Semantic Preference on Journal Articles: Corpus-Based Study of Adjectives in English

Ni Putu Meira Purnama Yanti[1]
[1]Universitas Warmadewa, Denpasar

e-mail: [1]meirapurnamayanti@email.com

Abstracts: The first part of this study is to investigate the semantic preference of the most frequent adjectives in journal articles through corpus-based analysis of the articles published in five journals in the three corpora (Humanities, Health Sciences, and Social Politics). The corpus-based analysis was carried out on the corpus to identify the most frequent adjectives that co-occurred in the five journals. By observing the concordance of the adjectives and analyzing the words they associated with, the semantic preferences of each adjective were determined. Then, the second part of this study aims to describe the differences and similarities between the three corpus (humanities, health sciences, and social politics) concerning the use of adjectives on journal articles from both areas. The results show that there are only three adjectives that collocate with the nouns, namely cognitive, different, and high. The use of the adjectives, especially "different" and "high" show the similarities and differences in the three corpora, meanwhile the adjective "cognitive" only exists in the Humanities corpus, thus it cannot be found the similarities and differences with other two corpora. In the corpus of Humanities and Health Sciences, the adjective "different" is used in the results and discussion section of the journal articles. Meanwhile, the differences are in the domain of each discipline. In the Health Sciences and Social Politics corpus, the similarities of the adjective "high" are in terms of indicating the measurement and quantification. While the differences are the use of the adjective in the part of journal articles of both corpora. Meanwhile, the differences are in the domain of each discipline. In the Health Sciences and Social Politics corpus, the similarities of the adjective "high" are in terms of indicating the measurement and quantification. While the differences are the use of the adjective in the part of journal articles of both corpora)

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

INTRODUCTION

The concept of semantic preference comes from the notion that "many uses of words and phrases show a tendency to occur in a certain semantic environment.". For example, the word (or lexical item) "large" frequently associates (or collocates) with words for "quantities and sizes", such as numbers, scale, part, amounts, quantities. Stubbs (2001) defines semantic preference as the relation, not between individual words, but between a lemma or word form and a set of semantically related words. Stubbs adds that an item shows semantic preference when it co-occurs with "a class of words which share some semantic feature". Semantic preference describes the phenomenon in which a particular lexical item collocates frequently with a series of items that belong to a semantic set (Begagic, 2013). Semantic preference is believed to be dependent on the register, context, and domain, it is probably shared among speakers of a given community. Therefore, people who seek to become part of their academic community need to be familiar with the language patterns being used. These patterns include word choices, word associations, phrases, and others. As far as people as researchers are under constant pressure to read and publish academic texts of considerable size, research in semantic preference becomes of crucial importance (Selmistraitis, 2020). Predominantly the medium of academic texts is the English language. The choice of English academic words and collocations proper for the very specific linguistic environment usually does not cause great difficulty for native speakers. However, it is not the case for nonnative ones. Non- native speakers, students, secondary level teachers, scientists often struggle with the choice of the correct word suitable for texts written in different academic disciplines. Pairs of related words often are a stumbling block on the road to favorable result in composing academic texts because the semantic preference of these words is neglected. The research regarding semantic preference is not extensive and the range of its practical application in academic studies and academic work is rather narrow. To fill this gap and complement the studies on the importance of semantic preference, thus the current research discusses

semantic preference about the most frequent adjective used in five journal articles and the comparison between two corpora.

METHODS

The corpora used in this research comprise articles collected from five journals in the fields of Humanities, Health Sciences, and Social Politics. The Humanities corpus contains seven journal articles published in the three following journals: International Journal of Lexicography (IJL), Journal of Human Rights Practice (JHRP), Journal of Language Evolution (JLE), totaling 33.973 words. The Health sciences corpus includes ten journal articles published in one journal, namely International Journal for Quality in Health Care (IJQHC), with a total of 28.832 words. Whereas, Social Politics corpus consists of three articles in one journal, that is, Social Politics International Studies: Gender, State, and Society (SPI), totaling 12.593 words. Thus, the total of the words in the five journals are 75.398 words. In the first part of this research, the software AntConc 3.5.8 was used to identify the most frequent adjectives that co-occurred in the five journals in the fields of Humanities, Health Sciences, and Social Politics. This paper presents the analysis of five of the most frequent adjectives identified in the sub-corpora, namely different, specific, cognitive, significant, and high. Table 1 shows the frequency of the adjectives in each journal separately and their total frequency of occurrence in the Humanities, Health Sciences, and Social Politics corpus. The adjectives are presented in alphabetical order.

Table 1. Frequency of Adjectives on Journals in Humanities Corpus

ADJECTIVES	**IJL**	**JLE**	**JHRP**
	92	**292**	**44**
Cognitive	0	116	0
Different	12	88	4
High	8	16	24
Specific	0	72	12
Significant	72	0	4

Table 2. Frequency of Adjectives on Journals in Health Sciences Corpus

ADJECTIVES	**IJQHC**
	336
Cognitive	28
Different	108
High	64
Specific	108
Significant	28

Table 3. Frequency of Adjectives on Journals in Social Politics Corpus

ADJECTIVES	SPI
	88
Cognitive	0
Different	52
High	24
Specific	4
Significant	8

In the five sub-corpora, the five adjectives are among the most frequent ones with some differences in the order of frequency of occurrence in each journal, as shown in Table 4 - 5.

Table 4. Frequency Order of Adjectives on Journals in Humanities Corpus

ORDER	IJL	JLE	JHRP
1	Significant	Cognitive	High
2	Different	Different	Specific
3	High	Specific	Significant
4	Specific	High	Different
5	Cognitive	Significant	Cognitive

Table. 5. Frequency Order of Adjectives on Journals in Health Sciences Corpus

ORDER	IJQHC
1	Different
2	Specific
3	High
4	Cognitive
5	Significant

Table 6. Frequency Order of Adjectives on Journals in Social Politics Corpus

ORDER	SP I
1	Different
2	High
3	Significant
4	Specific
5	Cognitive

After selecting the most frequent items to be investigated, the nouns associated with each adjective, with a co-occurrence frequency of 3 or above, were identified and grouped into the semantic sets. These sets were analyzed to determine the semantic preferences of the adjectives within the five journals in the fields of

Humanities, Health Sciences, and Social Politics. In this part of the analysis, the concordance lines of the adjectives were studied manually, making it possible to observe the meaning of the words in context and to place them into the semantic field. For instance, by checking the list of collocates and the concordance lines one of the adjectives in the three corpora, "specific", it could be observed that some words that are associated with were human, system, and intervention.

Since the study assumes that semantic preference is register- and domain-dependent, data analysis, and interpretation was conducted by considering the meanings and communicative functions of the associations in the relation to the academic register and research area. That is, semantic preferences were determined concerning the specific context of journal articles in the specific academic fields. In this respect, the semantic preferences identified in each corpus (Humanities, Health Sciences, and Social Politics) were compared to verify how distinct the language in the three fields are concerning the use of the five adjectives. For instance, when comparing the use of "specific" in the three corpora. In the Humanities corpus, it is more commonly used with words related to describe the particular system of human communication. In the Health Sciences corpus, it is used with the words to clarify medical purposes or even systems. Meanwhile, the corpus of Social Politics, it associates with words to express particular things regarding social and political relations. This comparative analysis was conducted as the second part of the study, and in case of finding some differences between the language used in the three areas of the corpus, it attempted to relate the communicative functions of the adjectives to the sections comprising journal articles, namely introduction, methods, results, discussion, and conclusion.

RESULTS AND DISCUSSION

All the words related to the five adjectives were examined in the context and grouped into semantic sets. Each semantic set was interpreted and analyzed in the case of semantic preferences. Tables VII-XI indicate the most common words associated with the adjectives and their semantic preferences. In each table, the most common collocations are displayed in the groups that conform to their semantic sets, with the number of occurrences in the three corpora (Humanities, Health Sciences, and Social Politics) in parentheses. The written form of each noun included in the tables corresponds to the most frequent form (singular or plural) of the noun used in the corpus. For example, if in the table the noun appears in the singular form, it means it occurred more often in the singular form in the corpus. If the noun is displayed in the table in its plural form it means it mostly appeared in the corpus in its plural form.

Semantic Preference in Humanities Corpus

1. Cognitive

It was based on this assumption that the context of the use of each association adjective + collocate was analyzed to place each word into the most appropriate semantic field or semantic preference. For instance, the noun "science" breaks into a set of words related to general research terms because it occurred in combinations such as cognitive science.

Table 7. Collocates and Semantic Preference of Cognitive in Humanities Corpus

COLLOCATES	SEMANTIC PREFERENCES
Human (1), skills (2), trends (1), gestures (1), evolution (1)	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.

a. Linguistics as part of the **cognitive sciences**.
b. 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.

2. Different

Table 8. 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.

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

Semantic Preference in Health Sciences Corpus

1. Different

Table 9. Collocates and Semantic Preference of Different in Health Sciences Corpus

COLLOCATES	SEMANTIC PREFERENCE
Types (4)	Categorization

Table 9 illustrates the word related to the field of Health Sciences that are associated with the adjective "different". These associations were used to indicate the distinction of medical procedure (different types of surgical procedures), to refer to one of a medical instrument that is different from others (different types of ventilators), and especially to indicate various kinds of medical risks (different types of risks).

a. Clinical staff working in ICUs and at the new National Health Service field hospitals could have been asked to use **different types of** ventilators with known risks of accidentaly pressing the wrong buttons or misleading information on screens.
b. PAP has been demonstrated in multiple randomized controlled trials and meta- analyses to reduce the risk of SSIs across **different types of** surgical procedures.
c. The rapid rise in the number of people infected with COVID-19 might have brought a change in perception of risk for HAI and AMR in as far as an imbalance has been created in the relationship between three **different types** of risks that the IPCS normally tries to manage.

2. High

Table 10. 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.

a. It is worth noting that the BTA criteria applied in **high-risk** industries to determine what can be considered to be a full system 'barrier' may expose the weaknesses and limitations of the proposed 'strong systemic barriers' (e.g.standardized procedures or cognitive aids) that are claimed will prevent specific secondary care 'never events' if successfully implemented.
b. The technical development of medicine in general is not accompanied by the organizational system that supports the delivery of **high-quality**.

Semantic Preference in Social Politic Corpus

1. High

Table XI 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.

Table. 11 Collocates and Semantic Preference of High in Social Politics Corpus

COLLOCATES	SEMANTIC PREFERENCES
Levels (5)	Quantification

a. The relations and practices of violence that may be interpersonal, interstate, or intergroup are interconnected and form a single institutionalized domain of violence; for example, the finding that there is a correlation between countries with **high levels** of interpersonal homicide and **high levels** of militarization.
b. Neoliberal states often develop into security states deploying **high levels** of coercion and violence to address the **high levels** of violence their policies have generated (Walby 2009;Walby et al. 2015;Walby et al.2017).
c. Conservative is "conservative" a third type of public gender regime, or does "neoliberal" sufficiently capture all public forms involving **high levels** of inequality and thin democracy.
Comparative Analysis among Humanities, Health Science, and Social Politics Corpus

The second part of this study was a comparative analysis among Humanities, Health Sciences, and Social Politics corpus aimed to describe the differences and/or similarities regarding the use of the five adjectives in the articles published in the journals from the fields.

1. Cognitive

In the Humanities corpus, the adjective "cognitive" was used with words to express the mental process of understanding humanities case, as in the sentence "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". The adjective "cognitive" was also used with the words studies, system, science to indicate general research terms. Meanwhile, the use of adjectives "cognitive" with noun were not used in two other corpora (Health Sciences and Social Politics). Thus, the semantic preference regarding the collocations adjective + noun also could not be found in these corpus fields.

2. Different

The use of adjectives different in the Humanities corpus was to differentiate the human communicative system. Meanwhile, the use of adjective different in the Health Sciences corpus was to show the distinction of medical procedures, instruments, or even risks. Examples of some of these combinations are different kinds of semiotic units and different. kinds of semiotic system in the Humanities corpus, and different types of surgical procedures, different types of ventilators in Health Sciences. Associations between different and the words: kinds and types were found in the journal articles from both areas, especially in the discussion section. Whereas, the use of adjective "different" with the nouns were not found in the Social Politics corpus. To sum up, there are some similarities and differences concerning the semantic preferences in the Humanities and Health Sciences corpus. In both corpora, the adjective "different" and its combinations were mainly used in the discussion section of the journal articles. Meanwhile, the differences were on the use of adjectives "different" in the domain of each discipline.

3. High

The use of adjective "high" with noun was only used in the two corpora (Health Sciences and Social Politics) to refer to measurement and quantification, even though this use is much more frequent in the Health Sciences corpus to indicate the degree of health care systems. The associations with high in the

Health Sciences were found in the introduction section, while, the associations with high in the Social Politics were frequently found in the results and discussion sections. In conclusion, the adjective “high” in both corpora were similar in terms of indicating the measurement or even quantification. Meanwhile, the differences were the use of the adjective in the part of journal articles of both corpus. In the Health Sciences corpus, the adjective “high” was mainly used in the introduction section, whereas, in the Social Politics corpus was used in the section of result and discussion. In addition, the occurrence of the adjective “high” in the Health Sciences corpus is more frequent than Social Politics corpus.

CONCLUSION

Based on the five adjectives selected, there are only three adjectives that collocate with the nouns, namely cognitive, different, and high. The spread of these adjectives is also not uneven in the journal articles of the three corpora. In the Humanities corpus, there is the use of adjectives “cognitive” and “different”. In the Health Sciences corpus, there are the use of adjectives “different” and high". Meanwhile, in the Social Politics Corpus, there is only the use of adjective “high”. The use of the adjectives, especially “different” and “high” show the similarities and differences in the three corpus. In the corpus of Humanities and Health Sciences, the adjective “different” and its combination is similarly used in the results and discussion section of the journal articles. Meanwhile, the differences of the adjective “different” are in the domain of each discipline; in the Humanities corpus, the adjectives “different” is used to indicate the distinction of human communicative system, whereas, in the Health Sciences, the adjectives “different” is frequently used to express the various types of medical procedures, risks, or even medical instruments. In the Health Sciences and Social Politics corpus, the similarities of the adjective “high” in both corpus are in terms of expressing the measurement and quantification. While the differences are the use of the adjective in the part of journal articles of both corpus. In the Health Sciences corpus, the adjective “high” was mainly used in the introduction section, whereas, in the Social Politics corpus was used in the section of result and discussion.

REFERENCES

Begagic, M. (2013). Semantic Preference And Semantic Prosody of The Collocation Make Sense. *Jezikoslovlje Journal*, 3: 403-406. doi: http://hrcak.srce.hr.

Stubbs, M. (2001). Text, corpora, and problems of interpretation: A response to Wid-dowson. *Applied Linguistics*, 22: 149–172. doi: https://academic.oup.com.

Selmitraitis, L. (2020). Semantic Preference, Prosody and Distribution of Synonymous Adjectives in COCA. *GEMA Online: Journal of Language Studies*, 3: 1-18. doi: https://ejournal.ukm.

https://academic.oup.com/journals/pages/arts_and_humanities

https://academic.oup.com/journals/pages/social_sciences

https://academic.oup.com/journals/pages/medicine_and_health