

Oita Text Forum
30 November – 1 December 2013

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

Corpus-based study of scientific research article titles

Research article titles act as a primary node for scientific researchers to decide whether or not to read the full article. Given the increasing importance of keyword Internet searches, it is surprising to find only a handful of studies analysing the structure of research article titles. Scholars have compared variation across disciplines (Haggan, 2004), explored the taxonomy (Hartley, 2007), correlated citation count and humour in titles (Sagi & Yechiam, 2008), and investigated their syntactic features (Wang & Bai, 2007; Anthony, 2000). However, no studies in the field of information science were uncovered in an extensive literature search of peer-reviewed journals in the last decade. This paper therefore aims to identify the functional and structural traits of scientific research article titles published in information science

A corpus of 500 research article titles was compiled by selecting the first 100 titles published from January 2012 in 5 highly-ranked IEEE journals (mean 5-year impact factor ranking of 3.8). The corpus was automatically tagged with parts-of-speech tags and manually tagged according to overall structure, namely: sentential, interrogative, nominal, V-ing, prepositional or compound. Nominal titles were subcategorised according to the type of modification (pre or post) and number of heads (uni, bi or multi) in the noun phrase. Compound titles were analysed in terms of the hanging title relationship (e.g. problem/solution).

Preliminary results appear to concur with studies conducted in different fields regarding the syntactic structure, yet show some notable differences. The majority of titles were V-ing or nominal, 15% of which started with an indefinite article. Although there were no sentential or interrogative titles in this corpus, approximately 10% were compound and 5% prepositional.

Detailed results will be shared and pedagogical implications suggested.

References

- Anthony, L. (2000). Characteristic features of research article titles in computer science. *IEEE Transactions on Professional Communication*, 44 (3), 187-194.
- Haggan, M. (2004). Research paper titles in literature, linguistics and science: Dimensions of attraction. *Journal of Pragmatics*, 36 (2), 293-317.
- Hartley, J. (2007). There's More to the Title than Meets the Eye: Exploring the Possibilities, *Journal of Technical Writing and Communication*, 37 (1), 95-101.
- Sagi, I., & Yechiam, E. (2008). Amusing titles in scientific journals and article citation. *Journal of Information Science*, 34, 680-687.
- Wang, Y., & Bai, Y. (2007). A corpus-based syntactic study of medical research article titles. *System*, 35, 388-399.