Learner corpora and theory-building in second language acquisition research

Prof. Rosamond Mitchell
University of Southampton

Abstract

The creation and analysis of learner corpora is a well established field in applied linguistics, which provides important insights into the nature of learner language (Granger et al., 2015). Learner corpora are large scale electronic collections of naturalistic or semi-naturalistic texts produced by language learners, which can be automatically analysed to detect a variety of linguistic characteristics. Most commonly, through “contrastive interlanguage analysis” (Granger, 1996), learner corpus researchers analyse lexical, grammatical and phraseological patterns in learner productions, and make comparisons either with native speaker patterns, or between groups of learners from different L1 backgrounds, or at different proficiency levels. This research plays an increasingly important role in informing language assessments and the production of pedagogic materials.

Numerous commentators have argued for the potential of learner corpora also to contribute to second language acquisition (SLA) research (Granger, 2009; Myles, 2015). However, SLA research is typically concerned to test precise hypotheses about the drivers of language development, connected to a particular theoretical position, and consequently SLA researchers have commonly preferred to work with specially elicited, experimental data. Thus, the two fields have been slow to come together (Hasko, 2013; McEnery et al., 2019).

In this talk we will consider recent examples of SLA research using learner corpora which set out to test hypotheses deriving from a number of different theoretical positions, including generativist, usage-based skill acquisition and language socialization theories. We will discuss the characteristics of the corpora which SLA researchers found most useful, and how they complemented the use of experimental data, and we will draw conclusions regarding learner corpus design for interdisciplinary research purposes. Finally, we will discuss the implications of the “open science” movement (Marsden, 2019) for learner corpus design, management, promotion and wider use in SLA research.

Lecture in English
References


Short bio statement

Professor Rosamond Mitchell is an Emeritus Professor of Applied Linguistics in Modern Languages and Linguistics at the University of Southampton where she has been working since the 1980s. After many years in the School of Education, she joined Modern Languages and Linguistics in 2004, where she was founding director of the Centre for Applied Language Research.
She has two major areas of research expertise: second language acquisition, and foreign language classroom learning and teaching. In second language acquisition, she has particular interests in the development of corpus based approaches for the study of learner development, and has worked extensively on the development of learner corpora in French and Spanish. In foreign language education, she has particular research interests in: classroom interaction and its role in language learning; critical thinking in language education; foreign language pedagogy for early learners; and informal learning by advanced learners during residence abroad.