Mohylevska V.A.,
Lecturer, the Department of Foreign Languages
National A cademy of Internal A ffairs

PEDAGOGICAL APPLICATIONS OF CORPORA

Although corpus linguistics has become quite influential in English language teaching in recent years, the actual use of corpora (large, searchable collections of real language, electronic texts) by practicing teachers still seems to be fairly limited. In their introduction to data-driven learning (DDL) in the classroom, Gilquin and Granger declare, for example, "one reason for not doing DDL might simply be that the teacher does not know enough about corpora and the possibility of using corpora in the classroom" [6, p.366].

According to Reppen, a corpus is "a large, principled collection of naturally occurring texts (written or spoken) stored electronically" [11, p.2]. Good reasons to use such corpora (plural) are that they help identify both "linguistic and situational co-occurrence patterns" and offer "a ready resource of natural, authentic, texts for language learning," in addition to removing some of the guessing of native speakers of languages, whose intuitions are "often ill-informed" [11, p.4]. Thus corpora and DDL are of great value to both native- and non-native-English speaking teachers who question their own intuitions of English language use. A corpus linguistics approach to teaching and learning is largely frequency-based (using a cost-benefit view - see Barker, 2007) with an emphasis on authentic, real-life examples that are to be examined in context. While the machines of years ago may have caused

frustration for most people, corpus linguistics has flourished in the last 20 years with technological advances now providing more userfriendly access to corpora, especially on the Internet. Corpus linguistics and DDL require a corpus, a computer, and concordancing software to analyze the corpus (usually built-in to online corpora sites), plus specific questions to research and answer, and a process for studying the results.

While corpus linguistics is used in many linguistic areas [9], language learning and teaching is a primary one, especially for ESL/EFL. The applications of corpora in teaching are myriad, including vocabulary learning and teaching, phraseology, register, English for academic and specific purposes, and materials design. Key classroom approaches for DDL with corpora include using word lists, concordance lines for examples of real language use, texts tagged with parts of speech, and examples of register use in language [11].

An important concept to note is that DDL requires that teachers and students take on new roles in class. As Gilquin and Granger clearly state, for teachers it is important to have adequate knowledge about corpora and some of the potential options for using them in the classroom, but DDL also implies something of "a less central role...than in traditional teaching," because the teacher facilitates learning and helps learners arrive at answers to their questions [6, p. 366]. As a result, teachers cede the role of expertise to the corpus, "take risks, and agree to 'let go' and let the student take pride of place in the classroom" [6, p. 367]. Similarly for students, DDL requires some training, as well as guidance in developing questions, determining what resources are available, and the means for understanding and evaluating information that corpora provide.

Perhaps the first way that teachers can use corpora for ESL/EFL teaching is in creating and using word frequency lists. In addition to using already established English word lists and the new Academic Vocabulary List [5], for example, with the online Corpus of Contemporary American English (or COCA), teachers can prepare specific lists of English words and then use them either in the preparation of lessons and materials for particular classes or for students to study. Beyond individual words, however, teachers can also use corpora to help students find collocations and teach lexical chunks [8]. Like Kathpalia and Ling, collocations can help students develop their proficiency in English, and that using collocations helps learners go beyond individual words and enables them to see that language tends to work in larger phrases or chunks [7]. An excellent resource for much more detail on such approaches to vocabulary teaching using corpora is McCarthy, O'Keefe, and Walsh's (2010) helpful book [10].

Yet corpora are not only useful for vocabulary teaching. Another way that corpora are valuable is in how they help in examining grammar patterns in English. As alluded to earlier, corpora and concordance lines with real English examples can help both teachers and their students consider specific words and phrases in spoken or written contexts, where they can study preposition use, verbs and complements, and adjective + noun and/or other combinations that will assist English language learners and teachers consider their own speaking and writing. A final task that should be mentioned is that students and teachers may use corpora to compare students' English use (spoken or written vocabulary, grammar, phraseology, etc.) with that of native English speakers (e.g., using the COCA) or other ESL/EFL students (e.g., with the English as a lingua franca Vienna-Oxford International Corpus of English, or VOICE, corpus).

In conclusion, corpora will not only revolutionize the teaching of subjects such as grammar in the 21st century [2], they will also fundamentally change the ways we approach language education, including both what is taught and how it is taught. As Gavioli and Aston argue, corpora should not only be viewed as resources which help teachers to decide what to teach, they should also be viewed as resources from which learners may learn directly [4].

References

1. Aston, G. (ed.) (2001) Learning with Corpora. Houston, TX: Athelstan.

2. Conrad, S. (2000) 'Will corpus linguistics revolutionize grammar teaching in the 21st century?'. TESOL Quarterly 34: 548-60.

3. Barker, D. (2007). A personalized approach to analyzing 'cost' and 'benefit' in vocabulary selection. System, 35,   523-533. doi:10.1016/j.system.2007.09.001

4. Gavioli, L. and Aston, G. (2001) 'Enriching reality: language corpora in language pedagogy'. ELT Journal 55/3: 238-246.

5. Gardner, D., & Davies, M. (2014). A new academic vocabulary list. Applied Linguistics, 35, 305-327. doi:10.1093/applin/amt015

6. Gilquin, G., & Granger, S. (2010). How can data-driven learning be used in language teaching? In A. O'Keefe & M. McCarthy (Eds.), The Routledge handbook of corpus linguistics (pp. 359-370). London: Routledge.

7. Kathpalia, S. S., & Ling, K. S. (2008). Is collocation the way to language proficiency? TESL Reporter, 41(1), 1-11.

8. Lessard-Clouston, M. (2013). Teaching vocabulary. Alexandria, VA: TESOL International Association.

9. McCarthy, M., & O'Keefe, A. (2010). Historical perspective: What are corpora and how have they evolved? In A. O'Keefe & M. McCarthy (Eds.), The Routledge handbook of corpus linguistics (pp. 3-13). London: Routledge.

10. McCarthy, M., O'Keefe, A., & Walsh, S. (2010). Vocabulary matrix: Understand- ing, learning, teaching. Andover, U.K.: Heinle Cengage.

11. Reppen, R. (2010). Using corpora in the language classroom. Cambridge: Cambridge University Press.