

# Visual groundedness as an organizing principle for word class: Evidence from Japanese

Coleman Haley

University of Edinburgh, Edinburgh, United Kingdom

Sharon Goldwater

University of Edinburgh, Edinburgh, United Kingdom

## Abstract

How do languages structure word classes, and is this organization arbitrary? This study explores groundedness—the degree of association between a word in an utterance and the meaning which the utterance describes—as a potential organizing factor. In particular, we look at visual groundedness, where meaning is approximated with an image, allowing tractable estimation with neural vision-and-language models. Prior work showed that nouns, adjectives, and verbs differ in groundedness cross-linguistically. We test whether groundedness describes the atypical word class structure in Japanese, where adjectives are split into i-adjectives (formally verb-like) and na-adjectives (noun-like). Analyzing the Crossmodal-3600 dataset with the PaliGemma model, we find that na-adjectives exhibit significantly ( $p=0.029$ ) higher visual groundedness, suggesting that the formal similarities of these classes to nouns and verbs reflect their semantics. This challenges the idea that linguistic categories are purely conventional or innate and arbitrary, and supports a view where language structure emerges from human perception.