

Abstract:

Lexical diversity touches many areas in linguistics (e.g. L1 and L2 teaching, aphasia diagnosis, forensic linguistics, quantitative linguistics in general). Its values are ascertained using many traditional metrics such as the number of types, the repetition interval of the same words, the semantic connectivity of words, etc. These metrics have been introduced without examining how they correlate with subjective assessments of lexical diversity. Thus, the first external subjective evaluation of the validity of these metrics was only carried out during the last decade in several papers by the teams around Scott Jarvis. In an experiment described in a March 2021 paper by American linguists Kristopher Kyle, Scott A. Crosley and Scott Jarvis, they had respondents rate the lexical diversity of about six hundred texts and compared their ratings with indices calculated by the traditional metrics mentioned earlier. The topic of lexical diversity and its validation has resonated strongly in the linguistic community since the paper was published. The aim of this paper is to replicate the experiment of Scott Jarvis et al. for English.

Keywords: lexical diversity, metrics of lexical diversity, repetition of the experiment, quantitative linguistics