

26 - 28 June 2026

Kuala Lumpur , Malaysia

# The Polysemy of “Bias” in Contemporary English: A Corpus-Based Disambiguation Analysis Across Social, Statistical, Juridical, Technical, and Everyday Registers

Huy Nguyen , Nguyen Nguyen Khang

*Van Hien University, Vietnam*

## Abstract

“Bias” is a polysemous word across social, statistical, legal, technical, and everyday registers. However, there is still a lack of systematic corpus-based evidence determining, including insufficient data to identify which collocational cues correctly distinguish each established sense, and which established sense dominates in naturalistic daily usage. Thus, this study investigates the seven senses of the word “bias” in the Cambridge Dictionary, including “prejudice,” “preference,” “statistical,” “bowls,” “fabric,” “verbal,” and “political and legal,” by way of analyzing 13,189 KWIC instances extracted from eight sub-corpora with a total of 255,907 tokens. The dataset includes seven sub-corpora stratified by sense and one benchmark sub-corpus of modern written and spoken English. Extracting concordance lines via AntConc (version 4.3.1) allowed for a collocate analysis within an L5-R5 window, evaluated through both Log-Likelihood and Mutual Information metrics. Statistical testing in R (version 4.5.1) began with Pearson’s chi-square tests of independence. To adjust for expected frequency violations, a Monte Carlo simulation with 10,000 replicates was implemented within JASP (version 0.18.3). The omnibus test yielded highly significant outcomes ( $p < .001$ , Cramér’s  $V = .62$ ), proving that the seven sense profiles occupy distinct boundaries within the node’s lexical space. Specifically, a chi-square goodness-of-fit test identified the “preference” sense as the predominant interpretation within well-classified daily usage cases ( $p = .029$ ). Finally, the study gives statistically grounded disambiguation cues that its positional collocate-weighting framework can apply to other high-frequency polysemous lexemes, offering direct practical value to lexicographic practice as well as to journalistic and legal communication.

**Keywords:** Bias; Collocational Analysis; Corpus Linguistics; Lexicography; Polysemy