

<u>Workshop ID :</u>	42
<u>Workshop Duration :</u>	Workshop - 1 Day
<u>Workshop Title :</u>	Corpus linguistics and the challenges of investigating language contact
<u>Workshop Leader :</u>	Bertus Van Rooy , North-West University Haidee Kruger , Macquarie University

Corpus linguistics as methodology has been very important to linguistic research across a range of fields as it expanded in the last half century. In the early years, the prototypical corpus was a balanced monolingual corpus that aimed to represent (a standard variety of) a language across different written and spoken genres, as exemplified by the Brown and LOB corpus (McEnery & Wilson 2001). Over time, the kinds of corpora extended to more specialised registers and applications, such as corpora for lexicography, historical, translation and learner language corpora, corpora of learner language and mega-corpora that are compiled from the web or from large data-scanning projects (McEnery & Wilson 2001), often not in the first instance aimed at linguistic research, but a much wider audience. The field has seen advances in the sophistication of methods and tools for data analysis, as well as increased sophistication in the use of statistics for finding and evaluating patterns in the data, sometimes driven by the availability of new corpora, sometimes by research questions that previous methods were unable to answer.

A promising, but still niche, area within corpus linguistics is the study of the effect of language contact and bilingualism. Some communities of scholarship already exist in fields such as learner corpus studies (Granger *et al.* 2015), the study of New Englishes, especially aligned with the International Corpus of English project (Hundt & Gut 2012), English as a Lingua Franca (Seidlhofer 2012), and corpus translation studies (Laviosa 2002). These areas draw on corpus linguistic methods, but also conceptual and methodological tools from other disciplines, such as applied linguistics, second language acquisition, sociolinguistics, translation studies and world Englishes.

Methodological and conceptual plurality and diversity can be a virtue and open up new and unexpected angles and vistas. However, a question that should be asked is whether these various perspectives converge on consistent methods and findings that can be justified in rigorous ways. Moreover, research on different kinds of corpora that have their origins in situations of language contact are largely conducted in separate fields of inquiry, which begs the question of what they stand to gain from considering insights that cut across the disciplinary divides.

The purpose of this workshop is to consider various research traditions that use corpora of language produced in contexts of language contact, to take stock of the achievements and to reflect on the gains and limits of the various traditions of scholarship that combine with corpus linguistics. At the same time, the ways in which current methods within corpus linguistics are appropriated and adjusted are assessed, with a view to developing the most appropriate and rewarding ways in which corpus methods can contribute to insight in contact-influenced varieties of languages.