THE ROLE OF CORPUS BASED LEARNING IN ESP CLASSROOM

Pulatova Durdona Ravshanovna
Doctoral student of Tashkent State
Pedagogical University named after Nizami

Abstract. This article highlights the importance of corpus based learning in teaching specific terminologies to nonphilological students in developing their professional communicative competence.

Key words: corpus based learning, terminology, corpus, concordance lines

In second-language education, vocabulary acquisition has drawn significant interest. Researchers have discovered a significant connection between vocabulary size and a student's writing and reading skills (Staehr, 2008), and it is estimated that learners must understand 98% of the words in a text for conceptual understanding (Hu & Nation, 2000; Schmitt, Jiang, & Grabe, 2011), or 6,000-7,000 word families for general speech comprehension (Hu & Nation, 2000; Schmitt, Jiang, & Grabe, 2011). Teaching language can be a challenging prospect in light of reports like these (Nation, 2006). The task becomes considerably harder in special-purposes teaching environments.

The task becomes considerably more challenging in special-purposes classroom practices. They have to master new technical terminology, which can represent up to 30% of discipline-specific texts (Chung & Nation, 2003; Coxhead, 2018), and many words students have already learned take on particular meanings, such as solution in chemistry or force in physics (Mudraya, 2004). Most importantly, students must understand how vocabulary choices influence meaning and which words are appropriate in different registers. Students must be able to express and think on disciplinary topic using words as instruments with flexibility and correctness (Koda, 2018; Nagy & Townsend, 2012; Woodward-Kron, 2008).

Recognizing the need to address vocabulary in specific contexts, researchers have developed word lists for a wide range of situations, along with English for Academic Purposes (Coxhead, 2000; Gardner and Davies, 2014), Engineering English (Graham & Osment, 2013; Ward, 2009), Medical English (Wang, Liang, & Ge, 2008), and, most recently, Plumbing English (Coxhead & Demecheleer, 2018). These and other word lists have provided insights on words that are peculiar to certain areas of language use.

However, because most word lists contain over 1,000 terms, these word lists primarily set a target for students' knowledge and, practically, a longer-term target (Watson Todd, 2017). Instructors are still unsure about how to teach vocabulary and which words to start with.

Large word lists are notably problematic for more in-depth vocabulary instruction approaches like corpus based learning. Corpus based activities involve students interacting with linguistic data from a corpus in order to instill rules or patterns in the language (Smart, 2014). Over the last 30 years, Corpus based approach has received increasing attention and has shown great promise in assisting students in recognizing and applying lexicon-grammatical patterns (Huang, 2014), revealing how words function in real contexts, strengthening students' noticing skills, increasing engagement and motivation (Boulton, 2009), and providing students with resources that

they can use independently (Charles, 2014). Despite these promising benefits, Corpus based learning takes time and is often more suited to "refining usage in context" than learning large amounts of words (Boulton, 2012).

Corpus based learning has recently acquired popularity as a method for vocabulary education in specialized situations (e.g., Graham & Osment, 2013; Önder Özdemir, 2014; Vyatkina, 2016). Corpus based learning can be carried out in a variety of ways, ranging from oriented, printer friendly activities (e.g., Huang, 2014) to internet activities like Compleat Lexical Tutor and sketch engine (Cobb, 1997) or more open-ended, student-led corpus study (e.g., Lee & Swales, 2006), but at its core, Corpus based learning consists of exposing students to language examples from a corpus and asking them to make observations and conclusions based on the corpus data. Corpus based learning necessitates that students interact deeply with words and observe usage patterns.

Because of its strengths in showing word usage in context, Corpus based learning is especially promising for ESP (English for specific purposes) instruction. Effective writing in ESP requires students to know more than just the meaning or usage of a word in general English; students must also have fuller word knowledge, which includes the ability to use collocations, familiarity with lexico-grammar, and an understanding of the meaning created by word choices in the target context. In Medical English, corpus investigation might discover important collocations such as give treatment or provisional diagnosis (Önder Özdemir, 2014). It can also offer light on the ramifications of word choice, such as how the seemingly innocuous phrases guarantee and best are rarely used in engineering documents because they imply an erroneous and potentially dangerous level of assurance (Conrad, 2017).

Students must be exposed to real-world situations in order to gain an understanding of appropriate word usage and choice in their target context. Linguistic examples from that environment, and they want guidance or tools to assist them in acquiring appropriate knowledge to apply in the future. It represents their own writing. Corpus based learning satisfies each of these requirements. It gives students a wealth of context-specific language data. Corpora can be adapted to extremely particular circumstances (for example, see the corpus of wine-tasting comments used by English students).Tourism in Hou, 2014), so that the language is appropriate to the communication aims of the pupils. Concordancers and other corpus tools (or instructor-prepared concordance lines) assist students in distilling material and extracting relevant knowledge.

Although many Corpus based learning applications include a considerable amount of interaction with terminology and requires students to seek knowledge for themselves, only a small amount of vocabulary can be taught in this method (Boulton, 2012). Much of the studies on vocabulary instruction through Corpus based learning have concentrated on 10 or fewer lexical items per week, with some focusing on as low as three or four lexical items per week (e.g., Vyatkina, 2016; Yunus & Awab, 2014). Cobb (1997) is an exception; he successfully taught 20 words each week using controlled online Corpus based learning activities, but this quantity would likely be overwhelming for more hands-on or open-ended Corpus based learning implementation. Teachers who wish to reap the benefits of Corpus based learning must make informed decisions.

Teachers who want to reap the benefits of corpus based activities must make educated word choices. Greater intentionality in word decisions might help you save time, avoid boredom, and feel less frustrated. Finding specific vocabulary to teach in ESP classrooms might be difficult. Language teachers are not usually experts in the fields in which their students intend to utilize

English. It is difficult for instructors to determine which words will be most valuable for students to comprehend, as well as extremely improbable that an instructor will understand the complexities of how the words are employed and the purposes they play in the area. Similarly, professionals in a subject are frequently unprepared to identify vocabulary since they may not be entirely aware of the words they use on a daily basis. These specialists can identify technical language linked to the subject that students need to know, but as Mudraya (2006) has argued, technical words are typically adequately covered alongside content in textbooks and are easier for students to acquire. Instructors must find less evident words that play essential functions in a discipline and grasp how these words function in that context.

Previous studies on vocabulary selection has normally been categorized into two groups: (1) studies of expert corpora to identify words important for achieving communicative goals or participating in a specific context, and (2) studies of learner produced language to identify words that students have not yet mastered. There appears to be minimal research overlap between these two techniques, which is disappointing because both sets of research emphasize an important aspect of assessing kids' vocabulary needs. The current study integrates methodologies from these two fields of inquiry.

The original corpus of research concentrated on determining specialized vocabulary through expert corpora analysis. Researchers identified words based on their frequency of use, distribution across a corpus, and how strongly they are associated with a specific context compared to general use. Coxhead (2000) and Gardner and Davies (2014) both specified criteria for the minimum number of times a term must appear in the academic corpora utilized in their studies when building academic word lists. Their corpora were split by academic specialties or subject areas to control dispersion. In order to exclude words with narrow applicability, they set minimums for the number of occurrences in each corpus subset. Coxhead excluded words that appeared in the General Service List (West, 1953), a list of regularly used words in general English, to rule out words that were prevalent in general use. Gardner and Davis employed a comparative frequency technique, requiring a word to be at least 50% more frequent in their academic corpus than in their general English corpus. This comparative-frequency method is frequently achieved using keyword analysis.

The research on word list creation using expert corpora has established principled methods that shaped the present study; however, as mentioned previously, word lists are typically too long to be covered adequately in a language class. The lists leave instructors still needing to choose a smaller subset of words to focus on, which can be challenging without expert knowledge or additional data. Word lists also fail to consider students' current vocabulary knowledge and needs. To avoid wasting time with in-depth instruction of words that students are already able to use competently, students' vocabulary knowledge must be assessed.

The second branch of research has examined learner language to determine student needs. According to Gilquin, Granger, and Paquot (2007), analyzing learner corpora is crucial but often overlooked in pedagogical research. A corpus of learner-produced language allows researchers to assess learners' proficiency with certain components of language, such as vocabulary usage.

In evaluating learners' linguistic demands, specialist comparisons have already been frequently utilized (e.g., Conrad, 2017; Flowerdew, 2003; Hartig & Lu, 2014). Because language teachers frequently lacked enough ability to completely judge the efficiency of the students'

language in a given situation, they can compare student language samples to expert language samples, identify disparities, and address these gaps through instruction.

In short, it was found out that while analyzing learner language is crucial, selecting vocabulary merely based on which words students do not use effectively does not account for which terms are important in the ESP context. Teachers may spend a lot of time teaching terms that aren't very useful while ignoring more important words. A methodology that incorporates findings from both areas of research will assist ensure that words are both significant in the ESP setting and have proved difficult for ESP students.

REFERENCES

1. Anthony, L. (2015). AntConc (version 3.3.4) [Computer Software]. Tokyo, Japan: Waseda University http://www.laurenceanthony.net.

2. Boulton, A. (2009). Data-driven learning: Reasonable fears and rational reassurance. Indian Journal of Applied Linguistics, 35, 81-106. https://hal.archivesouvertes.fr/hal-00326990v1.

3. Boulton, A. (2012). Corpus consultation for ESP: A review of empirical research. In A. Boulton, S. Carter-Thomas, & E. Rowley-Jolivet (Eds.), Corpus-informed research and learning in ESP: Issues and applications (pp. 261-291). Amsterdam: John Benjamins. https://doi.org/10.1075/scl.52.11bou.

4. Charles, M. (2014). Getting the corpus habit: EAP students' long-term use of personal corpora. English for Specific Purposes, 35, 30-40.

5. Cobb, T. (1997). Is there any measurable learning from hands-on concordancing? System, 25, 301-315. https://doi.org/10.1016/S0346-251X(97)00024-9.

6. Conrad, S. (2015). Register variation. In D. Biber, & R. Reppen (Eds.), The Cambridge handbook of English corpus linguistics (pp. 309-329). Cambridge: Cambridge University Press.

7. Conrad, S. (2017). A comparison of practitioner and student writing in civil engineering. Journal of Engineering Education, 106, 191-217. https://doi.org/10. 1002/jee.20161.

8. Conrad, S. (2018). The use of passives and impersonal style in civil engineering writing. Journal of Business and Technical Communication, 32, 38-76. https://doi.org/10.1177/1050651917729864. Coxhead, A. (2000). A new academic word list. Tesol Quarterly, 34, 213-238.

9. Coxhead, A. (2018). Vocabulary and English for specific purposes research: Quantitative and qualitative perspectives. Routledgehttps://doi.org/10.4324/ 9781315146478.

10. Coxhead, A., & Demecheleer, M. (2018). Investigating the technical vocabulary of Plumbing. English for Specific Purposes, 51, 84-97. https://doi.org/10.1016/j. esp.2018.03.006.

11. Davies, M. (2008). The corpus of contemporary American English: 450 million words, 1990–present. http://corpus.byu.edu/coca/.

12. Donesch-Jezo, E. (2010). Corpus concordancing in teaching academic discourse writing to medical students. In Proceedings of the international conference ICT for language learning.

13. Flowerdew, L. (2003). A combined corpus and systemic-functional analysis of the problem-solution pattern in a student and professional corpus of technical writing. Tesol Quarterly, 37, 489-511. https://doi.org/10.2307/3588401.

14. Gardner, D., & Davies, M. (2014). A new academic vocabulary list. Applied Linguistics, 35, 305-327. https://doi.org/10.1093/applin/amt015. Gaskell, D., & Cobb, T. (2004).

15. Graham, D., & Osment, C. (2013). Interactive web-based learning of corpus-generated phrases. AsiaCALL Online Journal, 9, A1-A13. http://asiacall.info/acoj/ acoj-2014/.

16. Hartig, A. J., & Lu, X. (2014). Plain English and legal writing: Comparing expert and novice writers. English for Specific Purposes, 33, 87-96. https://doi.org/10. 1016/j.esp.2013.09.001.

17. Hou, H. I. (2014). Teaching specialized vocabulary by integrating a corpus-based approach: Implications for ESP course design at the university level. English Language Teaching, 7, 26-37.

18. Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. Reading in a Foreign Language, 13, 403-430.