

Corpus-Based Study of English Vocabulary: Playful Language in Alice's Adventures in Wonderland

Yu-Chih Chen

National Yang Ming Chiao Tung University, Institute of TESOL, Taiwan

Abstract

Corpus Linguistics (CL) has become a potent instrument in language instruction, helping researchers analyze authentic language use by studying corpora. CL's ability to bridge the gap between academic theory and real-world language use has made it especially useful in vocabulary acquisition. This method is particularly effective when exploring creative language, which engages learners in meaningful and enjoyable ways. Lewis Carroll's *Alice's Adventures in Wonderland* provides a rich context for exploring playful language, defined by neologisms, puns, and metaphors. These elements offer insight into how language play enhances reader engagement and vocabulary development. This study applies CL techniques to analyze playful language in *Alice's Adventures in Wonderland*. A mixed-methods approach will be used, integrating quantitative and qualitative corpus linguistic methods. Python-based text processing will extract non-standard words, Grammarly will validate creative word forms, and WordSmith will perform corpus analysis. The text will be compared to the British National Corpus (BNC) to identify words unique to Carroll's work. Extracted playful words will undergo preprocessing, including tokenization, stopword removal, and lemmatization. WordSmith will then conduct word frequency, collocation, and concordance analysis to explore the distribution and patterns of playful language. This research aims to contribute to the field of English language teaching by expanding the use of CL in literary analysis. It will offer empirical insights into how playful language in literature can engage readers and inform material design for language educators and researchers.

Keywords: corpus linguistics (CL); cognitive engagement; literary analysis; textual analysis; vocabulary development