Abs-65

Charlotte Taylor, University of Portsmouth

Searching for similarity: The representation of boy/s and girl/s in the UK press in 1993, 2005, 2010

This aim of this paper is to raise the methodological importance of searching for similarity and stasis as well as difference and change in corpus-assisted discourse analysis. To date, most research in corpus approaches to discourse studies has largely focussed on difference, which carries the risk of neglecting the similarities (as noted in Baker 2006: 182) and leading the researcher to finding the changes or differences that s/he set out to look for.

This tendency to focus on difference in comparative work is frequently reflected in the corpus linguistic software and in the paper I look at what automated searches for similarity are possible and how these may be employed, for example Sketch Engine's Sketch Difference also includes collocates that are shared between two lexical items.

One notion which has been developed to cover the gap is consistent collocates (c-collocates), introduced in the Lancaster RASIM project to describe the lexical items which collocated with refugees/asylumseekers/immigrants/migrants in at least seven out of the ten annual subcorpora (described in Gabrielatos and Baker 2008). Baker (2009) further addressed the issue by introducing the concept of lockwords ie items that have a similar frequency across corpora as a parallel to keywords, that is those items which are significantly different between corpora. Another source for approaches to similarity comes from the field of forensic linguistics, which has focussed on similarity far more extensively, particularly in the area of plagiarism studies and detection.

The paper forms part of a larger project which uses the methodology of Modern Diachronic Corpus Assisted Discourse studies (MD-CADS, see the special issue of Corpora, 2010) to analyse the representation of GIRL and BOY in the UK press and investigates whether the definitions, characterisations and contexts of presentation have changed or remained stable over the last two decades. In doing so, it builds in particular on previous corpus studies of gender terms such as Holmes & Sigley (2000), Sigley & Holmes (2002), Pearce (2008), Caldas-Couthard & Moon (2010) and Baker (2010). More specifically, I analyse the occurrences of GIRL and BOY in the entire output of the Times, Guardian and Telegraph from 1993 and 2005, and then extend that analysis to a smaller corpus of the same newspapers from 2010. In this presentation, I will be limiting discussion to the analysis of the constants and similarities both in the time period, for instance the continued usage of GIRL to refer to women, and between the search items, for instance the way in which BOY is increasingly used with GIRL collocates.

References

Baker, P. (2006). *Using Corpora in Discourse Analysis*. London: Continuum.

Baker, P. (2009) 'The BE06 Corpus of British English and recent language change.' *International Journal of Corpus Linguistics*. 14:3 312-337.

Baker, P. (2010) 'Will *Ms* ever be as frequent as *Mr*? A corpus-based comparison of gendered terms across four diachronic corpora of British English.' *Gender and Language* 4.1: 125-129.

Caldas-Coulthard, R. & R. Moon. '"curvy, hunky, kinky": Using corpora as tools for critical analysis'. *Discourse & Society* 21(2): 99-133.

Gabrielatos, C. & P. Baker (2008). 'Fleeing, sneaking, flooding: a corpus analysis of discursive constructions of refugees and asylum seekers in the UK Press 1996-2005)' *Journal of English Linguistics* 36:1 pp. 5-38.

Holmes, J. and Sigley, R. (2001) 'What's a Word like *Girl* Doing in a Place Like This?', in A. Smith and P. Peters (eds.), *New Frontiers of Corpus Linguistics*, pp. 247–63. Amsterdam: Rodopi.

Pearce, M. (2008) 'Investigating the Collocational Behaviour of MAN and WOMAN in the BNC using Sketch Engine', *Corpora* 3(1): 1–29.

Sigley, R. and Holmes, J. (2002) 'Looking at *Girls* in Corpora of English', *Journal of English Linguistics* 30(2): 138–57.