

Victoria Kamasa, PhD
Adam Mickiewicz University in Poznań
vkamasa@amu.edu.pl

Corpus Linguistics for Critical Discourse Analysis: What can we do better?

“More recently, it seems that use of CL techniques is becoming increasingly popular in critical approaches to discourse analysis” (Baker et al., 2008: 274-275). Both the methods used and the research problems are very diversified. Methods vary from bare analysis of frequencies (Alcaraz-Ariza, 2002; Mautner, 2007) through analysis of collocations (Don et al., 2010; Lischinsky, 2011; Mautner, 2007; Weninger, 2010) or keywords (Bachmann, 2011; Don et al., 2010; Lukac, 2011; Weninger, 2010) to the analysis of key semantic domains (Prentice, 2010). Whereas research subjects range from national identity issues (Don et al., 2010, Freake et al., 2010, Prentice, 2010), through different social problems (Kirejczyk, 1999, Lukac, 2011) to social construction of businesswomen (Koller, 2004) or economic crisis (Lischinsky, 2011). Despite the growing interest, the variety of methods and the diversity of research subject, the steadily growing body of corpus-supported CDA studies has not been critically reviewed in order to identify most vulnerable points of the research practice and suggest some improvements. We attempt to fill this gap in the presented paper.

Our review is based on over 30 papers in which authors declared to use some technique of corpus linguistics for some form of Critical Discourse Analysis (CDA/CL). We analyze the methods used by paper's authors as well as the results provided by those methods in order to propose some points for improvement. The analysis concentrates on two major issues:

- the relation between used methods, received results and postulated conclusions e.g. the degree to which the results support the conclusion;
- the relation of the research practice to the benefits of using CL for CDA. Some of those benefits have been pointed out in the literature (Mautner, 2009; Marko, 2008; Hardt-Mautner, 1995; Lischinsky, 2011).

As a result we describe seven main points in which such CDA/CL analysis may be improved. The first one concerns corpus design: we show how decisions taken on this stage may limit the results. Secondly, we refer to the usage of statistics and demonstrate how some results can be improved by extending the number of issues statistics is used for. Moreover, we point to some inconsequence which may take place during the research process concerning both sticking to rules declared by the author and paying attention to numbers such as word frequencies. Another problem we discuss can be called “mind-reading problem”: while the results concern the proprieties of text, the conclusions regard cognitive states of the users. We also refer to the so-called cherry-picking problem (Breeze 2011), which is postulated to be solved by the usage of CL techniques (Lischinsky, 2011). Finally, we discuss briefly the role of researcher's intuition and show some stages of CDA/CL research in which the intuition continues to play crucial role. For every of these points we present examples form research practice.

As a conclusion we offer some suggestions of improvements which may be beneficial for the growing community of CDA/CL practitioners in order to fully use the potential of corpus linguistics' tool to reveal socially important discursive constructions.

References:

- Alcaraz-Ariza, M. Á. (2002). Evaluation in English-Medium Medical Book Reviews. *International Journal of English Studies*, (vol. 2 (1)), 137–153.
- Bachmann, I. (2011). Civil partnership – “gay marriage in all but name”: a corpus-driven analysis of discourses of same-sex relationships in the UK Parliament. *Corpora*, (Vol. 6 (1)), 77–105.

- Baker, P. (2006). *Using corpora in discourse analysis. Continuum discourse series*. London: Continuum.
- Baker, P., Gabrielatos, C., Khosravinik, M., Krzyzanowski, M., McEnery, T., Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK Press. *Discourse & Society*, 19(3), 273–306.
- Breeze, R. (2011). Critical Discourse Analysis and Its Critics. *Pragmatics*, (21:4), 493–525.
- Degano, C. (2007). Dissociation and Presupposition in Discourse: A Corpus Study. *Argumentation*, 21(4), 361–378.
- Don, Z. M., Knowles, G., & Fatt, C. K. (2010). Nationhood and Malaysian identity: a corpus-based approach. *Text & Talk - An Interdisciplinary Journal of Language, Discourse & Communication Studies*, 30(3), 267–287.
- Freake, R., Gentil, G., & Sheyholislami, J. (2010). A bilingual corpus-assisted discourse study of the construction of nationhood and belonging in Quebec. *Discourse & Society*, 22(1), 21–47.
- Hardt-Mautner, G. (1995). 'Only Connect.': *Critical Discourse Analysis and Corpus Linguistics*. Retrieved from <http://ucrel.lancs.ac.uk/papers/techpaper/vol6.pdf>
- Kirejczyk, M. (1999). Parliamentary Cultures and Human Embryos: The Dutch and British Debates Compared. *Social Studies of Science*, 29(6), 889–912.
- Koller, V. (2004). Businesswomen and war metaphors: 'Possessive, jealous and pugnacious'? *Journal of Sociolinguistics*, (8/1), 3–22.
- Lischinsky, A. (2011). In times of crisis: a corpus approach to the construction of the global financial crisis in annual reports. *Critical Discourse Studies*, 8(3), 153–168.
- Lukac, M. (2011). Down to the bone: A corpus-based critical discourse analysis of pro-eating disorder blogs. *Jezikoslovlje*, (12.2), 187–209.
- Marko, G. (2008). *Penetrating language: A critical discourse analysis of pornography*. Tübingen: Narr.
- Mautner, G. (2007). Mining large corpora for social information: The case of elderly. *Language in Society*, 36(01), 51–72.
- Mautner, G. (2009). Checks and balances: how corpus linguistics can contribute to CDA. In R. Wodak & M. Meyer (Eds.), *Methods of critical discourse analysis* (pp. 122–144). London: SAGE.
- Prentice, S. (2010). Using automated semantic tagging in Critical Discourse Analysis: A case study on Scottish independence from a Scottish nationalist perspective. *Discourse & Society*, 21(4), 405–437.
- Weninger, C. (2010). The lexico-grammar of partnerships: corpus patterns of facilitated agency. *Text & Talk - An Interdisciplinary Journal of Language, Discourse & Communication Studies*, 30(5), 591–613.
- Yasin, M. S. M., Hamid, B. A., Keong, Y. C., Othman, Z., & Jaludin, A. (2012). Linguistic Sexism In Qatari Primary Mathematics Textbooks. *GEMA Online Journal of Language Studies*, (Volume 12(1)), 53–68.