



SEMANTIC PREFERENCE ON JOURNAL ARTICLES: CORPUS-BASED STUDY OF ADJECTIVES IN ENGLISH

Meira Purnama

Master of Linguistics Program, Warmadewa University
meirapurnamayanti@gmail.com

Abstrak

The first part of the study may examine the semantic preference of the most common adjectives in scholarly publications by doing a corpus-based analysis of the articles published in five journals across the three corpora (Humanities, Health Sciences, and Social Politics). Using corpus-based analysis, we determined the most often occurring adjectives that occur together across all five periodicals. By analyzing the similarity between adjectives and the words to which they were most often applied, we were able to understand the adjectives' individual preferences in terms of meaning. The second section of this research analyzes the contrasts and similarities between the three corpora in terms of the adjectives used in scholarly publications from both domains (humanities, health sciences, and social politics). There were only three adjectives found to collocate with both nouns, and they are all related to cognition: cognitive, different, and high. While the words "different" and "high" are used to emphasize the dissimilarities and similarities across the three corpora, the word "cognitive" appears only in the Humanities corpus, making comparisons and dissimilarities between the corpora impossible. Journal articles in the humanities and health sciences utilize the term "different" in the findings and discussion sections. However, there are significant distinctions across the fields due to their respective markets of emphasis. In the lexicon of the social sciences and the health sciences, the adjective "high" has etymological roots with other terms that signify measurement and quantification. Therefore, the adjective is utilized quite differently in both corpora.

Keywords: Semantic preferences, adjective, humanities, health sciences, social politics.

Abstrak

Bagian pertama dari penelitian ini dapat memeriksa preferensi semantik dari kata sifat yang paling umum dalam publikasi ilmiah dengan melakukan analisis berbasis korpus dari artikel yang diterbitkan dalam lima jurnal di tiga korpora (Humaniora, Ilmu Kesehatan, dan Sosial Politik). Dengan menggunakan analisis berbasis korpus, kami menentukan kata sifat yang paling sering muncul yang muncul bersamaan di kelima majalah. Dengan menganalisis kesamaan antara kata sifat dan kata-kata yang paling sering digunakan, kami dapat memahami preferensi individu kata sifat dalam hal makna. Bagian kedua dari penelitian ini menganalisis perbedaan dan kesamaan antara ketiga korpora tersebut dalam hal kata sifat yang digunakan dalam publikasi ilmiah dari kedua domain (humaniora, ilmu kesehatan, dan sosial politik). Hanya ada tiga kata sifat yang ditemukan berkolokasi dengan kedua kata benda, dan semuanya terkait dengan kognisi: kognitif, berbeda, dan tinggi. Sementara kata-kata "berbeda" dan "tinggi" digunakan untuk menekankan perbedaan dan kesamaan di antara tiga korpora, kata "kognitif" hanya muncul dalam korpus Humaniora, membuat perbandingan dan perbedaan antara korpus menjadi tidak mungkin. Artikel jurnal dalam ilmu humaniora dan kesehatan menggunakan istilah "berbeda" di bagian temuan dan pembahasan. Namun, ada perbedaan yang signifikan di seluruh bidang karena penekanan pasar masing-masing. Dalam leksikon ilmu sosial dan ilmu kesehatan, kata sifat "tinggi" memiliki akar etimologis dengan istilah lain yang menandakan pengukuran dan kuantifikasi. Oleh karena itu, kata sifat digunakan secara berbeda di kedua korpora.

Kata kunci: Preferensi semantik, kata sifat, humaniora, ilmu kesehatan, sosial politik.

I. INTRODUCTION

The semantic preference hypothesis says that many uses of words and phrases "tend to occur in a particular semantic context." For instance, the adjective "large" is often associated with "quantities

and sizes," such as numbers, scale, parts, amounts, and quantities. Numerous scholars in the field of corpus linguistics have carefully studied the concepts of semantic preference and semantic prosody as corpora have grown in size and methods for extracting various lexical items for various purposes have been developed (Sidupa, 2019). Stubbs (2021) defined semantic preference as the association between a lemma or word form and a set of words with a shared meaning. According to Stubbs, an item exhibits semantic preference when it co-occurs with "a class of phrases that have some semantic quality." When a certain lexical item collides with a collection of semantically related terms frequently, a phenomenon is also known as semantic preference develops (Begagic, 2013). Semantic preference is thought to be shared among speakers in a certain community and to depend on register, context, and domain. People who wish to be involved in the educational institutions must thus be familiar with how it is often used.

Wei et al., for instance, look at prosody and semantic preference (2014). These patterns may be identified in word relationships, phrase structure, and word choice, among other things. Over the last three decades, there has been a lot of study done on semantic prosody. Since it has been used to a number of secondary meanings, none of which are identical to the others, such as connotation, evaluation, appraisal, pragmatic force, irony, and sincerity, the word "semantic prosody" is still controversial (Philip, 2011). Longer text passages are impacted by semantic prosody, albeit preference is expressed at the semantic set level (Wang & Zhou, 2022). Collocates of the node impact semantic preference instead of the node word's semantic prosody. Semantic prosody "substantially contributes to the creation of semantic preference" and "determines the entire environment that restricts the desirable possibilities of the node item," according to Partington (Partington, 2004).

As long as scholars are always under pressure to read and write dense academic materials, semantic preference research is essential (Selmitraitis, 2020). It is accepted that English and Chinese have contributed to "achieving equivalency between lexical items of the two languages with matching counterparts." English is the language that is most often used in academic writing. For native English speakers, finding academic English phrases and collocations appropriate for a very specific linguistic circumstance is often not too difficult (Afzal, 2019). Non-native speakers, however, have a distinct set of challenges. Non-native speakers, students, professors, and scientists often struggle to choose the proper term in papers written in a variety of academic fields. When writing academic papers, it's usual to forget to consider the connected word pairs' semantic preferences.

Semantic Because the study assumes that semantic preference is register- and domain-dependent, the meanings and communicative functions of the associations were taken into consideration in relation to the academic register and research area. In other words, semantic preferences were developed in light of the specific context of journal articles from the specific academic disciplines. The semantic preferences found in each corpus (Humanities, Health Sciences, and Social Politics) were compared to determine how different the language is in the three fields with regard to the use of the five adjectives (Lahlou, 2020). Think about comparing the three corpora's use of the word "specific." When other words that describe the particular type of human communication are also present, it is more frequently used in the corpus of humanities. In the Health Sciences corpus, it is combined with other words to clarify medical systems or even purposes. Meanwhile, the corpus of social politics employs language to convey specific concepts regarding social and political relations (Vaccari et al., 2015). In order to determine whether there were any differences in the language used in the three corpus areas, the second part of the study included a comparative analysis, which attempted to relate the communicative functions of the adjectives to the sections of journal articles, such as the introduction, methods, results, discussion, and conclusion.

There aren't many academic studies and initiatives that could employ preference since it hasn't gotten much investigation. The current study compares the use of adjectives with various semantic preferences in two corpora (social science and humanities) and discusses semantic preference for the adjective that appears most frequently in five journal articles in order to close this gap and complete the studies on the significance of semantic preference. The semantic preferences of 10 adjectives are explicitly examined in this research. The semantic preferences of each adjective might be ascertained by examining the concordance between the adjectives and the nouns they were paired with. The second section of this research contrasts and analyzes the three corpora in terms of how adjectives are used in both academic and popular journal articles (humanities, health sciences, and social politics).

II. METHODS

The corpora utilized in this research include articles from five publications in the social sciences, health sciences, and humanities. Seven journal articles comprising 33,973 words from the three publications International Journal of Lexicography (IJL), Journal of Human Rights Practice (JHRP), and Journal of Language Evolution are included in the Humanities corpus (JLE). The International Journal for Quality in Health Care produced 10 journal articles totalling 28,832 words for the Health sciences corpus (IJQHC). Contrarily, the Social Politics corpus consists of three pieces with a combined word count of 12,593 that were printed in the journal Social Politics International Studies: Gender, State, and Society (SPI). The result is a total of 75,398 words spread over the five publications.

In the first stage of this study, using the program AntConc 3.5.8, the most prevalent adjectives that co-occurred in the five publications in the domains of Humanities, Health Sciences, and Social Politics were identified. This research analyzes five of the sub-most corpora's used adjectives: distinct, particular, cognitive, substantial, and high. Table I displays both the overall frequency of incidence of the adjectives in the corpus of the humanities, health sciences, and social politics as well as their frequency of occurrence in each journal individually. The adverbs are arranged alphabetically.

III. RESULTS AND DISCUSSION

Semantic sets were constructed when each word connected to one of the five adjectives was investigated in its context. Each semantic set was analyzed and investigated in order to determine semantic preferences. The adjectives' most common synonyms and favored semantic implications are shown in Tables VII–XI. The most common collocates are shown in each table in the groupings that correspond to their semantic sets, with the total number of occurrences in each of the three corpora (Humanities, Health Sciences, and Social Politics) in parentheses. The written form of each noun included in the tables corresponds to its single or plural most frequent use across the corpus. If a word is included in the singular form in the table, for example, it was probably used more often in the corpus. A noun's plural form implies that this was how it occurred most often in the corpus.

Semantic Preferences in Humanities Corpus

Cognitive

In order to assign each word to the most suitable semantic field or semantic preference, the context of the use of each association adjective + collocate was examined. For instance, because the noun "science" appeared in combinations like cognitive science, it breaks into a set of words connected to general research terms.

Collocates and Semantic Preference of Cognitive in Humanities Corpus

COLLOCATES	SEMANTIC PREFERENCES
Human (1), skills (2), trends (1), gestures	Words related to humanities
Studies (2), system (1), science (4)	General research terms

The adjective "cognitive" was used in the corpus to express knowledge or mental process of understanding. The analysis of the use of "cognitive" through journal articles are mostly expressed in the abstract sections.

- 1) Linguistics as part of the **cognitive sciences**.
- 2) We claim that the evolution of bodily mimesis allowed for the use of signs, and the social-**cognitive skills** needed to support them to emerge in hominin evolution.

Different

Collocates and Semantic Preference of Different in Humanities Corpus

COLLOCATE	SEMANTIC PREFERENCE
-----------	---------------------

Kinds (5)	Classification or type
-----------	------------------------

The adjective “different” was associated with classification or type. Combinations such as “different kinds” seem to be used in the corpus to differentiate the human communicative system.

- 1) A crucial, though theoretically, underdeveloped distinction is that between two **different kinds of** semiotic units; signals and signs.
- 2) The combination of several **different kinds of** semiotic systems within an integrated communicative system.

Semantic Preferences in Health Sciences Corpus

Different

Collocates and Semantic Preference of Different in Health Sciences Corpus

COLLOCATES	SEMANTIC PREFERENCE
Types (4)	Categorization

The words associated with the adjective "different" in the context of the field of health sciences are shown in Table IX. These associations were used to distinguish between different types of medical procedures (e.g., different types of surgical procedures), to refer to a particular type of medical instrument (e.g., different types of ventilators), and, most importantly, to distinguish between different types of medical risks (different types of risks).

1. Clinical staff could have been asked to use various ventilators with known risks of inadvertently pressing the wrong buttons or viewing inaccurate information on screens while working in ICUs and the new National Health Service field hospitals.
2. PAP has been shown to lower the risk of SSIs across a variety of surgical procedures in numerous randomized controlled trials and meta-analyses.

Insofar as an imbalance has been created in the relationship between the three different types of risks that the IPCS typically tries to manage, the rapid increase in the number of people infected with COVID-19 may have changed how risk for HAI and AMR is perceived. **High**

Collocates and Semantic Preference of High in Health Sciences Corpus

COLLOCATES	SEMANTIC PREFERENCE
Risk (4), quality (4)	Measurement

Table X describes the variety of words related to the field of Health Sciences that are associated with the adjective “high”. These associations were used to indicate the measurement regarding health care systems. The association with “high” was mainly found in the introduction section of the research.

1. It is important to note that the BTA criteria used in high-risk industries to determine what qualifies as a full system "barrier" may reveal the flaws and restrictions of the proposed "strong systemic barriers" (such as standardized procedures or cognitive aids), which are claimed to, if successfully implemented, prevent specific secondary care "never events."
2. The organizational framework that supports the delivery of high-quality care has not kept pace with the technical advancement of medicine in general.

Semantic Preferences in Social Politics Corpus

High

Table below illustrates the word related to the field of Social Politics that are associated with the adjective “high”. These associations were used to emphasize the quantitative analysis concerning the social case, that is, varieties of gender regimes. The associations with “high” were mainly found in results and discussion part of the research.

Collocates and Semantic Preference of High in Social Politics Corpus

COLLOCATES	SEMANTIC PREFERENCES
Levels (5)	Quantification

1. The relationships and acts of violence—whether interpersonal, interstate, or intergroup—are interconnected and constitute a single institutionalized domain of violence; an illustration of this is the discovery that militarization and interpersonal homicide rates are positively correlated.
2. To combat the high levels of violence their policies have caused, neoliberal states frequently transform into security states that employ high levels of coercion and violence (Walby 2009;Walby et al. 2015;Walby et al.2017).
3. Conservative Does "neoliberal" adequately describe all public forms involving high levels of inequality and slender democracy, or is "conservative" a third type of public gender regime?

Comparative Analysis among Humanities, Health Sciences, and Social Politics Corpus

In order to describe the variations and/or continuities in the use of the five adjectives in the articles published in the journals from the fields, the second part of this study compared corpora from the humanities, health sciences, and social politics fields.

Cognitive

The use of signs and the social-cognitive abilities required to support them are said to have emerged throughout the development of hominins as a result of the emergence of physical mimesis, as stated in the sentence "We claim." The term "cognitive" was used in conjunction with other adjectives to describe the way in which humanities cases were understood mentally. The word "cognitive" was also used alongside the phrases "studies," "system," and "science" to signify generic research terminology. However, the adjective "cognitive" was not used with a noun in two other corpora (Health Sciences and Social Politics). Adjective + noun collocations were not preferred in these corpus areas according to semantic analysis.

Different

In order to differentiate between the various human communicating systems, the Humanities corpus uses several descriptors. In the Health Sciences corpus, the adverb different was utilized to differentiate between diverse medical techniques, equipment, and even hazards. In the corpus of the humanities, examples of some of these combinations include diverse semiotic systems and semiotic units, while in the health sciences, examples include various surgical procedures and ventilators. Particularly in the discussion section, journal articles from both domains revealed correlations between the phrases distinct and sorts and types. However, there were no examples of the adjective "different" being used with nouns in the Social Politics corpus. In conclusion, there are some parallels and variations between the semantic preferences of the Humanities and Health Sciences corpora. The adjective "different" and its combinations were more commonly used in the discussion portion of journal articles in both corpora. The adjective "different" was utilized differently in each field at the time.

High

The two corpora (Health Sciences and Social Politics) were the only ones to use the adjective "high" with nouns when discussing measurement and quantification, despite the fact that this usage was noticeably more common in the Health Sciences corpus when discussing the level of healthcare systems. While the associations with high levels in the social sciences were frequently found in the results and discussion sections, those with high levels in the health sciences were typically found in the introduction. In conclusion, the word "high" indicated measurement or even quantification in a similar way in both corpora. However, the adjective usage in the journal articles of both corpora varied. The word "high" was mostly used in the introduction section of the Health Sciences corpus, whereas it was mostly used in the result and discussion sections of the Social Politics corpus. Additionally, the adjective "high" appears more frequently in the corpus of health sciences than it does in the corpus of social politics.

IV. CONCLUSION

Out of the five that were chosen, only three adjectives—cognitive, different, and high—collocate with the nouns. These adjectives are distributed consistently throughout the journal articles in the three corpora. The humanities corpus contains the words "cognitive" and "distinct." The corpus of the health sciences contains the words "distinctive" and "high." The Social Politics Corpus, however, only includes the term "high." Adjectives are used to highlight the parallels and contrasts between the three corpora, notably "different" and "high." In the findings and discussion sections of journal papers in the humanities and health sciences, the term "different" and its combination is commonly used. In contrast to the corpus of the humanities, where they are employed to differentiate the human communication system, the health sciences corpus commonly uses the adjectives "different" to describe the numerous sorts of medical treatments, hazards, and devices. The word "high" is used to describe measurements and quantification in corpora of the social sciences and health. Journal articles from the two corpora employ adjectives differently, but these variations are exclusive to adjective use. While it predominated in the findings and discussion portion of the Social Politics corpus, the term "high" predominated in the introduction section of the Health Sciences corpus.

REFERENCES

- Afzal, N. (2019). A Study on Vocabulary-Learning Problems Encountered by BA English Majors at the University Level of Education. *Arab World English Journal (AWEJ)*, 10(3), 81–98.
- Begagic, M. (2013). Semantic Preference and Semantic Prosody of The Collocation Make Sense. *Jezikoslovlje Journal*, 3, 403–406.
- Lahlou, H. (2020). The Motivations for the Semantic Change in the Category Green in Arabic: A Synchronic and Diachronic Corpus-Based Analysis. *Open Journal of Social Sciences*, 8(1).
- Nafilaturif'ah, N., & Poluwa, M. I. (2021). Semantic Prosody and Preference of "Healthy" and "Unhealthy" Collocations in Covid-19 Corpus. *Journal of Linguistics, Literature and Language Teaching*, 5(2).
- Partington, A. (2004). Utterly Content in Each Other's Company: Semantic Prosody and Semantic Preference. *Semantic Preference. International Journal of Corpus Linguistics*, 9(1), 131–156.
- Philip, G. (2011). *Colouring Meaning: Collocation and Connotation in Figurative Language*. John Benjamins Publishing Company.
- Selmitraitis, L. (2020). Semantic Preference, Prosody and Distribution of Synonymous Adjectives in COCA. *Journal of Language Studies*, 20(3), 1–8.
- Stubbs, M. (2021). Text, Corpora, and Problems of Interpretation: A Response To Wid-Dowson. *Applied Linguistics*, 22(2), 49–172.
- Vaccari, C., Valeriani, A., Barberá, P., Bonneau, R., Jost, J. T., Nagler, J., & Tucker, J. A. (2015). Political Expression and Action on Social Media: Exploring the Relationship Between Lower- and Higher-Threshold Political Activities Among Twitter Users in Italy. *Journal of Computer-Mediated Communication*, 20(2), 221–239.
- Wang, J., & Zhou, J. (2022). A Corpus-Based Study of Semantic Categorizations of Attracted Adjectives to the it BE ADJ clause Construction. *SAGE Journals*.