

Lexical diversity in human- and LLM-generated text

Kelly Kendro

Northern Arizona University, Flagstaff, Arizona, United States

Jeffrey Maloney

BYU-Hawaii, Laie, Hawaii, United States

Scott Jarvis

Northern Arizona University, Flagstaff, Arizona, United States

Abstract

Despite the widespread adoption of public-facing large language models (LLMs) over the past several months, we still know little about the complexities of machine-generated language in comparison to human-generated language. To better understand how lexical complexity differs between human- and LLM-produced texts, we elicited responses from four commercially-available LLMs (ChatGPT 3.5, ChatGPT 4.0, Claude, and Bard), and compared them to writing from humans from different backgrounds (i.e., L1 and L2 English users) and education levels. We also investigated whether the LLMs demonstrated consistent style across targeted prompts, as compared to the human participants. Through an analysis of six dimensions of lexical diversity (volume, abundance, variety-repetition, evenness, disparity, dispersion), preliminary results suggest that LLM-generated text differs from human-generated with regards to lexical diversity, and texts created by LLMs demonstrate less variation than human-written text. We will discuss the implications of these differences for future research and education in applied linguistics.