Corpus linguistics software: enhancing learners and educators' experience

The use of corpus linguistics (CL) in second language learning and teaching is now well-established in the field of language acquisition (Granger 2009) and is said to be one of the most prominent applications of CL (Bennett 2010). In this field, CL methods are mainly used as a learning tool and as a reference source. A common practice when CL is a learning tool is the use of hands-on activities, in which students work directly with authentic data (O'Keeffe, McCarthy, and Carter 2007), characterised by Johns (1991) as data-driven learning. CL has also been used as a reference to inform material production and to describe learners' language (Meunier 2010). However, there is still resistance to CL from both learners and educators (Romer 2006). Studies have reported that users have difficulties in formulating queries, understanding the tools functioning and interpreting the results (Yoon and Hirvela 2004). Presenting data through visual aids has been shown to facilitate information insight and to enhance the user experience (Few 2004). This presentation outlines some recent work on improving the experience of language learners and educators with CL software, through the development and implementation of a new data visualization. This consisted of three steps: (a) identifying the target audience and understanding their needs; (b) development and implementation of the visualization; and (c) user assessment of the newly developed tool. User needs were assessed via (a) literature investigation into papers reporting corpus-based methods and (b) a contextual design approach (Beyer and Holtzblatt 1998), allowing observation of how users interact with CL software in their own environment. These observations were then used as a starting point for the development of a new visualization for dispersion. Key issues for a successful data visualization, such as its functionality, aesthetics and accuracy (Cairo 2016) were also considered. The new functionality was implemented in CQPweb (Hardie 2012), an open-source piece of software for corpus linguistic analysis.

References

Bennett, G. (2010). Using Corpora in the Language Learning Classroom. Ann Arbor, MI: University of Michigan Press. http://doi.org/10.3998/mpub.371534

Beyer, H., & Holtzblatt, K. (1998). Contextual design: defining customer-centered systems. Morgan Kaufmann.

Cairo, A. (2016). The truthful art: data, charts and maps for communication. New Riders.

Few, S. (2004). Show me the numbers: designing tables and graphs to enlighten. Analytic Press.

Granger, Sylviane. (2009) The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation. In: Aijmer, K., Corpora and Language Teaching, Benjamins: Amsterdam and Philadelphia, p. 13-32.

Hardie, A. (2012). CQPweb — combining power, flexibility and usability in a corpus analysis tool. International Journal of Corpus Linguistics, 17(3), 380-409. http://doi.org/10.1075/ijcl.17.3.04har

Johns, T. (1991). Should you be persuaded - Two Samples of data-driven learning materials. In T. Johns & P. King (Eds.), Classroom Concordancing (Vol. 4, pp. 1–16).

Meunier, F. (2010). Learner Corpora and English Language Teaching: Chekup Time. Anglistik: International Journal of English Studies, 21(1), 209–220.

O'Keeffe, A., McCarthy, M., & Carter, R. (2007). From Corpus to Classroom: language use and language teaching. New York: Cambridge University Press.
Römer, U. (2008). Corpora and language teaching. Corpus Linguistics Volume 1, 112–131. http://doi.org/10.1016/j.esp.2003.12.001
Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. Journal of Second Language Writing, 13(4), 257–283. http://doi.org/10.1016/j.jslw.2004.06.00