

A combined insight into formulaic language: formulaic strings and lexical bundles in Chinese university students' written texts

Jiaoyue Chen (jc6e10@soton.ac.uk)

PhD Candidate in Applied Linguistics, Faculty of Humanities, University of Southampton, SO17 1BJ

Abstract

Previous research on formulaic language has the biased focus either on the formulaicity or frequency as the defined feature of this ubiquitous language phenomenon by drawing on corpus linguistics approach and psycholinguistic account independently. At the same time, the question of how formulaic language is used and identified by language users in the EFL contexts is not fully addressed. This study aims to explore formulaic language used in Chinese university students' written texts by drawing on both corpus-based approach and psycholinguistic account. Thus a combined identification method will be used: student perception and corpus extraction.

Two learner written corpora are compiled by the English written texts produced by the Year 1 and Year 3 students through one semester from a university in China. The Year 1 corpus has 300 texts collected from four writing tasks with 56058 words; the Year 3 corpus compiled by 251 texts under four writing tasks, with 64186 words. Formulaic strings are underlined by 12 students once they finish the texts. Four-word lexical bundles are extracted from the two corpora based on a set of criteria. The identified formulaic language (formulaic strings and for word lexical bundles) are analysed in the structural categories. Furthermore, formulaic language used by the two groups of students (Year 1 and Year 3) is compared.

The results show that formulaic language does exist in the student written texts. The formulaic strings identified by the students include five main structural categories, i.e. templates, formulas, sentence builders, phrases, idiomatic phrases. It suggests that formulaic language exists at both micro-level (for example, templates) and macro-level (such as idiomatic phrases and phrases), rather than only at micro-level. By using two approaches to the identification of formulaic language, there are some discrepancies in the distribution and numbers regarding the structures of formulaic strings and lexical bundles. This reveals different views on formulaic language between the researchers, who relied on the corpus-based approach to investigate the language product from an objective viewpoint, and the users of the language, who usually only focus on the language use in a specific situation. The findings also indicate that more advanced speakers overall tend to use more and longer formulaic expressions.

At the same time, this study probes the potential difficulties in formulaic language learning and using for the EFL students by analysing the formulaic language use. In addition, it calls for the repositioning of formulaic language in English language learning and teaching in the EFL context.