

**2nd Guangdong University of Foreign Studies Forum on Applied Linguistics
7-9 December 2012.**

Name: John Blake
Institute: Japan Advanced Institute of Science and Technology
Email: johnb@jaist.ac.jp

Research abstracts: A diachronic systemic functional analysis

English has become the *de facto* language for scientific articles (Simionescu & Simion, 2004). Given the high lexical density of research abstracts (Halliday & Martin, 1993) and the assertion by Swales (1990) that mastery of drafting abstracts is the *rite de passage* of entry into the discourse community, non-native English speakers (NNES) appear to be put at a linguistic disadvantage. This study is a diachronic investigation into the linguistic features of research abstracts produced over a career spanning 25 years for one eminent NNES engineering professor with numerous publications.

A small corpus of approximately 25,000 words from 134 scientific research abstracts (125 in English and 9 in Japanese) was collected from publications appearing in prestigious IEEE journals and conference proceedings in which the professor was an author or co-author. The moves and steps in each abstract were manually identified. The corpus was tagged using the UAM Corpus Tool (O'Donnell, 2008).

A document similarity assessment tool, namely Ferret 5.0 (Lyon, Malcolm and Dickerson, 2001) was harnessed to identify the degree of text recycling. The corpus was then analysed at textual, interpersonal and ideational levels and the results correlated to numerous factors, such as authorship, sub-discipline and stage of career.

The key preliminary finding is that the lexicogrammatical complexity of abstracts increased significantly over time with more recent abstracts harnessing more descriptive Circumstance blocks.

The provisional pedagogical implications from this study are that novice scientific abstract writers could master this genre faster by adopting a strategy of focussing on thematic development, incorporating more detailed descriptions and borrowing innocuous phrases commonly used within the discourse community of the discipline.

References

- Lyon, C., Malcolm, J., & Dickerson, B. (2001). Detecting short passages of similar text in large document collections. In *Proceedings of Conference on Empirical Methods in Natural Language Processing*. SIGDAT Special Interest Group of the ACL.
- Halliday, M.A.K., & Martin, J.R. (1993). *Writing Science: Literacy and Discursive Power*. London: Falmer Press.
- O'Donnell, M. (2008, June). Demonstration of the UAM CorpusTool for text and image annotation. *Proceedings of the ACL-08: HLT demo session (Companion volume)* (pp.13-16). Columbus, Ohio: Association for Computational Linguistics.
- Simionescu, M., & Simion, E. (2004). *ALLEA Biennial Yearbook Critical Topics in Science and Scholarship* (pp.129-133). Amsterdam: ALLEA.
- Swales, J. (1990). *Genre Analysis*. Cambridge: Cambridge University Press.