

UDC 316.42

<https://doi.org/10.33619/2414-2948/74/34>

CORPUS-BASED DATA-DRIVEN LEARNING TO DEVELOP SENIOR STUDENTS' RESEARCH WRITING SKILLS: PRACTICAL INSIGHTS

©*Nurmatova G., ORCID: 0000-0002-4301-8089, Ph.D., Navoi State Mining Institute, Navoi, Uzbekistan, g.nurmatova77@gmail.com*

ОБУЧЕНИЕ СТУДЕНТОВ СТАРШИХ КУРСОВ РАЗВИТИЮ НАВЫКОВ НАУЧНОГО ПИСЬМА С ИСПОЛЬЗОВАНИЕМ ЭЛЕКТРОННОЙ ИНФОРМАЦИИ (DDL) И КОРПУСА: ПРАКТИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ

©*Нурматова Г., ORCID:0000-0002-4301-8089, Ph.D., Навоийский государственный горный институт, г. Навои, Узбекистан, g.nurmatova77@gmail.com*

Abstract. The article introduces corpus-based DDL (Data-driven Learning) technologies in teaching ESP (English for Specific Purposes). The aim of the author is twofold: to offer English language teachers to design grammatical and lexical activities to develop senior students' research writing skill and to assist senior students to construct scholarly field-related sentences. For the purpose of the study, the author used a mini manually compiled corpus of robotics, one of the branches of mechanical engineering and demonstrated practical instructions of corpus-based grammatical and lexical insights with DDL technologies. In spite of some limitations and future research, the findings of the study can contribute language teachers to develop senior students' productive (writing) skills via designing corpus-based data driven materials as well as improve students to construct grammatically and lexically correct sentences for succeeding in their further research and career growth.

Аннотация. В статье представлены технологии DDL (Data-Driven Learning) на корпусной основе обучающихся ESP (английский для специальных целей) студентов старших курсов развитию навыков научного письма. Автор ставит перед собой двоякую цель: предложить преподавателям английского языка разработать грамматические и лексические упражнения для развития у студентов старших курсов навыка научного письма, которые помогут сконструировать предложения, относящиеся к предметной области. Для исследования автор использовал составленный вручную мини-корпус робототехники, одного из разделов машиностроения, и продемонстрировал практические инструкции грамматического и лексического понимания с использованием технологий DDL в корпусе. В данной работе присутствуют некоторые отклонения и ограничения, которые открывают путь для дальнейшего исследования, результаты которого могут применяться англоязычными преподавателями для развития продуктивных (письменных) навыков у студентов старших курсов посредством разработки материалов, основанных на корпусе данных, а также строить грамматически и лексически правильные предложения для научного и профессионального роста.

Keywords: corpus, DDL, field-related lexis, research writing skills.

Ключевые слова: корпус, DDL, отраслевая лексика, навыки научного письма.

Introduction

Scientific communication is increasingly mediated by the English language. References and citation to English language publications now constitute 85% of all citations in French academic journals; in the Science Citation Index, English now accounts for 95% of all publications [6]. Although teaching field-related (STEM) English is inevitably time and effort consuming for English language teachers, today achievements of science and technology has resulted in the enrichment of language lexis and even turning into vernacular words of speech. Therefore, today we have to adhere to Eugen Wustern's, an Austrian terminologist's statement: "*A language should be learnt in relation with development of science and technology*" [8].

Throughout ten years of teaching experience at Navoi State Mining Institute, I observed the students' interests and needs for the English language and practiced designing various types of activities to develop their both receptive and productive skills. I found out that the students are tend to obtain information of their field of study in the English language more easily as it comforts to their content knowledge and catches their interest to grasp it in the English language. Moreover, most senior level student, masters and PhDs are more willing to study and write scholarly papers in English with the aim to introduce and share their innovative ideas to other countries' researchers. Consequently, this problem was vitally reflected in learners' writing skills. Indeed, designing activities for both receptive and productive skills should represent an authentic language of real researcher. However, the process of designing activities may differ; if, for example, for reading skills, there is a plenty of authentic material that can be easily matched with the language level of a student and a content of his field of study, developing and designing activities for writing skills, especially research writing skills can be quite time consuming and may need lots of efforts and moreover, field knowledge of a student. Therefore, it is not an easy task for a language teacher, especially for a non-native language teacher to deal with this matter [1, 2].

As M. J. Cotter (as cited in O.Mudraya 2006) says: "Since ESP teachers are rarely experts in the learners' field of knowledge and sometimes they are not native speakers of the language, they will be faced with two types of problems:

1. producing examples in a language that is foreign to them and
2. in a field of knowledge that they do not master to formulate statements that exemplify a given construction relying only on intuition in the case of ESP in particular [4].

Thus, teaching English field-related lexis in research context is becoming one of the most actual matters for ESP teachers that overwhelm them to search for more advanced technologies and methods. For the purpose of this study is to introduce corpus-based DDL technologies for designing research writing materials for ESP and ESL classroom activities of senior students, which can assist non-native English language teachers to produce lexically and grammatically authentic examples to develop research writing activities [5].

Theoretical and methodological principles of the research

Corpus

Corpus is generally defined as a large collection of authentic texts in electronic format. Corpus-based language teaching has been praised as a revolution in teaching by [9, 12]. Moreover, corpus linguistics has opened new possibilities for terminology. Digital corpora and corpus manager and analysis software allow working with big number of documents; extract comprehensive datasets necessary to examine cognitive, linguistic and communicative dimensions of terminology [11].

Usually specialized adhoc corpora are compiled for terminology research which represents the use of language in the selected specialized domain [11]. Therefore, terms in their turn can be

regarded as field-related lexis and/or as any lexis of linguistics is able to represent grammatical and lexical characteristics.

For the present research, a specialized corpora has been compiled for terminology extraction: it consists of twelve domains of engineering scholarly articles of AntCorGen corpus. Figure one presents how corpus of a particular domain can be compiled in seconds.

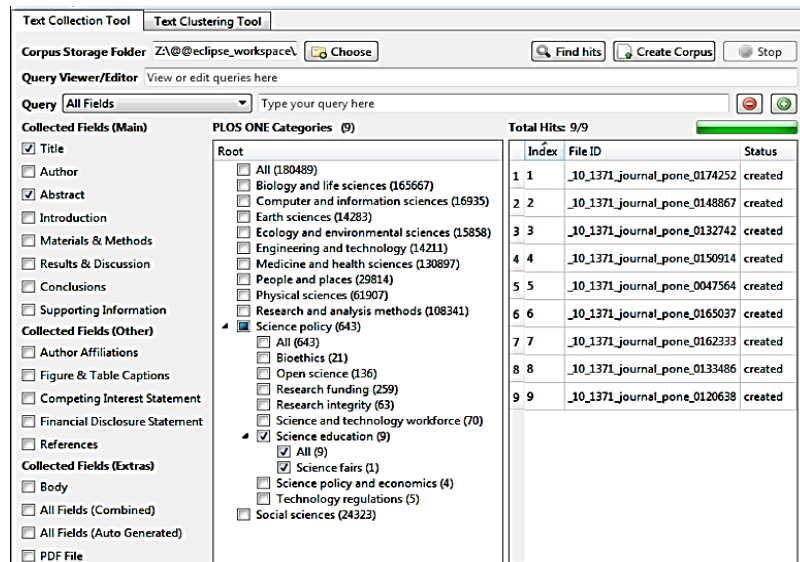


Figure 1. AntCorGen's corpus compilation process

The choice of this type of corpus has several reasons: its platform, PlosOne source, is freely accessible; it provides analyzing terms and field-related lexis in authentic scholarly publication, additionally, it includes articles of the authors from beginners to expert researchers that enable adjusting its communication setting to appropriate audience i.e. master and doctoral students. Out of these twelve domains I selected robotics, one of the branches of mechanical engineering domain for designing research writing activities.

Data-Driven Learning

Data-driven learning (DDL) is an approach to foreign language learning and was first applied at one-to-one tuition as English for Academic Purposes at Birmingham University in 1991 to a large auditory of students with specially prepared handout materials (studwood.ru). As Johns (1991), the “father” of DDL, advocated the learning-centered value of DDL, calling “every student a Sherlock Holmes mean that any student can be an explorer of a language perceiving most language aspects. However, it is not possible to apply DDL technology without corpus-based techniques and its tools, because corpus represents an authentic material and its tools assist in identifying grammatical or semantical patterns of the natural language. As Gabrielatos (2005) and Hadley (2002) point out: “The use of corpus linguistics in language learning is called Data-Driven Learning” [3, 5]. Therefore, DDL and corpus software fulfil each other for effective ESP classroom activities. In this study I am going to share with some of the corpus-based DDL grammatical and lexical patterns extraction from created corpus of robotics (domain of mechanical engineering), where we observed the behavior of terms and field-related lexis for designing research writing activities. I used one more software of A.Laurence, AntConc's concordance tool which is suitable for language analysis as well as for educational purposes.

Corpus-Based DDL

There are several computer programs for corpus analyses among which the most popular are corpus.byu.edu, AntConc, WordSmith Tools and some others. The wide range of software serve for generating a corpus of any field and can be very effective for developing teaching materials as well as for delivering sessions [13]. For this study as I mentioned above I used AntCorGen and AntConc software developed by professor A. Laurence of Waseda University, Japan.

Corpus size may vary as it depends on the aim of your delivering session and the syllabus covering a set of lessons for this topic. If, for example, an instructor intends to present vocabulary of a series of lessons about a particular topic, the size of that corpus can be as large as possible. Tools of the software such as frequency, concordance and file view are very effective for developing classroom materials to assist in analyzing the collected corpus [10]. Word frequency tool can present the frequencies of all words in a corpus enabling a careful observation for analyses (Figure 2) and collocations (Figure 3).

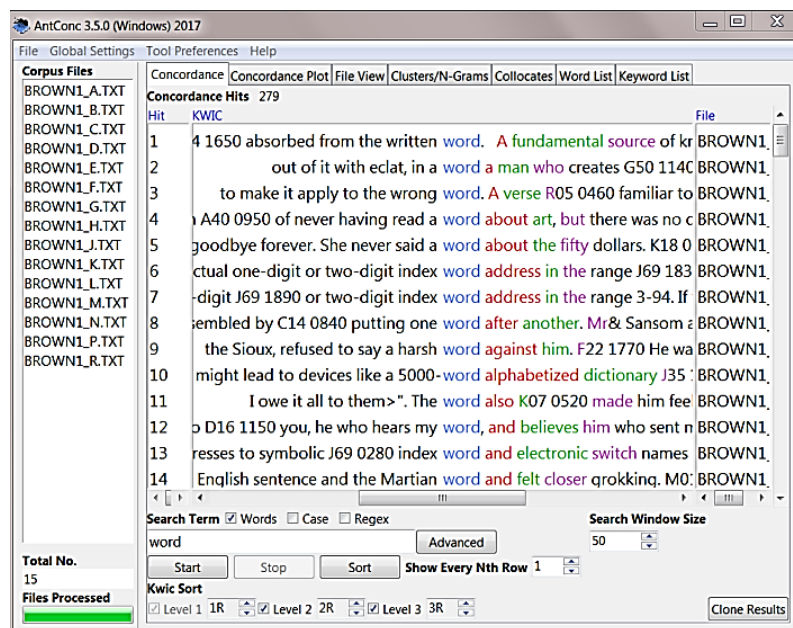


Figure 2. AntConc's display of concordance tool

As mentioned above observation is very important as we can see here what grammatical patterns could be included in our activities. Ant Conc software was successfully used for this purpose. It is important to look carefully at sentences in a corpus and notice grammatical patterns before using concordance tools. Some grammatical stance present lexical closure and saturation and style of the language and can be particular interesting to construct sentences in research context. Below I am introducing some tips to notice grammatical and stylistic construction of sentences in research article. There are examples and explanations from Robotics corpus that show how it is possible to notice grammatical forms and patterns that will assist both to extract terms and terminological combinations related to students' field of study and to construct grammatically correct sentences:

I. To be+indefinite articles (*is (a/an) /are*). This combination is important for two reasons. First, *is a* can:

- a) explain subject and verb agreement

Efficient exploration in high-dimensional spaces is a major challenge in building learning systems.

b) is+a can also help mine terms:

The PR2 is a human-scale robot with an omnidirectional wheeled base, a torso that translates vertically, two arms with grippers, a pan/tilt head with cameras, and various other sensors, such as tactile sensors.

Robot-supported therapy is a rehabilitation method allowing patients to train their arm-hand with high intensity, a large amount of practice and minimal use of therapists' time.

c) is+PII

The algorithm is designed to have a reduced computational complexity in order to be applied to low performance embedded systems, minimizing, as a result, both cost and power consumption.

d) is+adj/adv. Here we can see very interesting field-related adjectives that can be useful for mechanical engineering student.

The attractive properties of HMA lie in the fact that it is thermoplastic and thermoadhesive.

e) is that clause

One advantage of this approach is that it is simple to implement and interpret.

This samples of grammatical stance

II. Clause relatives: *that, which, where, when, why, how* etc;

These implementations are generally known as EKF-SLAM [9–11]. However, one of the main problems with EKF-SLAM is that it requires having geometric models of the environment, which limits its use to environments where such models are available. An alternative to these models are the so-called scan-correlation procedures, where the maximum alignment between two sets of data is estimated.

III. *to* is a very interesting for observation both as a preposition and as an infinitive verb:

The self-determined and self-directed exploration for embodied autonomous agents is closely related to many recent efforts to equip the robot with a motivation system producing internal reward signals for reinforcement learning in pre-specified tasks.

As it is seen from the above examples, we can observe not only grammatical stance, but also terminological combinations and/or terminological collocations. We take the highest frequent word and look at concordance tool. Collocation analyses also display an interesting insights: the highest frequency word “robot” does not have collocations, but as word combination in concordance tool it has lots of word combinations. Therefore it is possible to make differences between word combinations and collocations. According to glossary of CL by Baker (2006), collocations are stable combinations in speech. As G. Nurmatova (2021) notices terminological collocations can be regarded as field-related lexis and become steady used words among field experts if the frequency of the node word equals to the frequency of a collocate word (Table 1) [3, 8].

Table 1

COLLOCATIONS OF THE HIGHEST FREQUENCY WORDS

structure	<i>a passive mechanical structure</i>
system	<i>a robotic system</i>
output	<i>a sensing output</i>
controller	<i>a fabrication/attachment controller</i>
controller	<i>a motion controller</i>
arm	<i>a robotic arm</i>

I applied J. Pearson’s techniques (1998) to determine some signal words that can assist to extract terms in the created corpus. As terms can also be regarded as field-related lexis I found

designing lexical activities for term extraction extremely helpful [10]. The examples below illustrate how signal words serve as navigators for terms extraction:

1. *More/less* assists in identifying appropriate adjectives and adverbs that show research comparisons, observations, methods or novelty of the study.

The past two decades in robotics have seen the emergence of a new trend of control in robotics which is rooted more deeply in the dynamical systems approach to robotics using continuous sensor and action variables. This approach yields more natural movements of the robots and allows to exploit embodiment effects in an effective way for an excellent survey.

2. “And” and “or” conjunctions can be helpful to identify pair words and can be applied for activities especially related with synonymy:

Advantages of robot-assisted surgery include improved dexterity and accuracy, steep learning curve, and tele surgery.

3. General class word such as *method, function, model* and others can be useful for term extraction (J.Pearson, O.Muraya).

Robot-supported therapy is a rehabilitation method allowing patients to train their arm-hand with high intensity, a large amount of practice and minimal use of therapists' time.

Thus, concordance tool helps to construct collocation networks, which is very useful to extract field-related and terminological collocations.

Corpus-Based DDL in practice

The next steps will reflect how to develop corpus-based data-driven activities for research writing activities. For this, analyses of small text in corpus gave lots of advantages in identifying meaning of lexis. Since the number of tokens/words in such corpus is small, it enabled learners to analyze the field-related lexis more easily. First, we start with indefinite article, as it usually introduces a new concept of a word or word combination, and/or even a definition of the term.

I uploaded a small extract from an article and introduced its procedure step by step. This set of activities is called “terms extraction for describing the purpose of their function”, where students will learn how to introduce terminological concept or notion and explain its purpose of usage in their field-related research writing. It moved from grammar analyses to terminological vocabulary perception.

Step 1-2. The teacher/instructor writes out all word combinations with indefinite article, but introduces only nouns to students. After this, introduces these word combinations, that most of them are terminological collocations (Table 1).

Step 3. The teacher/instructor distributes the text with underlined word combinations and infinitive verbs in bold (see appendix A).

Step 4. The teacher/instructor writes out other *adjectives + noun* combinations in three columns with indefinite *a/an* articles, definite article *the*, and without articles (Table 2):

Table 2
USING ARTICLES WITH ADJECTIVES + NOUN FIELD-RELATED COMBINATIONS

A
<i>a</i> passive mechanical structure <i>a</i> robotic system <i>a</i> camera <i>a</i> sensing output <i>a</i> fabrication/attachment controller <i>a</i> motion controller <i>a</i> robotic arm

THE
the arising physical stimuli the sensor morphology the suitable motion the red lines the green lines the involved processes the physical interactions the target object the proposed technological solution the mechanical structure the attractive properties the thermoplastic and thermoadhesive nature of HMA the sensing characteristics
NO ARTICLE
useful geometrical information mechanical structures physical interactions passive mechanical structures sensing purpose increasing/decreasing material temperature different mechanical structure active sensing

Step 5. The teacher/instructor makes students to notice on cases of using articles before adj+noun. Explain when introducing a new notion or concept – a. for detailed and concrete description, use –the, for plurality or repeated ones- no article.

Step 6. The teacher/instructor asks students to put aside the text and distribute the text with gaps to fill in the articles (see appendix B).

Step 7. The teacher/instructor asks students to match the texts. They compare their answers themselves.

Step 9. Draw the students' attention on how terms are introduced in a sentences.

After analyses of adj+noun, the case where a new term or concept is introduced (a passive mechanical structure and a robotic system) and understanding the role of indefinite article, the students should pay attention how these terms are introduced in the sentence. For that, we have to notice infinitive verbs in the same sentence where the terms have been found (because they serve for expressing the purpose). After more detailed analyses we notice that the first term 'a passive mechanical structure' is WHAT and the second 'a robotic system' is BY WHAT. Both of them describe the purpose of their usage by infinitive verb. And then ask students to write a similar sentence i.e. introducing two terms of their field i.e. two notions (A term or terminological combination of adj+noun should be used) and their purpose of usage.

In order to sense a possibly unknown target object in uncertain environment, a passive mechanical structure is used by a robotic system to probe the object via suitable motion.

a passive mechanical structure is a term and in order to look at its meaning, we have to use concordances with *mechanical*. Then for explanation of this term, in a file view tool to look at the text where a passive mechanical structure is used once more. And here we can see that this term a passive mechanical structure is explained in several sentences:

In order to realize the concept, the proposed technological solution is to use **a** robotic arm that is able to repeatedly fabricate, dispose and manipulate passive mechanical structures for sensing

purpose. Hot Melt Adhesive (HMA) is chosen as the material for the mechanical structure. The attractive properties of HMA lie in the fact that it is thermoplastic and thermoadhesive.

Thus we have an understanding that a passive mechanical structure:

a) is for sensing purpose

b) its material is hma (hot melt adhesive) and in its turn we know that it is used for mechanical structure.

c) the properties of hma is thermoplastic and thermoadhesive

From the last *thermoplastic and thermoadhesive* adjectives we have an understanding of the consistency of a passive mechanical structure

By the same way we can analyze other adj+noun terminological combinations in corpus. They are *a robotic system, a sensing output, a fabrication/attachment controller, a motion controller, a robotic arm*. The explanation of some of them may be given in the same sentence, such as *a robotic arm*

In order to get an iron cast, an iron and carbon alloy is used by a heating temperature of more than 2,14⁰ C.

Students should notice lexical choice with the most frequent words: *sensing (6), mechanical (5), can (4), material, motion, physical, suitable, system*. Matching activity: the students should find the characteristics for each these combinations from the text and write their role. For example: sensing characteristics; sensing output; sensing purpose; sensing system; active sensing.

The teacher/instructor can also ask students to look at infinitive verbs in bold and can show them a slide projector of infinitive verb concordances for observing and eliciting *TO* as infinitive (Table 3).

Table 3

ACTIVITY FOR ELICITING *TO* INFINITIVE FROM CORPUS

mechanical structures and integrate them in situ	<i>to adjust</i> the sensor morphology and therefore
mechanical structures, and/or the suitable motion	<i>to initiate</i> different physical interactions.
suitable motion can be executed in order	<i>to obtain</i> suitable amount and type of desired....
structure is used by a robotic system robot and the target object	<i>to probe</i> the object via suitable motion. A
nature of HMA will enable the system use a robotic arm that is able	to realize the concept, the proposed technologic
of robotics active sensing system In order	to repeatedly fabricate different mechanical structures
the camera, while the green lines correspond	to sense a possibly unknown target object in
physical interactions. The red lines correspond	to the arising physical stimuli into useful geometry
and type of desired stimuli additionally, due	to the involved processes during the physical

These type of activities were applied for senior students of master and PhD students to structure field-related sentences with further application in their scholarly publication first in local/national level conferences and then in international level conferences.

Conclusion

In this paper, we introduced corpus-based DDL samples of grammatical and lexical activities that can be comforted into senior students' research writing classroom design and successfully applied for classroom activities. Although application of some software and tools is time consuming and/or may demand some additional training out of the classroom, the effectiveness of corpus-based DDL technologies were successfully applied and warmly welcomed in designing classroom

materials by English language teachers. In spite of these drawbacks, the basic idea of this paper is to aware teachers and learners about the possibilities of corpus-based DDL technologies to develop senior students' productive skills i.e. research writing effectively. Indeed, the progress track of research writing skills is possible to carry out in further research.

Appendix A

Basic concept of robotics active sensing system

In order to sense a possibly unknown target object in uncertain environment, a passive mechanical structure is used by a robotic system to probe the object via suitable motion. A camera will observe this physical interaction and transduce the deformation of the structure due to the arising physical stimuli into useful geometrical information as a sensing output. Based on the output, a fabrication/attachment controller and a motion controller can decide the necessity and the way to adjust the sensor morphology in situ, i.e. the shape, size and connection of the mechanical structures, and/or the suitable motion to initiate different physical interactions. The red lines correspond to the sensing output obtained from the camera, while the green lines correspond to the involved processes during the physical interactions between the robot and the target object.

In order to realize the concept, the proposed technological solution is to use a robotic arm that is able to repeatedly fabricate, dispose and manipulate passive mechanical structures for sensing purpose. Hot Melt Adhesive (HMA) is chosen as the material for the mechanical structure. The attractive properties of HMA lie in the fact that it is thermoplastic and thermoadhesive. The material can be transformed between solid and liquid phases by increasing/decreasing material temperature, and the material in liquid phase exhibits adhesive property, while it forms bonding when solidified by cooling. More specifically, it is hypothesized that: (1) the thermoplastic and thermoadhesive nature of HMA will enable the system to repeatedly fabricate different mechanical structures and integrate them in situ to adjust the sensor morphology and therefore the sensing characteristics (2) once the sensor morphology is adjusted, active sensing via suitable motion can be executed in order to obtain suitable amount and type of desired stimuli (3) additionally, due to the use of a robotic system, these two processes can be executed autonomously.

APPENDIX B

Basic concept of robotics active sensing system

In order to sense a possibly unknown target object in uncertain environment, ___ passive mechanical structure is used by ___ robotic system to probe the object via suitable motion. ___ camera will observe this physical interaction and transduce the deformation of the structure due to ___ arising physical stimuli into ___ useful geometrical information as ___ sensing output. Based on the output, ___ fabrication/attachment controller and ___ motion controller can decide the necessity and the way to adjust the sensor morphology in situ, i.e. the shape, size and connection of the mechanical structures, and/or the suitable motion to initiate ___ different physical interactions. ___ red lines correspond to ___ sensing output obtained from the camera, while ___ green lines correspond to ___ involved processes during ___ physical interactions between the robot and the target object.

In order to realize the concept, ___ proposed technological solution is to use ___ robotic arm that is able to repeatedly fabricate, dispose and manipulate passive mechanical structures for sensing purpose. Hot Melt Adhesive (HMA) is chosen as the material for ___ mechanical structure. The attractive properties of HMA lie in the fact that it is thermoplastic and thermoadhesive. The material can be transformed between solid and liquid phases by ___ increasing/decreasing material temperature, and the material in liquid phase exhibits adhesive property, while it forms bonding when solidified by cooling. More specifically, it is hypothesized that: (1) the thermoplastic and thermoadhesive nature of HMA will enable ___ system to repeatedly fabricate different mechanical structures and integrate them in situ to adjust the sensor morphology and therefore the sensing characteristics (2) once the sensor morphology is adjusted, ___ active sensing via ___ suitable motion can be executed in order to obtain suitable amount and type of desired stimuli (3) additionally, due to the use of ___ robotic system, these two processes can be executed autonomously.

References:

1. Anthony, L. (2006). Developing a freeware, multiplatform corpus analysis toolkit for the technical writing classroom. *IEEE Transactions on Professional Communication*, 49(3), 275-286. <https://doi.org/10.1109/TPC.2006.880753>
2. Anthony L. AntCorGen (Version 1.1. 2)[Computer Software] //Tokyo, Japan: Waseda University. Disponível em: Disponível em: <https://www.laurenceanthony.net/software> Acesso em. 2019. V. 25.
3. Baker, P. (2006). *Glossary of corpus linguistics*. Edinburgh University Press.
4. Corino, E., & Onesti, C. (2019, February). Data-driven learning: a scaffolding methodology for CLIL and LSP teaching and learning. In *Frontiers in Education* (Vol. 4, p. 7). Frontiers. <https://doi.org/10.3389/educ.2019.00007>
5. Gabrielatos, C. (2005). Corpora and Language Teaching: Just a Fling or Wedding Bells?. *Tesl-Ej*, 8(4), n4..
6. Hyland, K. L. (2009). English for professional academic purposes: Writing for scholarly publication. *English for specific purposes in theory and practice*.
7. Mudraya, O. (2006). Engineering English: A lexical frequency instructional model. *English for Specific Purposes*, 25(2), 235-256. <https://doi.org/10.1016/j.esp.2005.05.002>
8. Nurmatova, G. Kh. (2021). Problems on Corpus-Based Linguo-Statistic Study of Engineering Terms with application of AntConc and GraphColl Software Tools: *Dissertation for degree of Doctor of Philosophy in (PhD) in Applied Linguistics*. Fergana State University, 193
9. Corino, E., & Onesti, C. (2019, February). Data-driven learning: a scaffolding methodology for CLIL and LSP teaching and learning. In *Frontiers in Education* (Vol. 4, p. 7). Frontiers. <https://doi.org/10.3389/educ.2019.00007>
10. Pearson, J. (1998). *Terms in context* (Vol. 1). John Benjamins Publishing.
11. Rackevičienė, S., Valūnaitė Oleškevičienė, G., & Cheiker, K. (2020). Terminology in Media Discourse: A Case Study of Terms Denoting Phobia Types. *Research in Language*, 18(4), 359-380. <https://doi.org/10.18778/1731-7533.18.4.01>
12. Sinclair, J. M. (Ed.). (2004). *How to use corpora in language teaching* (Vol. 12). John Benjamins Publishing.
13. TALC 2018 Pre-Conference Workshop, UK: Cambridge University Press, 2018/7/18-21.

Список литературы:

1. Anthony L. Developing a freeware, multiplatform corpus analysis toolkit for the technical writing classroom // IEEE Transactions on Professional Communication. 2006. V. 49. №3. P. 275-286. <https://doi.org/10.1109/TPC.2006.880753>
2. Anthony, L. (2019). AntCorGen (Version 1.1. 2)[Computer Software]. Tokyo, Japan: Waseda University. Disponível em: Disponível em: <https://www.laurenceanthony.net/software> Acesso em, 25.
3. Baker P. Glossary of corpus linguistics. Edinburgh University Press, 2006.
4. Corino E., Onesti C. Data-driven learning: a scaffolding methodology for CLIL and LSP teaching and learning // Frontiers in Education. Frontiers, 2019. V. 4. P. 7. <https://doi.org/10.3389/educ.2019.00007>
5. Gabrielatos C. Corpora and Language Teaching: Just a Fling or Wedding Bells? // Tesl-Ej. 2005. V. 8. №4. P. n4.
6. Hyland K. L. English for professional academic purposes: Writing for scholarly publication // English for specific purposes in theory and practice. 2009.

7. Mudraya O. Engineering English: A lexical frequency instructional model // English for Specific Purposes. 2006. V. 25. №2. P. 235-256. <https://doi.org/10.1016/j.esp.2005.05.002>
 8. Nurmatova G. Kh. Problems on Corpus-Based Linguo-Statistic Study of Engineering Terms with application of AntConc and GraphColl Software Tools: Dissertation for degree of Doctor of Philosophy in (PhD) in Applied Linguistics. Fergana State University, 2021. P. 193.
 9. Corino E., Onesti C. Data-driven learning: a scaffolding methodology for CLIL and LSP teaching and learning // Frontiers in Education. Frontiers, 2019. V. 4. P. 7. <https://doi.org/10.3389/educ.2019.00007>
 10. Pearson J. Terms in context. John Benjamins Publishing, 1998. V. 1.
 11. Rackevičienė S., Valūnaitė Oleškevičienė G., Čeiker K. Terminology in Media Discourse: A Case Study of Terms Denoting Phobia Types // Research in Language. 2020. V. 18. №4. P. 359-380. <https://doi.org/10.18778/1731-7533.18.4.01>
 12. Sinclair J. M. H. (ed.). How to use corpora in language teaching. John Benjamins Publishing, 2004. V. 12.
- TALC 2018 Pre-Conference Workshop, UK: Cambridge University Press, 2018/7/18-21.

*Работа поступила
в редакцию 20.11.2021 г.*

*Принята к публикации
28.11.2021 г.*

Ссылка для цитирования:

Nurmatova G. Corpus-based Data-driven Learning to Develop Senior Students' Research Writing Skills: Practical Insights // Бюллетень науки и практики. 2022. Т. 8. №1. С. 236-246. <https://doi.org/10.33619/2414-2948/74/34>

Cite as (APA):

Nurmatova, G. (2022). Corpus-based Data-driven Learning to Develop Senior Students' Research Writing Skills: Practical Insights. *Bulletin of Science and Practice*, 8(1), 236-246. <https://doi.org/10.33619/2414-2948/74/34>