Abs-257

Eric Hajime Jego (Nihon University School of Medicine – Tokyo, Japan)

Doctor talk: simple corpus methodologies for EMP teachers

Corpus methodologies have been used to create word lists and identify multi-word units to inform pedagogy for general English, academic English and medical English as well. The value of word lists in ESP education is well documented (Nation, 2001; Chujo and Nishigaki, 2006). From West's General Service List (1953) to Xue and Nation's University Word List (1984) to Coxhead's Academic Word List (2000) and Wang et al's Medical Academic Word List (2008), the road toward more specific word lists for use in EMP education has been established. The lists of the past have had favourable impact on EFL education. Despite this positive momentum, research on corpora of oral discourse is sparse and the need to develop a spoken academic word list for pedagogic purposes is clear (Chujo and Nishigaki, 2006). As Pastizzo et al. assert "the use of spoken word frequency counts is conspicuously absent in the literature" (2007: 1025). This area of EMP is particularly understudied in general and no major studies could be found which have attempted to create a spoken word list from a corpus of doctor discourse used in medical interviews for use in EMP education. Furthermore, the lists of past studies go to great levels of sophistication, for example, lemmatisation, tagging text for various items such as parts of speech, statistical analysis, and ensuring balanced representation in corpus construction all of which require corpus expertise. Although the benefits of identifying useful terms and expressions is established, little research exists which examines spoken corpora for medical English teaching purposes. This study seeks to propose simple procedures for identifying terms and expressions that doctors use with their patients which have potential pedagogical value for educators teaching medical interview English in EFL environments.

The methods described in this study use freely available materials including the following: AntConc corpus software toolkit, authentic video content from the emp-tmu.net website which includes real doctors interviewing real patients, and a stop list of the top spoken words from the British National Corpus. A corpus was compiled of all the doctor discourse from the website. These materials were used to create a word list and extract multi-word units of potential interest to English for medical purposes educators. While similar studies are very sophisticated and comprehensive, this study is distinct in that it describes simple and practical methods that even educators without extensive corpus expertise can apply.

References

Chujo, K. and Nishigaki, C. (2006) 'Creating Spoken Academic Vocabulary Lists from the British National Corpus', Practical English Studies, vol. 12, pp. 19-34.

Coxhead, A. (2000) 'A new academic word list', TESOL Quarterly, vol. 34, no. 2, pp. 213-238.

Nation, P. (2001) Learning vocabulary in another language, Cambridge: Cambridge University Press.

Pastizzo, M. and Carbone, R. (2007) 'Spoken word frequency counts based on 1.6 million words in American English', Behavior Research Methods, vol. 39, no. 4, pp. 1025-1028.

Wang, J., Liang, S. and Ge, G. (2008) 'Establishment of a Medical Academic Word List', English for Specific Purposes, vol. 27, pp. 442-458.

West, M. (1953) A general service list of English words, London: Longman, Green & Co.

Xue, G. and Nation, P. (1984) 'A university word list', Language Learning and Communication, vol. 3, no. 2, pp. 215-219.