A Corpus-based Study of Similes in British and American English

Riyad F Hussein
Department of Linguistics, Faculty of Foreign Languages
The University of Jordan
Amman, Jordan

Majdi Sawalha
Computer Information Systems Department
King Abdullah II School for Information Technology
The University of Jordan
Amman, Jordan

Abstract

This study aimed to investigate the forms of similes and the types of nouns following them and their frequency in different simile constructions. It also unveils the implications that can be drawn from the findings for English as Foreign Language (EFL) learning contexts. More specifically, this study attempts to find out the most frequent nouns following similes in different constructions in two generalized corpora, namely the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA). These similes were identified, retrieved and ranked in a descending order according to their frequencies per one million words. Results indicate that similes, unlike other multi-word units are changeable as different nouns with varying frequencies ranging from one to fifty three were shown to follow each of the similes under investigation. For instance, frequencies of nouns following the simile as good as were as follows: gold (53), (the) people (30), men (14), money (9), and cash (8). These results stood in contrast to the beliefs and opinions of some native speakers of English who view similes as fixed forms, very much like formulaic expressions and thus associated exclusively with certain nouns. The study calls on teachers and instructors to take these findings into account when teaching similes in the EFL context. Additional research is recommended on other similes such as, as sweet as sugar and as cold as ice, for instance in BNC and COCA to confirm or invalidate the findings reported in this research.

Keywords: BNC, COCA, corpora, corpus linguistics, EFL, similes