

## Lexical Predictors of Generative-AI Mnemonic Quality for Chinese Vocabulary

Recent advances in large language models have led to growing interest in using AI-generated mnemonics to support second language vocabulary learning. However, little is known about how lexical properties of Chinese words systematically constrain the quality of such AI-generated pedagogical content. This study investigates the extent to which key lexical features, such as structural complexity (word length), concreteness, and semantic transparency, predict the quality of AI-generated mnemonics for Chinese vocabulary items.

A dataset of 123 Chinese lexical items was constructed, including both single-character words and multi-character compounds. For each word, three mnemonics were generated using an identical prompt, yielding a total of 369 AI-generated mnemonics. Mnemonic quality was evaluated by trained raters using the Mnemonic Quality Rating Scale (MQRS), which assesses accuracy, clarity, key-cue capture, and proficiency alignment. Lexical properties were independently rated by trained raters with formal linguistic training in Chinese. To account for repeated mnemonic generation within words, linear mixed-effects models with random intercepts for lexical items were fitted.

Results from the full dataset indicate that both structural complexity and concreteness are significant positive predictors of mnemonic quality. Specifically, mnemonics generated for structurally more complex and more concrete words tend to receive higher quality ratings. At the model level, lexical fixed effects account for approximately 22% of the variance in mnemonic quality, while word-level random effects explain a substantial proportion of the remaining variance, highlighting pronounced differences across lexical items. A follow-up analysis focusing exclusively on multi-character compounds reveals a distinct pattern. Within this subset, semantic transparency emerges as a robust and significant predictor of mnemonic quality, whereas structural complexity and concreteness do not exhibit significant effects. This suggests that once words are multi-character compounds, the extent to which their meanings are transparently recoverable from constituent characters becomes the primary determinant of mnemonic quality.

Taken together, the findings suggest that the effectiveness of AI-generated mnemonics is not uniform across Chinese lexical items, but is systematically shaped by linguistic properties at both the monomorphemic and compound word levels. By identifying which lexical features facilitate or constrain high-quality mnemonic generation, this study contributes empirical evidence to ongoing discussions on Chinese word structure, semantic transparency, and AI-assisted language learning. More broadly, the results point to the importance of linguistically informed constraints in the principled design and pedagogical deployment of AI-generated learning materials for Chinese vocabulary instruction.