prior knowledge benefits and challenges posed by the lack of prior knowledge observed by the research participants, and aiming to formulate certain conclusions and recommendations for corpus annotation application while teaching foreign language at more advanced levels. A qualitative approach was used in the study in order to conceptualize within the framework of inductive content analysis the main factors related to prior knowledge for corpus annotation application while teaching a foreign language at more advanced levels.

The case participants were defined by their experience related to the research question. The research participants have been chosen on the basis of theoretical sampling that implies purposive seeking of the participants, who can shed light on the researched phenomenon and its emergent categories, dimensions, and properties. Two groups of students majoring in translation studies were taught by applying corpus annotation to raise their awareness of text coherence: one group of first-year students who have just started their studies and lacked the initial knowledge of syntax and discourse relations

within the text and a group of third-year students who had covered a course on syntax and discourse relations within texts before. Their teachers were questioned about their observations during the process of teaching/learning by applying corpus annotation. The essential details of the research participants of the study are presented in Table 1.

Table 1
Research Participants

Students			
Research participant ID	**Year of studies**	**Level of studies**	**Employed**
S1	1	Bachelor's degree	no
S2	1	Bachelor's degree (Erasmus student)	no
S3	1	Bachelor's degree	no
S4	1	Bachelor's degree	no
S5	1	Bachelor's degree	no
S6	1	Bachelor's degree	no
S7	1	Bachelor's degree	no
S8	1	Bachelor's degree	yes
S9	1	Bachelor's degree	no
S10	1	Bachelor's degree	no
S11	1	Bachelor's degree	no
S12	1	Bachelor's degree	no
S13	3	Bachelor's degree (Erasmus student)	yes
S14	3	Bachelor's degree (Erasmus student)	no
S15	3	Bachelor's degree	no
S16	3	Bachelor's degree	no
S17	3	Bachelor's degree	yes
S18	3	Bachelor's degree	no
S19	3	Bachelor's degree	yes
S20	3	Bachelor's degree	no
S21	3	Bachelor's degree	no
S22	3	Bachelor's degree	no
Teachers			
Research participant	**Position**	**Work experience at university**	**Level of study programs taught**
T1	Lecturer	25	Bachelor's degree
T2	Lecturer	20	Bachelor's degree

A method of semi-structured interviews for obtaining the empirical data was chosen for the research. This method is acknowledged as one of the most effective and most commonly used methods of obtaining data in the qualitative research framework (Silverman, 2013), as it provides a direct way of collecting the information about the researched phenomenon and also enables us to reveal the experience of the research participants presented in their own words (Kvale, 1996; Silverman, 2013).

At the beginning of the interview, the aim of the research was presented to the interviewees. Also, their active participation, objectivity, and sincerity were encouraged during the interviews. Additionally, at the beginning of the interview, while introducing the topic of the research, we created an atmosphere of comfort and trust by recognizing that the research participants are the experts of their own experience concerning the role of prior knowledge in applying corpus annotation while teaching a foreign language at more advanced levels. We did so by explaining that the research object has not been fully researched yet and that the voices of the research participants are extremely valuable in the research.

As it is recommended (Kvale, 1996; Silverman, 2013), at the end of the interviews the research participants were asked if they wanted to add anything which had not been touched upon during the interview, and in some cases, they provided some additional details. The appreciation and gratefulness for the participation in the research and the time devoted to the interview were also expressed. The interviews, on average, lasted half an hour, totaling approximately 15 hours of recorded material.

Data Analysis

The inductive qualitative content analysis was carried out by means of NVivo, which is a piece of software efficient for organizing and managing the data; however, the computer software does not do the analysis of the data for the researcher. NVivo could suggest automatic subdivision of huge amounts of raw data, but manual coding needs to be performed looking into the context of automatically suggested subdivision and identifying the patterns, deciding what could be collated, and extracting the meaning from the data.

The interviews have been instantaneously analyzed by applying the constant comparative method. The data have undergone several coding stages which embrace initial open coding, axial coding, and selective coding. Initial open coding started with a very close inspection in the initial stage creating sub-categories to describe all the aspects of the content. Then it was followed by axial coding during which the list of sub-categories was grouped under categories and memos were written about the generation of categories (Corbin & Strauss, 2014). Selective coding involved the process of abstraction leading to the formulation of the core category and methodically relating it to the corresponding

categories. Selective coding extends into a stage with increasingly more abstract and conceptual processing, ultimately leading to the construction of the core categories.

Research Findings

The process of abstraction of core categories relied on a synthetic technique of making connections between subcategories, categories and core categories to construct a more comprehensive scheme. The resulting scheme after the full coding process is presented diagrammatically in Figure 1. The iterative categorization produced a set of two broad core categories: "personal dimension" and "organizational dimension", which are organized in a hierarchical order. Furthermore, the first broad core category "personal dimension" contains the categories "prior knowledge benefit" and "difficulties without prior knowledge", while the other broad core category of "organizational dimension" encompasses the categories of "students' need for improvement of the learning process to be catered to by the teaching staff" and "suggestions on the year of studies for introducing annotation". The set of these two broad core categories and associated concepts describes the conditions, experiences, and consequences associated with the role of prior knowledge in applying corpus annotation while teaching a foreign language at more advanced levels.

The research reveals that organizational factors (suggestions on the year of introduction of corpus annotation and students' need for improvement of the learning process to be catered to by the teaching staff) are important in the use of prior knowledge in applying corpus annotation while teaching a foreign language at more advanced levels. Student research participants were enthusiastic about applying corpus annotation while teaching and learning a foreign language at more advanced levels. They viewed corpus annotation as a useful tool that provided deeper understanding of textual coherence and cohesion. However, they admitted the importance of prior knowledge in the process of applying corpus annotation while teaching a foreign language at more advanced levels and revealed the difficulties faced while learning without prior knowledge. The categories established in the research revealed the meaning of the prior knowledge in a conceptual way.

The third study year students - research participants with the prior knowledge stressed the advantages of having and applying their prior knowledge. They stated that it made their learning process easier, upgraded their prior knowledge, helped them to annotate the discourse relations, and assisted them while comparing and contrasting discourse relations in different texts. Table 2 represents the category of prior knowledge benefit, which includes such sub-categories as easier learning process, helping the annotation process, and helping the process of comparing and contrasting discourse relations.

Table 2
Benefits of Prior Knowledge

Meaning unit	Sub-category	Category
I had some initial knowledge of syntax and Transformational Generative Grammar, so it was useful for performing annotation and doing corpus design, and the learning process was much easier for me. S19	Easiness of learning process	Benefits of prior knowledge
I understood that if students have some background knowledge of syntax and pragmatics, it is easier for them; otherwise, students had difficulties in coping with the annotation scheme. T2		
I had some knowledge of similar work on sentence analysis in Transformational Generative Grammar Course so I thought that the annotation process very interesting. S13	Assisting annotation process	
I knew how to analyze sentences, how they are connected, and how to divide them into arguments. I can say that my syntax knowledge assisted me a lot while doing the project on corpus building and annotation. S15		
I could easily identify the connectives as I already knew them from my previous studies; I knew conjunction, condition, contrast, comparison. S14		
My previous linguistic knowledge and skills were of great assistance while studying corpus building and annotation, and also for the analysis of coherence relations while contrasting and comparing coherence relations in the corpus. S22	Assisting the analysis of coherence relations	
I enhanced my initial knowledge about the structure of the sentence and I also learned a lot about sentential meanings related to cohesion and sentential semantics. S20	Enhancing initial knowledge	
The most encouraging part was revising and analyzing the classification of sentential connectives. It was a good revision as I had learnt this material before. S17		

The research participants observed that having prior knowledge of syntax and textual coherence ensures an easier learning process, as one of the research participants stated that (s)he had some initial knowledge of syntax and Transformational Generative Grammar, so it was useful for performing annotation and doing corpus design, and the learning process was much easier for him/her. This is also confirmed by the opinion of the teacher research participant who understood that if students have some background knowledge of syntax and pragmatics, it is easier for them; otherwise, students had difficulties in coping with the annotation scheme.

Another point observed by the research participants proves that prior knowledge helps the annotation process. The research participants pointed out that the syntax knowledge was of great assistance while working on the annotation project since they knew how to divide sentences into arguments and how they were connected.

It was also revealed that prior knowledge helped during the process of comparing and contrasting coherence relations, as one of the research participants observed that his/her previous linguistic knowledge and skills were of great assistance while studying corpus building and annotation, and also for the analysis of coherence relations while contrasting and comparing coherence relations in the corpus.

The observations expressed by the research participants resonate with the ideas discussed by van Merrienboer et al. (2003) and van Merriënboer et al. (2019) that new information is more easily processed by an experienced person with the prior knowledge than by a person without experience. Another category, difficulty without prior knowledge, reveals that the absence of prior knowledge, on the contrary, made the learning process more difficult, which also supports the idea expressed above by van Merriënboer et al. (2019). The category includes such sub-categories as difficulty in understanding, more learning and novelty of the information (Table 3).

Table 3

Difficulties Because of Lack of Prior Knowledge

Meaning unit	Sub-category	Category
I had difficulties to understand and identify certain parts of speech in English because the information in the annotation manuals was too complex, it was not simple so the annotation manuals were really difficult for me to follow and apply. S2	Understanding difficulties	Difficulties because of lack of prior knowledge
We had a lot of new information to understand and learn so at the beginning it was really difficult to figure out how to perform annotation. S6	Extensive learning	
I needed to learn a lot of new information about various types of sentential relations and it was difficult, because I did not have any knowledge about sentential relations before. S8		
The whole project and especially text annotation was absolutely new for me and it was not so easy to master. S11	Information novelty	
I never knew about corpus analysis before so it was difficult to understand all the complexity of terms related to corpus annotation and even more complex to understand relations between the textual segments. S11		
We never had such a project before so it was difficult to perform it. S12		

Research participants point out that without prior knowledge they faced certain difficulties. First, they experienced difficulties in understanding the annotation process and had to undergo extensive learning. As one of the research participants observed, they had a lot of new information to understand and learn so at the beginning it was really difficult to figure out how to perform annotation.

The novelty of the information also caused difficulty; for example, that the whole project and especially text annotation were absolutely new and it was not so easy to master.

Student research participants also revealed their need for the improvement of the learning process to be catered to by the teaching staff, which involves one more category presented in Table 4.

Table 4
Needs of Students to Improve Teaching/Learning Process

Meaning unit	Sub-category	Category
I think we need much more time for the preparation of working on the annotation project before starting it. S9	Preparation needs	Needs of students to improve teaching/learning process
I think annotation makes teaching/learning process better and more effective, but we need some time for preparation and studying all the different sentential relations. S3		
I think at the beginning of the annotation project we need introductory sessions on the parts of speech in the English language, types of connectives and their functions in the text. S6	Need for meta-linguistic knowledge	
I think we need more introduction lectures as it could be useful for us to study detailed explanations of sentential relations and to understand how they differ from each other, how to identify and analyze them in the text. S8		
I would like to study more theoretical knowledge before starting our annotation project because it could help to better annotate or perform new practice. S1		
I would like more knowledge of basic syntax, semantics, and sentential relations. I think I need this basic knowledge in order to perform the project on corpus analysis, because language is a complex study area. S7		
I think it would be good to get a detailed explanation on annotation process and annotation related labels. S5	Need for explanation	
I would like everything to be extensively explained about parts of speech, about connectives, and sentential relations. S9		
I think it would be a good idea to start annotation with simpler tasks like smaller paragraphs and then move further to bigger parts texts, and only then analyze the whole text. S12	Simpler tasks	
I would like to annotate only a few paragraphs and try to mark every relation and analyze it and then it could be more efficient to annotate bigger texts. S11		

The research participants identified their need for preparation which would make their learning easier; for example, that they think annotation makes teaching/learning process better and more effective, but they need some time for preparation and studying all the different sentential relations. They also emphasized their need for introduction of

meta-linguistic knowledge and need for explanation which could enhance their learning. The research participants expressed their preference for simpler tasks as they believe it could make their learning more efficient moving from simpler tasks to more complex ones and that they would like to annotate only a few paragraphs and try to mark every relation and analyze it and then it could be more efficient to annotate bigger texts. The ideas pointed out by the research participants resonate with the observations by the authors van Merrienboer et al. (2003) and van Merriënboer et al. (2019) on the necessity of supportive information, procedural information and part-task practice. The category of research participant suggestions on the study year for the introduction of annotation also reveals that students need prior knowledge and preparation (Table 5).

Table 5
Suggested Study Year for Teaching Corpus Annotation

Meaning unit	Sub-category	Category
I think a course on corpus design and annotation should be offered in the second year of studies rather than in the first year of studies It is difficult to fully grasp the material because in the first year of studies we lack essential knowledge on syntax and pragmatics. I suggest moving the course towards the middle of the studies; let's say, second half of the second year. S22	Suggestion on the second year of studies	Suggested study year for teaching corpus annotation
I think corpus building and annotation should not be taught to the first-year students. We are just at the beginning of our studies so do not have a lot of background knowledge. In my opinion, the course could be good it is good for students who have already covered some introductory courses on linguistics. S15		
I feel that corpus design and annotation should be taught in the third semester because, now, in the second semester it was a bit too complicated for some students. S11		
I mean annotation could be taught to third-year students because then they have a more extensive pool of background knowledge necessary for corpus design and annotation. S17	Suggestion on the third year of studies	
It would be a good solution to do a course on corpus annotation in later study years, for example, the third year of studies. S14		
The introduction of the course on corpus design depends on the level of the students. However, I feel it is too early for bachelor's studies. T1	Suggestion for master's studies	
A course on corpus design and annotation is a useful course for students of linguistics in the last years of bachelor's studies But it should be better for master studies, or could be offered as a specialized master's program. T2		

The category reveals that the research participants of who were first-year students suggested introducing annotation and corpora application later, either in the second or the third year of studies. Their reflections are followed by the suggestion by their teachers of the possibility of introducing annotation and corpora application even to master's students. The suggestions reveal that van Merriënboer et al. (2019) might be right in their theoretical observations that human expertise is the availability of rich and automated knowledge schemas which help an experienced person with the prior knowledge to process new information.

Discussion and Conclusions

The current research reveals that initial linguistic knowledge is essential while teaching/learning a foreign language at more advanced levels, while applying corpus annotation and comparing and contrasting coherence relations. The research participants observed that prior knowledge of syntax and textual coherence ensures an easier learning process; it also helps the annotation process and the process of comparing and contrasting discourse relations. What is more, the lack of prior knowledge is admitted to lead to certain difficulties in the learning process; for instance, according to the student research participants, the novelty of the information causes difficulty in understanding for them and also leads to learning additional material to study the new information. The observations of the research participants resonate with the ideas expressed by van Merriënboer et al. (2019) that new information is more easily processed by an experienced person with the prior knowledge than by a person without experience. In addition, the research participants reflected on the year of introducing annotation and corpora application and expressed their suggestions for later introduction of corpus annotation, either in the second or the third year of studies, which is also related to the students' concerns about acquiring the prior knowledge.

The research participants also identified certain needs for enhancing the teaching/learning process. They identified their need for preparation, need for introduction of meta-linguistic knowledge and need for explanation; also, the research participants expressed their preference for simpler tasks as they believe simpler tasks could make their learning more efficient. The needs identified by the research participants resonate with the observations by the authors van Merrienboer et al. (2003); van Merriënboer & Kester (2014), and van Merriënboer et al. (2019) on the necessity of supportive information, procedural information and part-task practice. The identified needs also lead to the teacher's role as an activator, which was well-presented and analyzed by Hattie (2012) while discussing visible learning. The research participants' needs resonate with Hattie's (2012) idea that teachers focus on imparting new knowledge and then monitor the process how the students get fluent in using the new knowledge. The process of

imparting new knowledge seems to be closely related to the preparation and introduction of meta-linguistic knowledge and providing further explanation, and then, the stage of starting with simpler tasks and progressing towards the more complex ones this way building students' fluency in using the newly acquired knowledge.

The results of the current research lead us to suggest further research beyond the meaning of the initial knowledge in using a discourse-annotated corpus for teaching/learning a foreign language at more advanced levels. It would be important to program a follow-up course on corpora development and annotation, following the teaching methodology proposed by the 4C-ID model. It would be important to devise learning tasks that follow the principles advocated by the model, namely, to design classes of tasks that respond to the difficulties experienced by the participants in this study and following the principles of teaching complex learning.

References

Aston, G. (2001). Learning with corpora: An overview. In G. Aston (Ed.), *Learning with corpora* (pp. 7–45). Athelstan.

Charmaz, K. (2014). *Constructing grounded theory* (2nd ed). Sage. http://www.sxf.uevora.pt/wp-content/uploads/2013/03/Charmaz_2006.pdf

Cobb, Th., & Boulton, A. (2015). Classroom Applications of Corpus Analysis. In D. Biber-Reppen (Eds.) *The Cambridge handbook of English corpus linguistics* (pp. 478–97). Cambridge University Press. https://dx.doi.org/10.1007/9781139764377.027

Corbin, J., & Strauss, A. (2014). *Basics of qualitative research: Techniques and procedures for developing grounded theory.* Sage publications. https://dx.doi.org/10.4135/9781452230153

De Grave, W. S., Schmidt, H. G., & Boshuizen, H. P. (2001). Effects of problem-based discussion on studying a subsequent text: A randomized trial among first year medical students. *Instructional Science*, *29*(1), 33–44. https://doi.org/10.1023/A:1026571615672

Frerejean, J., van Merriënboer, J. J., Kirschner, P. A., Roex, A., Aertgeerts, B., & Marcellis, M. (2019). Designing instruction for complex learning: 4C/ID in higher education. *European Journal of Education*, *54*(4), 513–524. https://doi.org/10.1111/ejed.12363

Ghiglione, R., & Matalon, B. (2001). *O inquérito: Teoria e prática* [Inquisitive: Theory and Practice] (4th ed.). Celta Editora. http://books.google.pt/books?id=_1v-SAAACAAJ

Granger, S. (2015). Contrastive interlanguage analysis: A reappraisal. *International Journal of Learner Corpus Research*, *1*(1), 7–24. https://doi.org/10.1075/ijlcr.1.1.01gra

Gurlitt, J., Renkl, A., Motes, M. A., & Hauser, S. (2006). How can we use concept maps for prior knowledge activation-different mapping-tasks lead to different cognitive processes. In S. A. Barab, K. E. Hay, & D. T. Hickey (Eds.), *Proceedings of the 7th International Conference of the Learning Sciences* (pp. 217–21). Lawrence Erlbaum. https://www.researchgate.net/publication/336945328_How_can_we_use_concept_maps_for_prior_knowledge_activation_different_mapping-tasks_lead_to_different_cognitive_processes

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning.* Routledge. https://hozir.org/pars_docs/refs/26/25322/25322.pdf

Koester, A. (2010). Building small specialised corpora. In A. O'Keeffe, & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (1st ed.). (pp. 66–79). Routledge. http://doi.org/10.4324/9780203856949

Kvale, S. (1996). An introduction to qualitative research interviewing. Sage. https://doi.org/10.1016/S1098-2140(99)80208-2

Losey-León, M. A. (2015). Corpus design and compilation process for the preparation of a bilingual glossary (English-Spanish) in the logistics and maritime transport field: LogisTRANS. *Procedia-Social and Behavioral Sciences*, *173*, 293–299. https://doi.org/10.1016/j.sbspro.2015.02.068

Machiels-Bongaerts, M., Schmidt, H. G., & Boshuizen, H. P. (1995). The effect of prior knowledge activation on text recall: An investigation of two conflicting hypotheses. *British Journal of Educational Psychology*, *65*(4), 409–423. https://doi.org/10.1111/j.2044-8279.1995.tb01162.x

Melo, M. (2018). The 4C/ID-Model in physics education: Instructional design of a digital learning environment to teach electrical circuits. *International Journal of Instruction*, *11*(1), 103–122. https://doi.org/10.12973/iji.2018.1118a

Melo, M., & Miranda, G. L. (2015). Learning electrical circuits: The effects of the 4C-ID instructional approach in the acquisition and transfer of knowledge. *Journal of Information Technology Education*, *14*, 313–37. https://www.researchgate.net/publication/280553730_Learning_Electrical_Circuits_The_Effects_of_the_4C-ID_Instructional_Approach_in_the_Acquisition_and_Transfer_of_Knowledge

Scott, M., & Tribble, C. (2006). *Textual patterns: key words and corpus analysis in language teaching*. John Benjamins. https://doi.org/10.1075/scl.22

Silverman, D. (2013). *Doing qualitative research: A practical handbook*. Sage. https://www.researchgate.net/publication/279187183_Doing_Qualitative_Research_A_Practical

Smith, J. A., Flower, P., & Larkin, M. (2009). *Interpretative phenological analysis. Theory, method and research*. Sage. https://www.researchgate.net/publication/221670349_Interpretative_Phenomenological_Analysis_Theory_Method_and_Research

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, *31*, 261–292. https://doi.org/10.1007/s10648-019-09465-

Sweller, J., Ayres, P., & Kalyuga, S. (2011). Measuring cognitive load. In J. Sweller, P. Ayres, & S. Kalyuga (Eds.), *Cognitive load theory* (pp. 71–85). Springer. http://dx.doi.org/10.1007/978-1-4419-8126-4_7

Van Merriënboer, J. J., & Kester, L. (2014). The four-component instructional design model: Multimedia principles in environments for complex learning. In R. E. Mayer (Eds.) *The Cambridge handbook of multimedia learning* (2nd ed.). (pp. 104–48). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.007

Van Merriënboer, J. J., & Kirschner, P. A. (2013). *Ten steps to complex learning* (2nd. ed.). Routledge. https://web.mit.edu/xtalks/TenStepsToComplexLearning-Kirschner-VanMerrienboer.pdf

Van Merriënboer, J. J., Kirschner, P. A., & Kester, L. (2003). Taking the load off a learner's mind: Instructional design for complex learning. *Educational Psychologist*, *38*(1), 5–13. https://doi.org/10.1207/S15326985EP3801_2

Zanettin, F. (2011). Translation and corpus design. *Synap*s, *26*, 14–23. https://www.researchgate.net/publication/259638177_Translation_and_Corpus_Design

Diskurso anotuotas tekstynas užsienio kalbų mokymui(si): išankstinių žinių reikšmė

Giedrė Valūnaitė Oleškevičienė[1], Guilhermina Lobato Miranda[2], Liudmila Mockienė[3], Dalia Gulbinskienė[4]

[1] Mykolo Romerio universitetas, Ateities g. 20, 08303 Vilnius, Lietuva, gvalunaite@mruni.eu
[2] Lisabonos universitetas, Universiteto alėja, PRT-1649-013 Lisabona, Portugalija, gmiranda@ie.ulisboa.pt
[3] Mykolo Romerio universitetas, Ateities g. 20, 08303 Vilnius, Lietuva, liudmila@mruni.eu
[4] Vilniaus Gedimino technikos universitetas, Saulėtekio al. 11, 10223 Vilnius, Lietuva dalia.gulbinskiene@vilniustech.lt

Santrauka

Tekstynų kūrimas davė impulsą tyrimų sričiai, susijusiai su jų vartojimu kalbų mokyme ir mokymesi. Tekstynų taikymas mokant ir mokantis kalbų apima tam tikrų jų kūrimo ir anotavimo įrankių įsisavinimą ir besimokančiųjų kalbinių ir metalingvistinių gebėjimų plėtojimą. Mokymo ir mokymosi procesui būdingas išankstinių žinių schemų panaudojimas. Tyrimu siekiama ištirti pradinių žinių reikšmę mokant ir mokantis užsienio kalbos aukštesniu lygiu, kuriant ir naudojant diskurso anotuotus tekstynus. Pusiau struktūruotais interviu paremtas kokybinis indukcinis tyrimas atskleidžia dvi esmines dimensijas: organizacinę ir asmeninę, kurios struktūriškai reprezentuoja tyrimo dalyvių patirtį. Tyrimo dalyviai stebi išankstinių žinių naudą ir iššūkius, su kuriais jie susiduria dėl išankstinių žinių stokos, ir siūlo įvesti tekstyno anotaciją antraisiais ar trečiaisiais studijų metais. Tyrimo dalyvių įžvalgos rezonuojasi su teoriniais pagrindais, susijusiais su pradinėmis žiniomis mokant ir mokantis užsienio kalbų, ir sykiu pateikiami pasiūlymai dėl būsimų šios srities tyrimų ateityje.

Esminiai žodžiai: *diskurso anotuotas tekstynas, kalbų mokymas / mokymasis, išankstinės žinios.*