

A Corpus-based Online Learning System for Enhancing Hong Kong English Major Students' Academic Writing Skills

Lixun WANG

Department of English

The Hong Kong Institute of Education

Hong Kong, China

Lixun@ied.edu.hk

Many Hong Kong English major university students find English academic writing very challenging. One important reason is that meaning is often obscured by wrongly used academic lexico-grammatical items or inappropriately chosen words that make the expressions either unclear or inappropriate. Previous researches have suggested that the corpus-based Data-Driven Learning approach (Johns, 1991, 1993) is an effective way of helping learners to overcome their difficulties in learning lexical-grammatical items (Hadley, 2002; Cresswell, 2007; Johns, Lee & Wang, 2008; Bolton, 2010).

This paper introduces the development of a corpus-based online learning system at the Hong Kong Institute of Education (HKIEd). In order to systematically study the common errors made by the Hong Kong English major students in their English academic writing, a one-million-word English for Academic Purposes (EAP) Learner Corpus (a collection of HKIEd English major students' English essays) has been compiled. At the same time, a one-million-word EAP Professional Corpus (a collection of published English academic articles) has also been compiled as a reference corpus. The two corpora have been analysed both individually and in combination in order to (a) identify Hong Kong students' recurrent difficulties in using English lexico-grammatical items accurately and appropriately in the EAP Learner Corpus, (b) identify the possible sources of such difficulties by studying the differences and similarities of the relevant lexico-grammatical features in the two corpora, and (c) suggest possible remedies that may serve to reduce or eliminate such problems. Based on the analysis, two lists of words have been compiled: 1) those recurrent lexical errors or Hong Kong learners' peculiar lexico-grammatical features found in the learner EAP corpus; 2) the lexico-grammatical items crucial to the different functions that they serve in the professional EAP corpus: introducing topic, referring to literature, expressing opinions, drawing conclusions and summarizing, etc. We examined whether these items found in the professional EAP corpus also occur in the learner EAP corpus, and how often they occur, and whether the use of these items deviates from the examples in the professional EAP corpus.

Based on the two corpora and the analysis, a corpus-based online learning system has been established (<http://engres.ied.edu.hk/corpusbasedlearning>). Learners are able to search for particular lexico-grammatical items in the two corpora using a Web concordancing program we have developed. They can either search one particular corpus, or search both corpora simultaneously and see the concordancing results on the screen side by side. On the online learning system, they will also be able to view annotated notes concerning crucial or problematic academic lexico-grammatical items, and carry out various interactive online

remedial exercises to practise the crucial and problematic academic lexico-grammatical items. Learners' interaction with the system is tracked and recorded for future analyses.

It is hoped that the online learning system, drawing on current findings of corpus linguistics and lexico-grammatical studies, can offer a systematic form of training to students as to how to use lexico-grammatical items correctly and appropriately in academic writing.

References

- Bolton, A. (2010). 'Data-Driven Learning: taking the computer out of the equation.' *Language Learning* 60: 534-572.
- Cresswell, A. (2007). 'Getting to 'know' connectors? Evaluating data-driving learning in a writing skills course.' In Hildalgo, E., Quereda, L., & Santana, J. (eds). *Language and Computer: Corpora in the Foreign Language Classroom*. Amsterdam, Atlanta: Rodopi. 267-287.
- Hayley, G. (2002). 'An introduction to Data-Driven Learning.' *RELC Journal* 33: 99-124.
- Johns, T.F. (1991). 'Should you be persuaded: Two examples of Data-Driven Learning materials.' *English Language Research Journal* 4: 1-16.
- Johns, T.F. (1993). 'Data-driven learning: an update.' *TELL&CALL* 1993/2: 4-10.
- Johns, T.F., Lee, H. & Wang, L. (2008). 'Integrating corpus-based CALL programs in teaching English through children's literature.' *Computer-Assisted Language Learning* 21: 483-507.