CORPUS-BASED ANALYSES OF STEM LEXIS FOR ENGLISH FOR SPECIFIC PURPOSES (ESP)

Nurmatova Guzal Hakimovna
g.nurmatova77@gmail.com
Associate professor
Navoi State Mining Institute

Abstract: the article presents the advantages of corpus-based analyses for material design and developing activities to acquire STEM lexis in ESP context.

Key words: *corpus, ESP, STEM, activities, software, terms, frequency, concordances.*

Annotatsiya: ushbu maqola ESP doirasida STEM leksikasining korpus tahlillari asosida o’qitish va o’rganish afzalliklarini yoritib beradi.

Kalit so’zlar: *korpus, ESP, STEM, mashqlar, dastur, termin, chastota, konkordans.*

Аннотация: данная статья подставляет преимущества корпусного анализа для разработки материалов и заданий с целью освоения STEM лексики в обучение и преподавание ESP.

Ключевые слова: *корпус, ESP, STEM, упражнения, программа, термины, частота, конкордансы.*

Research and teaching English is becoming one of the most actual matters for ESP teachers. 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 (Hyland, 2007). Therefore, teaching field-related (STEM) English is inevitably requiring for English language teacher and Eugen Wustern, an Austrian terminologist is totally right when he says: *“A language should be learnt in relation with development of science and technology” (*E. Wüster).

A strong command of receptive that is reading and productive that is writing skills are essential to train for future engineering students as well as for other fields. Although both skills should represent an authentic language of real researcher, the process of developing activities may differ. If 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 writing skills,

especially research writing skills can be quite time consuming and may need lots of efforts. Therefore, it is not an easy task for a language teacher, especially for a non-native language teacher to deal with this matter.

As M.J.Cotter (2006) says: “Since ESP teachers are really 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,

So, what can help a non-native English language teacher to produce lexically and grammatically authentic examples to develop research writing activities? The advent of the computer, and modern technologies in the 21st century opened the way in linguistics. Corpus linguistics allows to make easy and quick analyses of authentic material. As A.O’Keeffe and Michael McCarthy say “Today, corpus have become a natural tool to present applied linguistics, powerful, easy to use and more than up to the tasks that researchers demand on them”

The wide range of software helps us to generate a corpus of any field and be very effective for developing teaching materials as well as for delivering sessions. As an example, I am dealing with one of the branches of mechanical engineering as robotics.

Having created a corpus in Robotics, we can see the number of tokens on the top.

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, you intend to present vocabulary of a series of lessons about Robotics, the size of your 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 an assist in analyzing the collected corpus.

Word frequency tool presents us the frequencies of all words in a corpus. This enables us a careful observation. Observation is very important as we can see here what grammatical patterns could be included in our activities. After careful examination we can move to concordances tool. Some grammatical stance not only present lexical closure and saturation and style of the language, but also can particular be interesting for constructing research writing activities for grammatical patterns.

Not only grammatical stance, but also terminological combinations and terminological collocations are worth to observe, as the highest frequent word of concordance tool. Collocation analyses can be also very interesting. In the analyzing corpus, for example, the highest frequency word “robot” does not have collocations, but as word combination in concordance tool it has lots of word combinations. Therefore we should be able make differences between word combinations and collocations. According to glossary of CL by Baker, collocations are stable combinations in speech. So in our case they can be regarded as terminological collocations that are steadily used among experts. The word “system”, for example, has the following collocations. This word has also lots of word combinations as well.

For task developing, analyses of small text in corpus can give lots of advantages in identifying meaning of lexis. The number of tokens/words in such corpus is small and it will enable us to analyze the field-related lexis more easily. For example, we can 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.

The 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 writing. It will move from grammar analyses to terminological vocabulary perception. And as Chung & Nation notice we will find out how Corpus linguistics can thus contribute to rendering learning a foreign language more effective since students will be faced with «real language» and the ESP teacher plays a fundamental role in teaching students to recognize technical vocabulary and understanding definitions (Chung & Nation, 2003).

Corpus-based analyses enables both a teacher and a learner to look at a research language in a new angle of aspect, such as noticing cases of using articles before adj+noun (for a student), or explaining a new notion or concept, drawing the students’ attention on how terms are introduced in a sentences by noticing infinitive verbs in the same sentence where the terms have been found (for a teacher).

References:

1. Mudraya. O. (2006). Engineering English: A lexical frequency instructional model. *English for Specific Purposes*, *25*, 235–256.

2. Cotter, M.J. (2006). Teaching terms: A Corpus-based approach to terminology in ESP classes. Lisbon School of Accountancy and Administration

(ISCAL). Portugal, p. 499-506.

3. Yilmaz m. & Özdem ertürk z. (2017). A Contrastive Corpus-Based Analysis of the use of reporting verbs by native and non-native ELT researchers. *Novitas-royal(research on Youth and Language), 11(2), 112-127*.

4. Zacharof M.P.&A.Charalambidou. (2018). An exploration of the Sub-Register of Chemical Engineering Research papers Published in English. *MDPI, 6 (30), 1-19.*

5. Zhang.M. (2013). A Corpus-Based Comparative Study of Semi-technical and technical vocabulary. *The Asian ESP Journal, 9 (2), 148-172.*