

A corpus-based quantitative approach to cross-linguistic archaism

Meng Ji
Imperial College London
85 Sterling Place, South Ealing, London W5 4RB, UK
m.ji@imperial.ac.uk

As a distinctive feature of modern linguistics, the use of quantitative empirical data becomes increasingly important in describing the behaviour and nature of linguistic events. However, it seems that such data-driven methodology has been rarely discussed in comparative literatures such as translation studies. The present paper constitutes an original attempt in testing the validity of applying a corpus-based quantitative approach to the study of cross-linguistic archaism in Cervantes' *Don Quijote* and its two modern Chinese versions (1978 & 1995). Despite being a universal feature of human languages, archaism may assume very different manners cross-linguistically, e.g. morpho-syntactic variations, grammatical deviations, semantic shifts, pragmatic alterations, etc. In contrast with what has been identified as the main forms of archaism in *Don Quijote* (Mancing, 1982: 15-9), i.e. the morphological substitution of *h* (such as *herida*, *hermosa*, *hacer*) with *f* (*ferida*, *fermosa*, *facer*); the second-person-plural verb forms retaining the *-d-* (*mostredes* for *mostréis*; *hayades* for *hayáis*, etc.); and the adjective *cautivo* used in the sense of *miserable* or *desdichado* in old Spanish, the effect of archaism in the two target texts of *Don Quijote* is achieved mainly through the use of archaic idioms in modern Chinese, which entails the use of lexical and grammatical elements exclusive of ancient Chinese. In an attempt to uncover the underlying patterns of archaism in the source and target texts, a number of corpus linguistic and statistical methods have been employed, including data sampling, data classification, chi-square significance test, non-linear regression scatter-plotting, etc. One of the most interesting patterns emerged from such a corpus-based quantitative approach to the parallel texts is that whereas the frequency of occurrence of archaic idioms in the two modern Chinese versions of *Don Quijote* has proved to be largely independent from the source text feature as containing an archaic use of language, the distribution of archaic expressions seems to show certain similarity between the two sets of target texts. Such textual evidence seems to suggest that the archaic language style intended by Cervantes among his contemporaries may have experienced a significant change when the original is translated into Chinese as shown in its two modern Chinese versions.

References

- Mancing, H. 1982. *The Chivalric World of Don Quijote: Style, Structure and Narrative Technique*, University of Missouri Press, pp. 15-19
- Oakes, M. 1998. *Statistics for Corpus Linguistics*, Edinburgh University Press
- Rico, F. (ed.) 1998. *Don Quijote de la Mancha*, Instituto de Cervantes, Madrid
- Ji, M. 2007. "Phraseology and Idiomaticity: a corpus-based approach to two modern Chinese versions of Don Quijote", in *Proceedings of the Conference and Workshop of Corpus and Translation Studies*, Shanghai Jiao Tong University