DOTTORATO IN STUDI LETTERARI LINGUISTICI E COMPARATI

PhD PROGRAMME IN LITERARY LINGUISTIC AND COMPARATIVE STUDIES

Research proposal:

Corpus linguistics and digital environments in English: from academic research to English language learning, testing and assessment.

Maria De Santo

L/LIN-12 Lingua e traduzione – Lingua inglese

Abstract

Corpus-based research is a significant and constantly expanding field of Linguistic Studies, offering diverse complex and intersected perspectives of language analysis. Remarkable advances have distinguished the progress of corpus linguistics over the years, from theoretical studies to applied methodologies, from corpus design to corpus exploitation tools, from monolingual to multilingual and specialist corpora (Facchinetti, 2005). Investigating corpora - collections of authentic texts stored electronically that can be used to discover information about language - gives insights into diverse languages from both a linguistic and a pedagogical viewpoint. Indeed, "corpus linguistics is not an end in itself but is one source of evidence for improving descriptions of the structure and use of languages, and for various applications, including the processing of natural language by machine and understanding how to learn or teach a language" (Kennedy, 1998).

The value of corpora in Second Language Acquisition in terms of both theoretical discussion and practical application is considerable. Nevertheless, whereas native corpora and corpus linguistic tools and methodologies have been used for a long time in the development of pedagogic materials, learner corpora - collections of texts produced by learners of a language – "are only just beginning to impact the field of language testing and assessment (LTA)." (Callies and Götz, 2015:1).

Learner corpus research has created a significant link between the two "previously disparate field of corpus linguistics and foreign/second language research" (Granger, 2002:1). This evolving

connection is central to this research project, whose main objective is to investigate the potential of corpus-based techniques to analyse learners' proficiency in L2 productions and how the results of this study can be applied to develop both teaching methodologies and assessment strategies that can enhance learners' linguistic and metacognitive skills.

The starting point of the research is the construction of an academic learner corpus, based on text productions of L2 learners of English. Data will be collected from e-tests, forums, chats and other Web 2.0 communicative and collaborative tools integrated into the University of Naples L'Orientale Moodle-based e-learning platform. The resulting corpus will, indeed, consist of a blend of text, images and sounds that will be transcribed, annotated and analysed with the help of software tools for corpus linguistic analysis. Results, obtained both with quantitative and qualitative methods of analysis, will be used to design teaching and learning materials tailored to a definite academic context as well as to project e-testing and e-assessment tools more reliable and closer to students' needs.

The theoretical framework of the present research proposal merges approaches and methodologies from the fields of corpus linguistics, second language acquisition and language testing, while fostering the integration of digital resources and environments, such as Learning Management System (LMS) platforms and Web 2.0 tools, with the methods of corpus-based linguistic analysis. Since their advent on educational sceneries, LMS-based digital environments have offered the opportunity to re-shape learning experiences, challenging educators and researchers with new opportunities to design, develop and experiment tools and resources for online teaching, testing and assessment. The University e-learning platform, based on Moodle one of the most widely used LMS at university level, will provide the environment and the tools for data collection, e-testing experimentation and the dissemination of the results of the research.

The double perspective that distinguishes the research proposal will also determine the nature of its impact and expected outcomes: the first, a readily available linguistic research product, the university Corpus of L2 Learners of English, the second a structured format for a language e-testing system tailored on the university specific educational context.

Furthermore, although the proposal refers to research, methodologies and procedures relevant to English linguistics and English Language Teaching (ELT), it addresses issues that can be also relevant to other languages taught at University level.

Keywords: Corpus linguistics, learner corpora, LMS, online environments, e-learning, e-testing, e-assessment