Paris EngCorpora15 – Abstract

Cesare Zanca

Dipartimento di Filologia e Critica delle Letterature Antiche e Moderne

Università degli Studi di Siena

Italy

Cesare.Zanca@unisi.it

Diversity in diverse Englishes: why should the students care?

This paper has a twofold approach to the investigation of large corpora to reveal facts about language use. On one side I consider the perspective of the researcher who collects a corpus, learns how to use dedicated software, engages with the technicalities of statistical and quantitative analysis and so on. On the other, I take the language students' perspective and their need to deal with specific and urgent linguistic problems for their writing or reading tasks.

The discussion is based on the exploration of the three large sub-corpora that make up the SiBol/Port diachronic corpus of British Newspapers collected "from different, but contemporary, periods in time" and "designed and compiled to be as similar as possible in order to eliminate any maverick variables" (Partington 2010 p.85)[1]. These corpora, collected in 1993, 2005 and 2010 will be compared to the brand new section of the Sibol-Port corpus: the 2013 sub-corpus, which is considerably larger and includes newspapers published in English in different parts of the world: UK, USA, India, Nigeria, Arab Emirates, China and Egypt.

Following the Modern Diachronic Computer Assisted Discourse Studies (MdCADS) methodology (Partington 2010: 85), these corpora will be used to conduct quantitative and qualitative investigations of specific discourse features and "the broader societal and political framework in which such discourse is embedded" (Schäffner 1997: 1), shunting back and forward from numbers and statistical information to qualitative findings and evaluations to identify "the relationship between instance and system, between the typical and the exceptional, between signal and noise" (Partington 2004: 12).

The integration of quantitative and qualitative considerations will be pointed out as a potential way to deal with one of the questions of the Engcorpora 2015 conference call for papers, namely the reliability of a corpus in the hypothesis-building process:

The particular discourse feature chosen for the conference is the idea of *diversity* and how it is expressed in different times, different contexts and different varieties of Newspaper English in the world. Apart from the use of the word *diversity* itself - which reveals a substantial shift in associated discourses[2] - other ways of expressing diversity are explored: from the use of the prefix non- followed by a social group identifier (eg.

---

[1] For a complete description see http://www3.lingue.unibo.it/blog/clb/?page_id=8
[2] An initial analysis of the word *diversity* in the Sibol-Port Corpus has been conducted and discussed in the 2013 and 2014 Languaging Diversity Conferences in Naples and Catania (Zanca 2014 forthcoming).

non-white non-Christian or *non-British*) to discourses related to the expression of *otherness* in terms of nationalities, values and social behavior.

The assumption is that the investigation of single lexical items or clusters in very large corpora can shed light on facts about how language is used and about temporal or spatial change in its use that would otherwise remain merely conjectural.

But is this kind of work possible for or felt as relevant by the average learner in a linguistic or language course? Following previous studies (Zanca 2011, 2013), a final part of the paper deals with this aspect, indicating a possible way to integrate corpus linguistics methodology in the average classroom.

Bibliography

Partington, Alan S. / Morley, John / Haarman, Louann (eds) 2004.*Corpora and Discourse*, Bern: Peter Lang.

Partington, A. 2010. '*Modern Diachronic Corpus-Assisted Discourse Studies (MD-CADS) on UK newspapers: an overview of the project'*, Corpora 5 (2), pp. 83–108.

Schäffner, C (ed.) 1997, *Analysing political speeches*. Current issues in language and society, vol. 3, Multilingual matters, Clevedon (UK).

Zanca C.(2011) *Developing translation strategies and cultural awareness using corpora and the web*. Tralogy [En ligne], Session 3 - Training translators / La formation du traducteur. Paris. URL. http://lodel.irevues.inist.fr/tralogy/index.php?id=116.

Zanca C.(2013) *Online Learning and Data Driven Learning in translation training and language teaching*. In C. Argondizzo (ed) Creativity and innovation in language education. LINGUISTIC INSIGHTS - Peter Lang AG. Bern.