

18th PCSF 2018
Professional Culture of the Specialist of the Future
ON PROSPECTS OF USING THE DDL APPROACH IN GSP
COURSE

Marina Kogan (a)*, Anna Kolotaeva (a), Anna Yaroshevich (a), Victor Zakharov (b)

*Corresponding author

(a) Peter the Great St. Petersburg Polytechnic University (SPbPU), Polytechnicheskaya 29, Saint Petersburg, 195251
Russia, m_kogan@inbox.ru

(b) Peter the Great St. Petersburg Polytechnic University (SPbPU), Polytechnicheskaya 29, Saint Petersburg, 195251
Russia, v.zakharov@spbu.ru

Abstract

This study focuses on prospects of using data-driven learning (DDL) methodology in teaching German for specific purposes (GSP) at Peter the Great Petersburg Polytechnic University. Unlike General English and English for specific purposes that are successfully taught with DDL worldwide, DDL does not seem to have paved the way to the GSP classroom yet. The paper analyses the reasons for this. One of the obstacles is that large general corpora do not contain field specific terminology, whereas specialized German corpora are not available. German scientific terminology is featured with high portion of compounds. Preliminary study of a small corpus of scientific papers built by the authors for this research proved this fact which makes it relevant to academic needs of learners of German who make a lot of mistakes in using and translating German compounds. The article presents lists of the most frequent nouns and some of the most frequent compound nouns that contain the nouns from the first list as stems generated using the AntConc freeware corpus analysis toolkit and a set of corpus-based teaching materials to help students to handle this lexical category. Further research should confirm the efficiency of DDL in teaching German compounds in the GSP classroom and effectiveness of the described way of integrating DDL in the GSP classroom: determining problem areas (gaps) in students' knowledge -> finding these structures in the corpus -> development of corpus-based teaching materials, and using them in the GSP course.

© 2018 Published by Future Academy www.FutureAcademy.org.UK

Keywords: AntConc, corpus analysis, corpus-based teaching materials, Data-driven learning (DDL), German for specific purposes (GSP), German compound nouns.



1. Introduction

In Russia the corpus-based approach to linguistics and language education dates back to the mid 1980s, when academician Andrey Petrovich Ershov initiated the development of Computer Fund of the Russian Language which resulted in creation of the Russian National Corpus (RNC), (available at <http://ruscorpora.ru/> since 2004). The influence of corpus linguistics on different branches of linguistics has become increasingly noticeable since then. This is especially true for development of large Russian corpora such as RNC and *One Speaker's Day* (Asinovsky, et al., 2009), lexicographic and lexical studies of the Russian language, sociolinguistics, translation studies (Beliaeva & Chernyavskaya, 2016), discourse analysis (Chernyavskaya, 2017, 2018), and some others. Foreign language pedagogy is one of the few branches in which prospects of using corpus-based approach remained generally underestimated by Russian language teachers unlike their colleagues abroad who have integrated corpus linguistics approaches into foreign language classroom since the late 1980s.

The term “data-driven learning” (DDL) was first introduced in 1990 by Tim Johns (Boulton, 2017, p.487). The word data usually refers to different text corpora. The great majority of publications on the use of corpora in language teaching and learning is in English and on the material of different English language corpora. In the XXI c. these studies for the English language are conducted at full tilt and their results are widely used in practice. Experts point out that these range from corpus linguistics itself and the development of corpora, tools and methodologies that are widely used in DDL, including work on language for specific purposes, frequency lists and learner corpora, error-correction and contrastive analysis, to corpus uses in syllabus design, materials and resources, testing and translation, as well as fields from lexis to acquisition studies and more general language learning theory

The comprehensive survey article by Boulton and Cobb (2017) gives the idea about the number and evolution of scientific publications in the DDL area. The authors mention that for the meta-analysis they selected 205 studies in English which contained an empirical evaluation of DDL, with the first research article published in 1989. The steady growth of the number of publications on the DDL can be seen for the past decades (from 1989 to present). Boulton and Cobb did not include in the meta-analysis ‘excellent qualitative DDL studies’ from the same period, the full extent of the fugitive literature, nor DDL studies published in languages other than English on purpose (Boulton & Cobb, 2017, p. 34). The corpus research contributed a lot into language pedagogy through designing syllabus, production of teaching and testing materials. The most acknowledged examples of this contribution are Longman Grammar of Spoken and Written English (LGSWE) (Biber et al. 1999), dictionaries and coursebooks produced as a result of the COBUILD project led by Sinclair (1987), Natural Grammar by Thornbury (2004) and others. Boulton puts in line with them a large bank of worksheets, an underestimated resource in his opinion, which can be found in numerous publications on DDL (Boulton, 2010).

1.1. Ways of DDL implementation

Discussing prospects of using the DDL approach in German for Specific Purposes (GSP) course it seems logical to briefly review the most common ways of implementing it in teaching English described in literature.

The data for this section which focuses on ways of DDL implementation in the classroom, types of corpora used, and the language skills and language forms targeted mainly came from the two sources: a book chapter (Cobb & Boulton, 2015) and a survey article (Boulton & Cobb, 2017).

In implementing DDL researchers and practitioners used varied corpora (large general corpora and small tailor-made ones) considering them as learning aids or reference resources and practiced different ways of learners' interaction with corpus materials. The *ways of interaction* with corpora are marked with a different degree of students' involvement. They include

- the student directly manipulating the available corpora and even mastering a special software (concordancer programs) which build a corpus from a collection of texts (so called *hands-on* approach);
- working with corpus-based hand-outs developed by a teacher (so called *hands-off* approach);
- and the combination of the two depending on pedagogical purposes.

75% of research papers selected for the meta-analysis used the *hands-on* approach. Contrary to some claims, the DDL approach seems to be most effective when using a concordancer hands-on rather than through printed materials. Tailor-made local corpora may be more effective than large public corpora (Boulton & Cobb, p.37-38).

DDL approach has been integrated into *General English*, *English for Specific Purposes* and *English for Academic Purposes* courses. As for development of *language skills* (listening, speaking, reading, writing, translation) using DDL approach, **writing** clearly dominates, though with mixed results, with translation following. The other skills remain largely underresearched (Boulton & Cobb, 2017, p.32). Among *language aspects* of language (vocabulary, lexicogrammar, grammar, discourse) more attention is given to lexicogrammar, i.e. the immediate surroundings of the target word (emphasis was less on meaning than on how the target word fits into its surroundings), alongside with extension of a vocabulary including phrasal verbs and other multiword units, and grammar (e.g., tenses or articles). Summing up their findings Cobb and Boulton note that the language objectives generally tend towards the level of vocabulary or lexico-grammar (including clusters and collocations, i.e. word usage in context), 'but there are attempts to use corpora in learning grammar and syntax, and even occasionally in phonetics or semantics' (Cobb & Boulton, 2015, p.483) and stress that DDL is 'a strong methodology for learning language per se, including lexicogrammar' (Boulton & Cobb, 2017, p. 34). Decisive conclusions for or against DDL approaches use in grammar or discourse are impossible because of the insufficient number of studies (Boulton & Cobb, 2017, p.38).

In Russia, researchers conduct DDL studies using small tailor-made corpora, built for the certain purpose\ task either by researchers (Pavlovskaya & Gorina, 2015) or students (Almazova & Kogan, 2013, 2014), big corpora (BNC, COCA) (Tarnaeva & Osipova, 2016, Strelkova, 2011, Sysoev, 2010), and parallel corpora (Sosnina, 2006, Sysoev & Kokoreva, 2013). In these studies, researchers focus on

- ways of making learners' language more idiomatic (Pavlovskaya & Gorina, 2015, Tarnaeva & Osipova, 2016);
- enhancing vocabulary for professional purposes (Almazova & Kogan, 2013, 2014, Sysoev & Kokoreva, 2013);
- specific grammar issues (Strelkova, 2011, Sysoev, 2010, Almazova & Kogan, 2013).

The Boulton and Cobb's meta-analysis concludes that the DDL **effective and efficient** in different learning environments with different categories of learners in pursuing different language goals.

1.2. Obstacles to DDL implementation

Boulton and Cobb (2017) note that there is a number of factors impeding wide use of DDL in the foreign language classroom. They are well-known in teaching community but rarely discussed in scientific papers devoted to this methodology. The following are typical obstacles:

- low ICT professional competence or lack of DDL training of language teachers (Leńko-Szymańska, 2017);
- chopped-off concordance lines which look very different from other text types and make it more difficult to gain meaning traditionally conceived via linear reading;
- language competence of most learners is well below the level of comfort reading/studying the authentic native language most corpora are composed of;
- DDL work requires substantial training, and the processes are time consuming when learners could simply be told or use pedagogically derived resources such as dictionaries (Boulton & Cobb, 2017, p.4);

and more related to this paper

- the lack of available research on using specialized corpora other than English (Boulton, 2010a; Yoon, 2016).
- the lack of corpora built with learners' needs in mind (e.g. oriented on an appropriate language level of learners).

Cobb and Boulton admit that despite the DDL approach being promising and efficient "traditional corpus consultation is in some ways a relatively marginal activity on a global scale" (Cobb & Boulton, 2015, p.496), which is especially true for languages other than English.

1.3. Using corpora in teaching German

In reviewing publications on using DDL in teaching German we started with an annotated list of references to all 205 studies which Boulton and Cobb included in their meta-analysis (Boulton & Cobb, 2017, p. 46) (available in Supporting Information Folder at the Publisher's website <https://onlinelibrary.wiley.com/doi/pdf/10.1111/lang.12224>). We discovered only 9 publications dealing with using corpora in teaching German chronologically ranged from 2001 to 2013. 4 of them were written by Vyatkina alone or with her co-author. Then we conducted a search for works by this author which resulted in two recent papers by Vyatkina both published in 2016 (Vyatkina, 2016a; Vyatkina 2016b). The content analysis of *German as a Foreign Language* journal on the journal's web-site (<http://www.gfl-journal.de/previous.php>) for the 5 year period (2013 – 2018) did not reveal any papers related to the topic in question. The on-line search with Google search engine with keywords such as *German corpus/corpora, data-driven, DDL, learning/teaching German* did not bring anything important for the problem. Most publications were in English and did not contain any teaching materials or recommendations on using DDL in German language classroom.

Below a brief summary based on papers available in Supporting Information Folder at the Publisher's website (<https://onlinelibrary.wiley.com/doi/pdf/10.1111/lang.12224>) demonstrates learners'

interaction with corpora, corpora types and language focus and gives a typical picture of using DDL in teaching German.

- American intermediate college level language learners worked with a series of the instruction materials developed by the teacher from a learner corpus. Language focus: German modal particles, *da* compounds pragmatic competence.
- 2-nd University students from different majors, beginners use .DWDS corpus to study German verb-preposition collocations.
- American intermediate college level language learners with German as their major or minor study German verb-preposition collocations using DWDS corpus.
- University students, beginners work with The German-English parallel INTERSECT corpus, hands-on, to study German modal particles and unknown words
- 2-nd, 3-d year university students, majoring in German used LeaP corpus, a phonetically annotated, fully text-to-tone aligned learner corpus of English and German to study phonetic properties of non-native speech.
- Junior high school learners used three small specialized corpora containing texts representing each one sub-genre of the "tourist information" genre where collocations, German adjective inflection are presented.
- High school learners focused on lexical and grammatical points from the coursebook using hands-on corpus activities. The author argues for a move from 'data-driven learning' to needs-driven corpora, corpus activities and corpus methodologies.
- Undergraduate students majoring in interpreting worked with the BACKBONE corpus, a multilingual pedagogic corpus with audio-visual content, investigating subject-specific terminology.

We can see that

- 1) none of the authors dealt with master University students from engineering majors;
- 2) for German for specific purposes small specialized corpora were built;
- 3) language focus was chosen because of either objective difficulties of the German language (e.g., verb-preposition collocations, German adjective inflection) or learners' needs to master unknown words or domain specific terminology.

As for integrating corpus-based techniques into the teaching of academic German vocabulary Jaworska concludes that it is 'still in its infancy' (Jaworska, 2015, p. 188). We assume that this is also true for teaching German in Russian institutions of higher education. In the following sections we will consider the syllabus of GSP course at SPbPU and justify steps we undertook to start using DDL in the GSP classroom.

1.4. German corpora for GSP course

There are quite a few German corpora and lexicographical resources available on-line. The most important among them is *Deutsche Reference Korpus* – DeReKo. It can be queried and analyzed free of charge by registered users via the system *COSMAS II* (<http://www1.ids-mannheim.de/kl/projekte/korpora>). A corpus of digital dictionary of the German language – *Das Digitale Wörterbuch Der Deutschen Sprache (DWDS)*, (<https://www.dwds.de>) – is also a very popular resource. It

provides the possibility of researching various uses of a word in well documented text corpora. The system is flexible with regularly updated information. The integrated portal *CorpusEye* (<http://corp.hum.sdu.dk/cqp.de.html>) provides access to various German corpora simultaneously. This list can be extended (see, for example, <https://www.uni-giessen.de/fbz/zmi/das-zmi/digitalhumanities/corpora>). But their analysis reveals that they are more suitable for theoretical linguists studying the German language in general or its particular aspects.

To our knowledge, there are no appropriate corpora resources, which would allow engineering students to benefit from using a DDL approach in developing writing skills or acquiring subject specific terminology in German. S. Schroth- Wiechert from Leibniz University of Hannover (LUH) decided to bridge this gap through the development of a specialized corpus as a resource for learners of academic and technical German. She concluded that the number of such corpora for learners of academic German is very small, are not easily accessible and the examples they contain are of low relevance to Master's degree students of a specific engineering field (e.g, Turbomachinery and Fluid Dynamics or Civil Engineering). She began to realize her idea as a part of the Strategic Partnership Program between Peter the Great St. Petersburg Polytechnic University and the Leibniz University of Hannover. As a result of the project the *Deutsch, English and Russkii* (DEaR) Corpus of German, English and Russian languages for engineering is to be developed. It will be an annotated corpus with different search capabilities available on-line for students and lecturers of the participant universities. The DEaR corpus is composed primarily from electronically published PhD and Master's degree dissertations written by engineering Master's and Post-graduate students, native speakers of either German, English or Russian and other technical texts.

Not being able to work with the DEaR corpus while it is under development, we built a small corpus of free-accessible research papers such as dissertations, scientific reports, articles from scientific journals and popular science magazines written in German (further Sample Corpus) and analysed it using *AntConc*, a freeware corpus analysis toolkit for concordancing and text analysis.

2. Problem Statement

Current Master's degree engineering programs in universities of Russia do not provide sufficient opportunities for foreign language learning. Thus, at Peter the Great St. Petersburg Polytechnic University (SPbPU) Master's degree engineering programs comprise a foreign language course of two academic hours a week within one semester. The educational process is to be structured and optimized in order to provide the opportunity for students to master both general communicative skills and language in their professional domains which involves the use of components of certain terminological systems in a foreign language, in order to understand professional literature and to participate in multilingual professional communication. The lack of teaching hours makes the crucial importance of independent learning optimization apparent.

In accordance with global trends, the number of SPbPU Master's degree students choosing German as their second foreign language declines annually. In view of the fact that in 2017-2018 the number of students who were studying German accounted for approximately 1% of students who learn English

within their Master's program, a group of German learners may include students of different major, which complicates the process of language teaching for professional purposes significantly.

Currently printed teaching materials such as coursebooks are complemented with online resources, e.g. authentic materials of the Deutsche Welle website. Articles about achievements in science and technology can be found in such sections as *Deutschland*, *Wissen und Umwelt*, *Wirtschaft*, and *Projekt Zukunft global Ideas*. Engineering students at SPbPU are mainly interested in the latest news in mechanical engineering, the motor industry, and nuclear engineering available on the DW website. These popular science materials contain technical details and professional vocabulary. The class work includes usage and revision of professional vocabulary and special collocations, information exchange on research issues, discussions and PowerPoint presentations. Mastering pre-fabricated structures is essential, as such structures are common in both spoken and written communication. In our opinion, DDL methodology could be used to enhance the present curriculum, which is based on the communicative approach. The experts' opinion about flexibility and variety of ways of implementing DDL, their conclusion that there is no single "right" way to use corpora in the classroom (Cobb & Boulton, 2015, p.497) encouraged us to undertake research to understand how to integrate the DDL approach in the GSP course to benefit engineering students of SPbPU.

3. Research Questions

The following research questions were explored:

1. What typical mistakes do Russian students of German do in their essay writing in GSP course?
2. Does the Sample Corpus contain examples which can help learners to correct their mistakes?

4. Purpose of the Study

The purpose of the study was threefold:

- to detect learners' problematic areas in GSP;
- to determine if the Sample Corpus relevant for the purpose of correcting typical students mistakes and if so, to design a pedagogical experiment on integrating the DDL approach into the GSP syllabus.

5. Research Methods

5.1. Detecting problematic language areas

Following Boulton (2010a), we expected to detect language problems from learners' own writings. We invited volunteers from the Tandem project (Stratonova, 2016) to write a short essay of up to 200 words answering the question: "If I had a chance to study at a German university, which subjects would I select and why?". Five students responded. Their essays were carefully checked and discussed with a native German speaker. Mistakes were numerous: with prepositions, articles, word choice, grammar mistakes related to verbs and link words, word order in simple and complex sentences etc.

One problematic area, compound nouns, was chosen to deal with further. Russian learners of German used the genitive phrases common for Russian scientific discourse instead of compound nouns

common for German scientific discourse, possibly, as a result of interlingual interference. Erroneous use of articles and prepositions in some cases was related with compound nouns.

5.2. Some remarks on German compound nouns

In German this type of word formation results in words of complex structure, combining at least two stems. Two-stem compound nouns (e.g., *Streckenabschnitt*) prevail though three-stem (*Vertikalachsenabschnitt*) and occasionally four-stem compound nouns (*Schmierspalthöhenberechnung*) are also used. Compound words represent a type of word formation and are a kind of linguistic economy.¹ Compound nouns can be classified according to morphological, syntactic and semantic aspects, the joining element, the number of stems, and some other factors.

Morphological classification is based on the part of speech of the last stem, so called ‘the basic word’

- substantival compounds (*Semesterarbeit, Achsenabschnitt*);
- -compound adjective (*feinfühlig, stressbeständig*);
- -verbal compound structure (*teilnehmen, umbauen*);
- -compound adverbs (*weiterhin*);
- compounds with numerals (*hundertfünf*).

Within the same part of speech they are distinguished (further classified) depending on the relation between compound stems (*Determinativkomposita* or *Kopulativkomposita*).

5.3. Corpus manager

For the corpus analysis we selected one of the most efficient concordancers AntConc, the software developed by L. Anthony, professor of Waseda University (Japan), available at his website². The arguments for the choice of this software are as follows:

- This is free software distributed under Creative Commons License;
- Compatible with different platforms;
- Has a user-friendly interface;
- Has a multitude of functions of automatic processing of texts in different codes and formats;
- Allows analysis with eliminating words from stop-lists, or focusing on lemma lists or negative keywords lists etc.

6. Findings

6.1. Collection and analysis of compounds in the Sample corpus

Using *AntConc* we produced a frequency list of our corpus vocabulary with function *Word List*. Then we selected the most frequent nouns and went on to search compounds containing these words as a stems.

The list of most frequent nouns includes core general vocabulary and lexis from scientific texts of different disciplines. The examples of the core general vocabulary are *Daten* (data), *Abbildung* (image), *Vergleich* (comparison), *Untersuchung* (research), *Einfluss* (influence), *Anzahl* (quantity), *Bereich* (field),

¹ https://is.muni.cz/do/rect/el/estud/pedf/ps12/wortbild/web/pages/02-1_komposition.html

² Laurence Anthony's AntConc, a freeware corpus analysis toolkit for concordancing and text analysis <<http://www.laurenceanthony.net/software/antconc/>>.

Ergebnisse (results), *Tabelle* (table). The frequent words from general scientific vocabulary are *Simulation* (simulation), *Berechnung* (calculation), *Verfahren* (process), *Temperatur* (temperature), *Gleichung* (equation), *Technik* (technology).

6.2. Procedure of collecting compounds

To find compounds we used a wildcard search. In this type of search words or parts of a word are substituted with wild card symbols thus bringing all forms of a word, e.g. German noun or adjective which has different case endings. When the asterisk symbol precedes or follows a stem, such a query allows finding all compounds containing the given stem. We searched compounds with *AntConc* tool *Cluster* with following settings: Cluster size – 1, minimal frequency – 10, minimal range (i.e. the number of different documents containing the word) – 5.

Using this algorithm we found compounds for each word from the list of most frequent nouns discussed above. We revealed that each word forms dozens, sometimes hundreds of compounds (*Bereich* (443), *Daten* (227), *Fall* (143), *Model* (293), *Fahren* (345)) though most of them have very low frequency occurring only once or twice in the whole corpus. This is especially true for compounds that appear in a narrow range of academic disciplines. Interestingly, that the results of our study of the Sample Corpus go in line with serious German corpus studies revealing that almost half (47%) of the word types are compounds, whereas most compounds are infrequent (83% of the compounds had a corpus frequency of 5 or lower) (Baroni, Matiasek, & Trost, 2002). The number of frequent compounds totals several dozens even in a relatively small Sample corpus. The examples are: *Berechnungsverfahren*, *Winkelbereich*, *Messbereich*, *Messverfahren*, *Brennverfahren*, *Temperaturbereich*, *Berechnungsmodell*, *Fertigungsverfahren*, *Frequenzbereich*, *Zeitbereich*, *Simulationsmodell*, *Rekonstruktionsverfahren*, *Kontaktmodell*, *Koordinatensystem*, *Gleichungssystem*, etc. So, for further research we selected compounds with a corpus frequency greater than or equal to 10. Discussion Procedure of collecting compounds

We think that compound nouns should be paid more attention in German for Specific Purposes course than it is done in typical GSP coursebooks. The challenges learners face studying this category of nouns arose from the variety of word formation patterns typical of this lexical category and translation problems. The specificity of translating German compounds into Russian is that they are often translated with a word combination or by descriptive translation. Translating such nouns into Russian can be difficult because of the incompatibility of declension forms in German and Russian. Moreover, the Russian language lacks constructions similar to absolute English constructions. The problem when translating German compound nouns is determined by a wide range of options and combinations of original morphemes, which is impossible in the Russian language. An excellent example is the pair *Radiowecker* (radio alarm) and *Weckerradio* (radio which has different functions and among the main ones is the alarm function). However, the relationship between components of compound nouns and their sequence in similar Russian constructions are different. For instance, *das Koordinaten system* is translated into Russian as a phrase where the genitive case is used: *system of coordinates*. These discrepancies can lead to students' difficulties in using and translating the vocabulary of scientific discourse

Another problem is compound polysemy. The same compound may be used as a general scientific term and denote some process in a specific scientific discipline. For example, *Lösungsverfahren* is

method of solving a problem. In chemical contexts it means a way of dissolving a substance. So, there is a risk that a learner can choose a wrong equivalent for translation.

In terms of word formation the same stems can occupy different positions in compounds both in the preposition of the second stems and in the postposition being the second/ last stem of a compound (Cf. **Untersuchungsergebnis** ‘research result’ and **Grundlagenuntersuchungen** ‘the exploration of the foundation’). The stem order in a compound is also important because it determines gender of the compound noun and as a consequence the choice of article. The compound has the same gender as the end stem. Thus, the word *der Produktionssystemeinfluss* (influence on production system) contains 3 stems, each of different gender: *die Produktion* is of feminine gender, *das System* is of neuter gender, and *der Einfluss* is of masculine gender, the whole compound is of masculine gender at that.

Our analysis shows that the Sample corpus contains quite a few examples of using compound nouns from general scientific vocabulary and with narrow specific meaning, with the former prevailing. It seems reasonable to select compound nouns of both types for GSP courses. It is worth noting that most coursebooks classify compounds according to the way of formation but not to their meaning and contexts of usage.

6.3.Data-based tasks for teaching compounds in the GSP classroom

With the tailor-made Sample corpus of German scientific texts at hand it is time to design ways of integrating the DDL approach into GSP syllabus for teaching compound nouns. The tasks could include the following set of exercises.

1) Exercises on **matching** a) compound stems; b) German compounds and their Russian equivalents.

2) Exercises **on translation** of sentences with compounds including the analysis of their immediate surroundings

3) Exercises **on filling in the gaps in sets of concordance lines**. These tasks are typical of textbooks declaring devotion to the DDL approach. Below there is an example of such a task (ex.1).

Ex.1. Fill in the gaps with a compound noun from the list that suits every sentence in the group
Gleichungssystem, Koordinatensystem, Fertigungstechnik

1) - Stattdessen werden sämtliche Leistungsbilanzen *in* ein lineares _____ überführt und in Form einer Matrixgleichung simultan gelöst.

-Hierzu wird das Problem *in* ein lineares _____ überführt.

2) - Aus diesen drei Bestimmungsgrößen kann *in* einem dreidimensionalen _____ der Behaglichkeitsbereich als Zielbereich für den Aufenthalt von Personen angegeben werden.

- Demgemäß kann ein Behaglichkeitsbereich *in* einem dreidimensionalen _____ angegeben werden.

-Wird in der FEM von einem kartesischen _____ ausgegangen, so wird sich im Allgemeinen keine eindimensionale Belastung ergeben.

3). - Dies führte bereits in der Vergangenheit zu Innovationsprozessen in den verschiedensten Bereichen der _____.

- Dies führte bereits in der Vergangenheit zu Innovationsprozessen in den verschiedensten Bereichen der _____.

Key: 1) *Gleichungssystem*, 2) *Koordinatensystem*, 3) *Fertigungstechnik*

4) A more sophisticated variant of **ex.1** implies that a compound is first formed from stems and then inserted into the gaps

5) Tasks on the analysis of the immediate surroundings of the keyword

Ex.2. Fill in the gap with the appropriate adjective. Note the preceding article/preposition.

1) eine ausreichend _____Standzeit

a) hohe b) hohen

2) Identifikation der _____ Systemparameter

a) entscheidender b) entscheidenden

3) bei Ausfall von _____Systemkomponenten

a) wichtigen b) wichtiger

4) Umlegung mit _____Systemzustand

a) veränderlichem b) veränderlichen

Key: 1) a, 2) b, 3) a, 4) a

The grammar focus of this exercise is that adjective ending depends on the preceding article or preposition

Ex.3. Fill in the gap with the article taking into account the compound noun case, preceding preposition, and sentence context.

- Fundamentverdrehung für _____ gesamten untersuchten Frequenzbereich.

- Die Probekörper erwärmen sich in _____ Übergangsbereich

- _____ Erprobungsabschnitt auf der Strecke Mannheim-Karls

Key: 1) *den* 2) *dem* 3) *ein*

6) Tasks on making up your own sentences with the compound nouns in question.[

7. Conclusion

The DDL-methodology appeared at the same time with linguistics corpus technology and is developing inside this field. Simultaneously it represents an independent direction within pedagogical didactics. As we said (section 1.2), DDL is not widely used in pedagogy and the wide introduction of DDL is hampered by a number of factors. This is especially true for teaching German. As for integrating the DDL approach into GSP course we can conclude that a special corpus should be collected first because sublanguages for specific purposes in some respects are remarkably different from General German. These corpora could then be used as a resource for development of teaching materials. We see our further research in the area in building a balanced German corpus containing special texts from different domains and different scientific genres. This corpus will allow us to study and describe the peculiarities of German for specific purposes. The following are interesting objects for research: verbs with prefixes, prepositional coordination, articles, conjunctions in complex sentences, etc. We hope that specially developed corpus-based teaching materials will help students to correctly use professional terminology, become more sensitive to issues of lexical collocatability, avoid both lexical and syntactic Russian calques while communicating in German, observe typical patterns of German sentence structure

and as a result ‘provide an effective formation of linguistic competence’ as other, e.g. Productive foreign language education technologies (Almazova, Eremin, & Rubtsova, 2016) do.

Acknowledgments

Authors would like to acknowledge colleagues from Leibniz Universität Hannover Dr. Sigrun Schroth-Wiechert and Anna Tilmans for support during this research and discussion of its results

References

- Almazova, N., & Kogan, M. (2013). Organizing polytechnic post-graduate students individual work on required reading corpora (within ESP Course). *Humanities and Science University Journal*, 6, 13-25.
- Almazova, N., & Kogan, M. (2014). Computer-Assisted individual approach to acquiring foreign vocabulary of students major. In P. Zaphiris & A. Ioannou (Eds.), *LNCS: Vol.8524: Learning and collaboration technologies. Technology-Rich environments for learning and collaboration. LCT 2014*. (pp. 248–257). Berlin, Germany: Springer, Cham. doi:10.1007/978-3-319-07485-6_25.
- Almazova, N.I., Eremin, Yu.V., & Rubtsova, A.V. (2016). Productive linguodidactic technology as an innovative approach to the problem of foreign language training efficiency in high school. *Russian linguistic Bulletin*, 3 (7), 50-54. doi: 10.18454/RULB.7.38
- Asinovsky, A., Bogdanova, N., Rusakova, M., Ryko, A., Stepanova, S., & Sherstinova T. (2009). The ORD speech corpus of Russian everyday communication «One Speaker’s Day»: Creation principles and annotation. In: V. Matoušek V.& P. Mautner (eds). *Lecture Notes in Computer Science: Vol. 5729ю Text, Speech and Dialogue. TSD* (pp. 248-257). Berlin, Heidelberg: Springer. doi: 10.1007/978-3-642-04208-9_36
- Baroni, M., Matiassek, J., & Trost, H. (2002). Predicting the components of German nominal compounds. In F. van Harmelen (ed.) *Proceedings of the 15-th European Conference on Artificial Intelligence (ECAI 2002)*, (pp.470–474). Amsterdam: IOS Press.
- Beliaeva, L., & Chernyavskaya, V. (2016). On methodology of knowledge presentation: terms and translation in seismic domain. *Journal of Siberian Federal University. Humanities & Social Sciences*, 12 (9), 2904-2912. doi: 10.17516/1997-1370-2016-9-12-2904-2912
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Pearson Education.
- Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, 67(2), 348–393. doi: 10.1111/lang.12224.
- Boulton, A. (2010b). Data-driven learning: Taking the computer out of the equation. *Language Learning*, 60(3), 534-572. doi: 10.1111/j.1467-9922.2010.00566.x.
- Boulton, A. (2010a). Data-driven learning: On paper, in practice. In T. Harris & M. Moreno Jaén (Eds), *Corpus linguistics in language teaching*. (pp. 17-52). Bern: Peter Lang.
- Boulton, A. (2017). Research timeline: Corpora in language teaching and learning. *Language Teaching*, 50(4), 483-506. DOI: 10.1017/S0261444817000167.
- Chernyavskaya, V. (2017). Metodologicheskie vozmozhnosti diskursivnogo analiza v korpusnoj lingvistike [Towards methodological application of discourse analysis in Corpus-driven Linguistics]. *Vestnik Tomskogo gosudarstvennogo universiteta. [Filologiya Tomsk State University Journal of Philology]*, 50, 135–148. doi: 10.17223/19986645/50/9.
- Chernyavskaya, V. E. (2018). Diskursivnyj analiz i korpusnye metody: neobxodimoe dokazatelnoe zveno? Obyasnitelnye vozmozhnosti kachestvennogo i kolichestvennogo podxodov [Discourse analysis and corpus approaches: a missing evidence-based link? Towards qualitative and quantitative approaches in language studies]. *Voprosy Kognitivnoy Lingvistiki*, 2, 31-37. [in Rus.]. doi: 10.20916/1812-3228-2018-2-31-37.

- Cobb, T., & Boulton, A. (2015). Classroom applications of corpus analysis. In D. Biber & R. Reppen (Eds) *Cambridge Handbook of English Corpus Linguistics* (pp. 478-497). Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139764377.027. 2015.
- Jaworska, S. (2015). Review of recent research (1998–2012) in German for Academic Purposes (GAP) in comparison with English for Academic Purposes (EAP): Cross-influences, synergies and implications for further research. *Language Teacher*, 48(2), 163-197. doi:10.1017/S026144481400038X.
- Leńko-Szymańska, A. (2017). Training teachers in data-driven learning: Tackling the challenge. *Language Learning & Technology*, 21(3), 217–241. Retrieved from <http://llt.msu.edu/issues/october2017/lenko-szymanska.pdf>
- Pavlovskaya, I.Y., & Gorina, O.G. (2015). Korpusno-kognitivnye metody obucheniya leksicheskoy storone rechi na inostrannom yazyke [Corpus-based cognitive oriented methods of learning and teaching lexical skills in a foreign language]. *Kognitivnye Issledovaniya Yazyka* [Cognitive Language Studies], 22, 552-554. [in Rus.]
- Sinclair, J. (Ed.) (1987). *Looking Up: An account of the COBUILD project in lexical computing and the development of the Collins COBUILD English Language Dictionary*. London and Glasgow: Collins ELT.
- Sosnina, E.P. (2006). O razrabotke i ispol'zovanii rossijskogo uchebnogo korpusa perevodov [On development and use of Russian learner translator corpus]. Trudy mezhdunarodnoj konferentsii «Korpusnaya lingvistika – 2006» [Proceedings of the International conference “Corpus Linguistics-2006”] (pp. 365–373). St. Petesburg: SPbSU. [in Rus.]. Retrieved from http://ling.ulstu.ru/linguistics/resourses/literature/articles/corpus_education_translation/.
- Stratonova, G. (2016). Interaktivnaya texnologiya distancionnogo obucheniya inostrannomu yazyku v vuze [Interactive distance learning technique for studying foreign language in the higher education]. In D.I. Kuznetsov (Ed.) *Gumanitarnaja obrazovatel'naja sreda tehničeskogo vuza*. Proceedings of the International Scientific and Methodical Conference (pp. 385-388). Saint Petersburg: Izd-vo Politehn. un-ta. [in Rus.]
- Strelkova, S. Yu. (2011). *Politics is ili politics are*: O nekotorykh sluchayakh formal'nogo i ponyatijnogo soglasovaniya v britanskom variante anglijskogo yazyka [*Politics is or politics are*: About some cases of formal and semantic noun-verb agreement in the British English]. *Inostrannye Yazyki v Shkole*. [Foreign languages at school], 10, 76-80. [in Rus.]
- Sysoyev, P.V. (2010). Lingvisticheskij korpus v metodike obucheniya inostrannym yazykam [Linguistic corpus in teaching foreign languages]. *Yazyk i Kul'tura*. [Language and Culture], 1(9), 99–111.
- Sysoyev, P.V., & Kokoreva, A.A. (2013). Obuchenie studentov professionalnoj leksike na osnove korpusa paralelnyx tekstov [Teaching professional terms to EFL students using parallel corpus]. *Yazyk i Kul'tura*. [Language and Culture], 1(21), 114 – 124. [in Rus.]
- Tarnaeva, L.P., & Osipova, E.S. (2016). Ispolzovanie resursov korpusnoj lingvistiki pri podgotovke perevodchikov v sfere professionalnoj kommunikacii [Students' cognitive activity in the process of using corpus-based linguistics resources when forming idiomatic competence of an interpreter in the sphere of professional communication]. *Voprosy Teorii i Praktiki* [Theory and Practice issues], 9-1(63), 205-209. [in Rus.]
- Thornbury, S. (2004). *Natural Grammar: The keywords of English and how they work*. Oxford: Oxford University Press.
- Vyatkina, N. (2016a). Data-driven learning for beginners: The case of German verb-preposition collocations. *ReCALL*, 28 (2), 207-226. doi:10.1017/S0958344015000269.
- Vyatkina, N. (2016b). Data-driven learning of collocations: Learner performance, Proficiency, and Perceptions. *Language Learning & Technology*, 20(3), 159–179. Available from <http://llt.msu.edu/issues/october2016/vyatkina.pdf>
- Yoon, Ch. (2016). Concordancers and dictionaries as problem-solving tools for ESL academic writing. *Language Learning & Technology*, 20 (1), 209–229. Available from <http://llt.msu.edu/issues/february2016/yon.pdf>.