The future of the International Corpus of English (ICE) project – New challenges, new developments

Robert Fuchs, Ulrike Gut (University of Münster) & Gerald Nelson (Chinese University of Hong Kong)

More than 20 years have passed since the late Sydney Greenbaum laid the foundations of the International Corpus of English (ICE) project (Greenbaum 1991). The project, with its aim to facilitate the comparison of the national standard and standardising varieties of English around the globe through the compilation of comparable 1 million word corpora, has been a resounding success – if numbers are anything to go by: the compilation of 13 subcorpora has been completed to date, with 13 more in the works, and Google Scholar indexes more than 3,200 publications making reference to the ICE project and its subcorpora.

Since 1991, the theoretical and practical contexts of research on varieties of English have changed, and with it the demands that researchers might make on a corpus that allows them to compare the national varieties of English used in countries around the world. Progress in computer technology allows us to use corpus data for analyses that might have been hard to envisage in the 1990s, such as the analysis of pragmatic and prosodic features (Kallen and Kirk 2012), as well as time-aligned annotation with tools such as ELAN (Brugman and Russel 2004, Wunder et al. 2010), which allows corpus users access to the original recordings for phonetic or prosodic analyses. Progress in corpus compilation theory makes corpus creation faster and more reliable (Voormann and Gut 2008), and new corpus compilation techniques permit the collection of data from internet sources to create mega-corpora, such as the Corpus of Global Web-based English (GloWbE, Davies and Fuchs 2015). In addition to technical developments, the evolution of varieties of English also prompts researchers to ask new questions or seek new answers to old ones. How should the ICE project, for example, treat national varieties that cannot easily be classified as English as a Second and English or English as a Foreign Language, such as English in Cyprus and the Netherlands (Buschfeld 2013, Edwards 2014)?

In this context, the workshop will provide a venue for discussing current issues in the compilation and use of (subcorpora of) the International Corpus of English. We welcome submissions for full papers and work-in-progress reports on

- how the (meta)data available in existing ICE subcorpora can be used to address new questions,
- how new corpus compilation techniques can make corpus collection more efficient or reliable, and permit the analysis of more language features than has been possible until now,
- how the ICE project can address recent change in how English is used around the world (such as across the ESL-EFL continuum and on the internet),

as well as related topics. Abstracts (of up to 500 words, excl. references) should be submitted to Robert Fuchs (robert.fuchs@uni-muenster.de) by 15th February 2015. Notification of acceptance will be sent out by 28th February 2015. The workshop will take place on 27th May 2015 as a pre-conference workshop of ICAME 2015 in Trier, Germany; for more information, see http://www.uni-trier.de/index.php?id=52275.
References


