



Research article

Why do Words with Negative Connotations Still Exist? A Corpus-Based Analysis of the Words 'Handicapped', 'Diffable', and 'Disability'

Yoga Yolanda^{1*} , Budi Setyono²

¹ Study Program of Indonesian Language and Literature Education, Faculty of Teacher Training and Education, University of Jember, Jember, East Java, Indonesia. *Corresponding author.

² Study Program of English Education, Faculty of Teacher Training and Education, University of Jember, Jember, East Java, Indonesia.

Abstract

This corpus-based study examines the persistence of negatively connoted words in the Indonesian, particularly focusing on *cacat* (handicapped). *Cacat* is compared to its synonyms, namely *difabel* (difable) and *disabilitas* (disability). The study employs a mixed-methods approach, using data from Indonesian corpora, specifically 'ind_mixed_2013' and 'Korpus Indonesia.' The analysis results indicate a gradual transition from the use of the word *cacat* to *disabilitas* in discussions about human conditions, while *cacat* still retains important metaphorical meanings in specific contexts and is irreplaceable. Recommendations encompass a review of language term absorption rules in Indonesian, stipulating that new words must be euphemistic and free from negative connotations, to be undertaken by the government.

Keywords: handicapped, difable, disability, synonyms, negative connotations



Peace, Justice and Strong Institutions

Introduction

Synonyms cannot replace each other in all contexts (Ajmal, Kumar, Ritonga, & Nukapangu, 2022). One of the reasons is that each synonymous word has its connotative meaning. Connotative meaning is the meaning possessed by a word or group of words based on feelings or thoughts (Herianah et al., 2022). An example is the word *blantik* (Masrurin, 2021). In Indonesia, *blantik* (broker) is a synonym for the words *calo* and *makelar*. Both *blantik*, *calo*, and *makelar* have their

Article History: Received: 15 September 2023. Revised: 08 December 2023. Accepted: 18 December 2023. Published: 19 December 2023.

Copyright: © 2023 by the author/s. License Aesthetix Media Services, India. Distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Citation: Yolanda, Y. & Setyono, B. (2023) Why do Words with Negative Connotations Still Exist? A Corpus-Based Analysis of the Words 'Handicapped', 'Diffable', and 'Disability'. *Rupkatha Journal* 15:4. <https://doi.org/10.21659/rupkatha.v15n4.15>

connotations. Therefore, on a topic of conversation, the appropriate word could be only one of the three words, such as the word *blantik* which is not considered suitable to refer to a person who mediates the sale of a house.

The meaning of a word or group of words can differ from one community group to another. Candra et al. (2018) say that connotative meanings can change based on culture and historical period. For translators, connotative meaning can pose a bigger problem than denotative meaning. Furthermore, a word can have different, even opposite, connotations for different groups of people. For example, the word concept of *anjing* (dog). Dog in Malay society has a negative connotation (Imran, 2011). In Arabic and the UK, the concept of a dog can have both negative and positive connotations (Sameer, 2016). In Western, the concept of a dog is even more likely to have a positive connotation (Lakoff & Turner, 1989).

In establishing social relationships, people tend to choose words that are more subtle or not taboo (Septiana & Rahmawati, 2021). Taboo words and words that indicate shame, fear, and dislike are words that have negative connotations. Therefore, people will try to be more euphemistic by choosing words with more positive connotations. For example, the news media *detik.com* in its uploads on Twitter chose the word *bui* (prison) rather than *penjara* (prison) or *sel* (prison), then chose the word *dicopot* (removed) or *dipulangkan* (repatriated) rather than *dipecat* (fired) or *di-PHK* (laid off) (Sinambela & Mulyadi, 2020).

Similar to the synonymous words in the examples above, in Indonesia there are the words *cacat*, *difabel*, and *disabilitas*. As synonyms, all three have their own connotations. In many parts of the world, using the word *cacat* (handicapped) is a crime. Likewise, in Indonesia, the word handicapped has the most negative connotations compared to the other two words. In fact, its use to refer to humans is considered inappropriate. One case occurred in 2020, namely regarding the mention of the word disability in the Job Creation Law. In full, the article reads, "What is meant by "supporting facilities" include facilities for the disabled, health facilities, public facilities, worship facilities, health posts, police posts, and fire extinguishers." In an article on *kompas.com*, Slamet Thohari, Chair of the Australia-Indonesia Disability Research and Advocacy Network (AIDRAN), said that the use of the phrase "people with disabilities" is a paradigm crime (Kamil, 2020). From this case, it can be said that people agree that the word disability is more disliked than the other two words. In fact, in the General Guidelines for Term Formation (PUI) published by the Ministry of National Education (2007, pp. 10–11), in the requirements for the formation of new Indonesian terms, terms with good connotations must be chosen. In other words, *cacat* and other words with negative connotations should have no place in society.

In the Indonesian context, various studies have been conducted to uncover comparisons of words that refer to disability, but there has been no research based on corpus data (see Ashadi & Premasari, 2020; Bolt, 2005; Devlieger, 1999; Haller, Dorries, & Rahn, 2006; Maftuhin, 2016; Pepper, 2016). In fact, using large empirical data will find areas of use of synonymous terms, and can reveal the reasons why negative connotation words still exist in society, as well as whether the government's steps to avoid negative connotation words in the formation of new Indonesian terms are appropriate.

By using a corpus linguistics approach in which there are frequency, collocation, concordance, n-grams, semantic preference, and prosodic semantics, synonymous words can be found similarities

and differences in their use (Ajmal et al., 2022; Jafarpour, Hashemian, & Alipour, 2013; Kruawong & Phoocharoensil, 2022; Phoocharoensil, 2021). The purpose of this study is to compare the use of synonymous words in Indonesia using a corpus linguistic approach. The synonymous words used in this study are *cacat*, *difabel*, and *disabilitas*. Through analysis of frequency, semantic prosody, collocation, and concordance, the frequency of use as well as the similarities and differences in the use of the three words were obtained.

Connotations of Disability Terminology

There are many studies that explore *the battle* of the words *cacat* (handicapped), *difabel* (difable), and *disabilitas* (disability). Broadly speaking, these studies use a social approach and highlight the struggle for disability rights. Pepper (2016) said that the term for people with disabilities has shifted. In his search, he found various references, namely *delicate*, *handicapped*, *invalid*, *flid*, *retard*, *spaz*, *the other*, *disabled people*, & *people with disabilities*. Pepper himself is more inclined to agree with the term *disabled person*, while in the UK, the term that tends to be used is *disabled people*. The basis is that, in the Social Disabilities Model, people with disabilities see themselves as people who experience barriers in society (Pepper, 2016).

Pepper's exploration above is done from the perspective of people with disabilities themselves. He discusses that changes in terms occur because they are based on painful feelings that people with disabilities feel towards existing terms. On the other hand, in 1999, Devlieger in his research using historical and semiotic approaches mentioned that in the United States, there were changes in terminology that caused differences in understanding disability as a phenomenon. *Cacat* is closely related to access, while *disabilitas* is closely related to ability. The use of *cacat* reflects an era in which accessibility was central to understanding and practice, while the use of *disabilitas* transforms it into individual ability (Devlieger, 2009).

Still in the context of the United States, Haller et al. (2006) conducted research on the term *disability* after the passing of the Law on Persons with Disabilities in the United States. The study traces how the word *disability* is used by journalists in the news media. Through content analysis of The Washington Post and The New York Times media uploads for a decade, it was found that the terminology of disability or the identity of the disability community continues to be formed, transformed, and maintained through news media exposure (Haller et al., 2006).

A study entitled *From Blindness to Visual Impairment: Terminological Typology and the Social Model of Disability* (Bolt, 2005) highlights the change in terms from *buta* (blindness) to *gangguan penglihatan* (visual impairment). Based on three terminologies: *ableism*, *disablement*, and *impairment*, Bolt proposes three hierarchical phases of change: *first*, anachronistic, *second*, regressive, and *third*, progressive. This hierarchical phase is the basis for terms such as *buta/kebutaan* (blindness) & *orang buta* (the blind) being rejected because they also denotatively mean *gangguan kebutaan* (visual impairment), a progressive term that fits the insights in The Social Model of Disability.

As I mentioned earlier, these studies use a human rights perspective to track changes in terms referring to disability. In general, words that already have negative connotations—tend to offend people with disabilities—are rejected or abandoned by society.

In Indonesia, Ashadi & Premasari (2020) said that people do not understand the difference between disability and handicapped. The societal perspective is only based on the cause: birth defect or accident effect. From the research, it can be understood that when someone uses the term disability, he may not intend to offend, it could be just because he does not understand the negative connotations that the term carries.

In addition, research that comprehensively discusses *cacat*, *difabel*, & *disabilitas* is a study entitled *Mengikat Makna Diskriminasi: Penyandang Cacat, Difabel, dan Penyandang Disabilitas* (Maftuhin, 2016). Using language data from *google.com*, this research reveals the fact that in the academic context, *penyandang disabilitas* (people with disability) compete fiercely to replace *penyandang cacat* (handicapped). In the future, optimistically, *penyandang disabilitas* will replace *penyandang cacat*. In addition, the term *difabel* will remain a second alternative term.

The tendency of society to use words with positive rather than negative connotations does not necessarily make words with negative connotations disappear. Both positive and negative connotation words still coexist. When people prefer the term *makelar* (broker), it does not mean that *calo* and *blantik* will disappear. The same goes for other words. Although they are synonyms, the words *cacat* (handicapped), *difabel* (difable), and *disabilitas* (disability) have their own connotations. *Cacat* is perceived more negatively than the other two. Although Maftuhin (2016) says that in the future *disabilitas* will replace *cacat*, the reality is that until now *cacat* is still used by the community.

Various research results provide more or less the same information: words that indicate offending or hurting people with disabilities will be abandoned. What is not captured in the various studies above is the reason why *cacat*, which has a negative connotation, still exists in society. Ashadi & Premasari's research (2020) concluded that public ignorance about the difference between *cacat* and *disabilitas* in Indonesia can be used as an answer, but since another study (Maftuhin, 2016) said that the word *cacat* in the future will be replaced by *disabilitas*, there needs to be another approach as a form of triangulation of this phenomenon. In addition, it is also necessary to trace the meaning of the three terms based on data on their actual use in the community to reveal why people use one term rather than the other.

Corpus-based Research on Synonymous Words

Previously, there has been no research on the words *handicapped*, *difable*, & *disability* using a corpus linguistic approach. Nevertheless, some studies were found using corpus linguistics to analyze synonymous words. Phoocharoensil (2021), through an analysis of prosodic semantics and collocation, found differences in the near-synonyms *persist* and *persevere*. The research shows that *persistence* tends to co-occur with words or phrases that express negative meanings associated with persistently unpleasant situations, whereas *perseverance* appears in contexts that show strong determination and great effort in completing difficult tasks. In addition, Kruawong & Phoocharoensil's research (2022) through a corpus-based approach also found similarities and differences in the use of the words *teach*, *educate*, & *instruct*. The findings of this study revealed that from the distribution pattern among text types, *teach* was much more widely and commonly used than *educate* and *instruct*. The frequency data also revealed that the three synonyms were

preferred in formal genres rather than in spoken conversations, such as in academic literature. Chaengchenkit (2023) explored the differences between three synonymous verbs, namely *cease*, *heat*, & *stop*. This study found that the words *cease* & *heat* appear more frequently in formal genres than *stop*, while *stop* appears more frequently in the TV & Movie Subtitles genre (non-formal). In addition to these three synonymous verb studies, Ajmal et al. (2022) explored the adjectives *beauty*, *handsome*, & *pretty*. Through a corpus linguistics approach, the researchers found similarities & differences in the use of the three synonymous adjectives. The three words are similar in core meaning, but different in detail meaning and collocation and of course they are not interchangeable in all contexts.

The use of corpus linguistics is important to explore the phenomenon of using synonymous words based on real (empirical) language data. Corpus linguistics allows researchers to reveal cultural and social information through typical analysis (Yuliawati, 2018). The use of corpus linguistics in this study is based on the assumption that in researching language there are two principles, namely (1) language is all about the creation of meaning and (2) language tends to be phraseological (Sinclair, 1991). Phraseology means that the meaning of words is basically formed from repeated associations based on the agreement of language users in a culture (Cheng, 2011).

That is what underlies a word does not pocket its meaning in isolation, but the word is meaningful when it associates or appears together with other words in a series of language structures with a large frequency so that it shows uniqueness (McEnery & Hardie, 2011). That's why in corpus linguistics there are concepts of collocation, semantic preference, and semantic prosody. Collocations are formulaic units of language that have become prominent in our understanding of language learning and use (Gablasova, Brezina, & McEnery, 2017). Collocation is a part of semantics that describes a part of a cultural field or reality in a particular universe that is realized by a set of lexical elements whose meanings are related (Iskandar, 2018). In simple terms, collocation can be referred to as a word that appears between other words (word associations or meaning fields). Furthermore, Semantic Preference is the relation between related words in the meaning field (Lindquist, 2009). For example, the word *money* often appears together with the words *pay*, *buy*, *income*, *salary*, and so on. Then, semantic prosody is a term that arises based on the understanding that a word is very dependent on the words that accompany it (Yuliawati, 2016). As a result, a word can have a negative connotation if associated with another word, but can have a positive connotation if associated with a different word. Through the concepts in corpus linguistics, this study aims to compare the frequency of use, the meaning of synonymous words based on semantic preference, and the differences in the use of synonymous words in Indonesia. The synonymous words used in this study are *cacat*, *difabel*, and *disabilitas*.

ind_mixed_2013 and Korpus Indonesia

In comparing the words "cacat," "difabel," and "disabilitas," two Indonesian language corpora were utilized: (1) ind_mixed_2013 by the Leipzig Corpora Collection (LCC) and (2) Korpus Indonesia (Koin) by the Ministry of Education, Culture, Research, and Technology (Kemdikbudristek). The selection of ind_mixed_2013 was based on its collection of 1.2 billion tokens gathered from various

texts on the internet, encompassing a wide range of written Indonesian language varieties, including fiction, nonfiction, scholarly papers, and popular articles. This collection of words was obtained up to 2013, necessitating the use of Korpus Indonesia, which was updated in October 2022 and contains 24,736,534 tokens. By utilizing these two corpora, the empirical data required for this research is abundant.

This study employs a mixed-method approach. The strategy used is sequential explanatory, which involves quantitative data collection, followed by quantitative data analysis, subsequent qualitative data collection, qualitative data analysis, and concludes with an overarching interpretation of the analysis (Creswell & Creswell, 2018).. The initial step involves revealing the frequency of word occurrences for the purpose of measuring the closeness of one word to another, also known as the level of collocational significance. The statistical test employed is Log-likelihood in the ind_mixed_2013 corpus by LCC and Koin by Kemdikbudristek. Subsequently, the statistical test provides data for qualitative analysis of word collocates and semantic preferences for the words "cacat," "difabel," and "disabilitas." The basis used for semantic categorization is The UCREL semantic analysis system (USAS). USAS is essentially an application for performing automatic semantic analysis with high accuracy (Rayson, Archer, Piao, & McEnery, 2004). While this application has been well-tested in numerous studies for English and offers various language versions (Ji, 2018), in this research, USAS is utilized as the basis for identifying semantic categories (word clustering) and subsequently describing the usage patterns of these words. The utilization of USAS as the basis for identifying semantic categories has been previously performed by Yuliawati (2018) & Goyak et al. (2021). Semantic preference analysis is conducted using collocation features in the corpora utilized. Additionally, for word clustering analysis, n-Gram and Concordance features are employed within the corpus application. During qualitative analysis, the assumptions of Cognitive Semantics (Evans & Green, 2018; Kovecses, 2010; Lakoff & Turner, 1989) are used.

Frequency

The frequency of the words *cacat*, *difabel*, and *disabilitas* in two Indonesian corpora, i.e (1) ind_mixed_2013 by the Leipzig Corpora Collection (LCC) and (2) Korpus Indonesia (Koin) by the Ministry of Education, Culture, Research, and Technology (Kemdikbudristek) shows differences. In LCC which consists of 1,206,281,985 tokens, the word *cacat* appears 41.036 times or 34,01 times per one million tokens, *difabel* 4.964 times or 4,11 times per one million words, and *disabilitas* 5.154 words or 4,27 times per one million tokens.

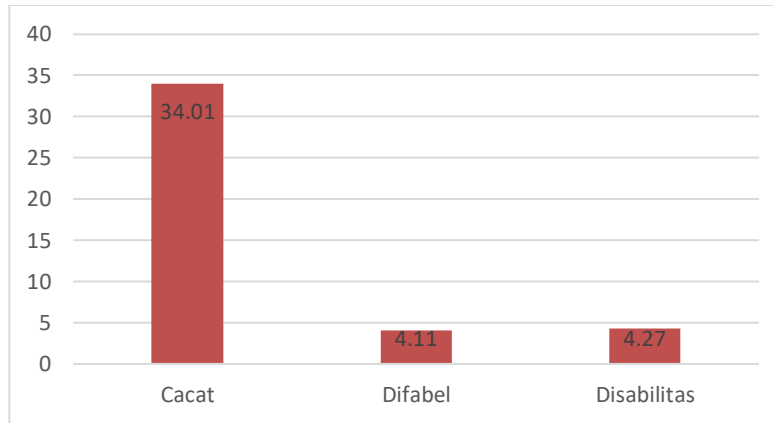


Figure 1: Frequency of occurrence of the words *cacat*, *difabel*, and *disabilitas* per one million words in the ind_mixed_2013 corpus

In Koin, which consists of 24.018.869 tokens, the word disability appears 686 times or 28,56 times per one million words, difabel appears 110 times or 4,57 times per one million words, and disability appears 1.239 times or 51,58 times per one million words (Figure 2).

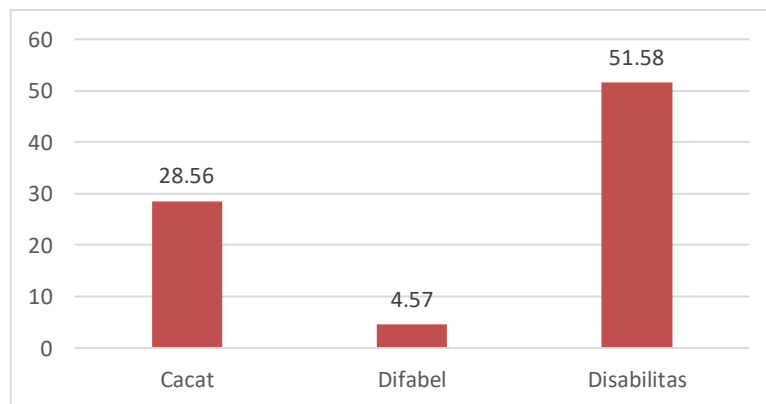


Figure 2: Frequency of occurrence of the words *cacat*, *difabel*, and *disabilitas* per one million words in the Koin

Collocates and semantic preferences

Despite the dynamic usage, the words *cacat*, *difabel*, and *disabilitas* are still productively used by society. Then, based on community usage, what is the meaning of the three synonymous words? To answer this, I utilize the concepts of collocation and semantic preference. Collocation refers to a word that appears among other words (word association or meaning field), the word appears in a sentence together with other words, while semantic preference is the relation between related words in the meaning field (Lindquist, 2009).

In this study, in identifying the semantic preference of the words *cacat*, *difabel*, and *disabilitas*, the collocates of each word are grouped using the USAS tagset. The USAS grouping was inspired by Tom McArthur's Longman Lexicon of Contemporary English (McArthur, 1982). This tagset has a nested structure with 21 main discourse areas which are further subdivided and with the possibility

of more detailed subdivisions in some specific cases. Each collocate found will be categorized into one of these 21 major discourse areas.

Cacat

Using the LCC software, the most frequent collocates of problematic words in the ind_mixed_2013 corpus were identified. LCC automatically calculates significance using Log-likelihood, but the software does not display the frequency of word encounters with their collocates. In contrast to LCC, the Indonesian corpus provides not only several techniques for calculating significance statistics but also the number of word encounters with their collocates.

Furthermore, using the Koin tool, the top 50 words that collocate with the word "cacat" were collected. These words have statistical significance in accompanying the word "cacat." The range used to generate the above tables consists of 5 words to the right and 5 words to the left of the word "cacat" in a sequence of words spoken or written by people.

The collected collocate data, when grouped based on the USAS tagset, can be categorized into 13 groups. Of the 13 categories, there is 1 category that only has 1 significant collocate with the word defect so that 1 category cannot be considered semantically related to the word defect. This is based on Stubbs' statement (Bernardini & Aston, 2002) that semantic preference is not a relationship between individual words, but rather a relationship between lemmas or other sets of words that have a semantic relationship. Therefore, the word defect has semantic preferences concerning 1) GENERAL AND ABSTRACT TERMS (primarily with respect to the existence, quality, quantity, and function); 2) THE BODY AND THE INDIVIDUAL (particularly with respect to anatomy and physiology as well as health and disease); 3) EMOTION (mainly about the emotional process); 4) GOVERNMENT AND THE PUBLIC DOMAIN (particularly the public sphere, work, and law); 5) ARCHITECTURE, HOUSING AND THE HOME (especially regarding construction or building materials); 6) MONEY AND COMMERCE (especially regarding work and business); 7) LIFE AND LIVING THINGS; 8) NUMBERS AND MEASUREMENT (especially regarding size and order); 9) SUBSTANCES, MATERIALS, OBJECTS AND EQUIPMENT (about basic materials); 10) SOCIAL ACTIONS, STATES, AND PROCESSES (about caring for each other); 11) TIME (regarding age); 12) PSYCHOLOGICAL ACTIONS, STATES, AND PROCESSES (especially regarding states and abilities).

Difabel

Based on the top collocate data of the word "difabel" in the two corpora, there are 12 semantic preferences related to the word *difabel*. There is 1 category that only has 1 word collocated with *difabel*, thus leaving 11 categories.

From the data presented, the semantic preference of the word *difabel* is related to 1) GENERAL AND ABSTRACT TERMS; 2) BODY AND INDIVIDUAL (mainly about anatomy, physiology, health, and individual in general); 3) GOVERNMENT AND PUBLIC DOMAIN (about law and government in general); 4) ARCHITECTURE, HOUSING AND THE HOME (mainly about house/building parts, area, and furniture); 5) SPORT (about role in sport); 6) MOVEMENT, LOCATION, TRAVEL, AND TRANSPORTATION (mainly about location/place and movement); 7) NUMBERS AND

MEASUREMENT (mainly about numbers, distances, and measurement in general); 8) EDUCATION (mainly about educational institutions and roles in education); 9) SOCIAL ACTIONS, STATES, AND PROCESSES (mainly about personality traits, community organization or participation, and power relations); 10) TIME (mainly related to age); and 11) PSYCHOLOGICAL ACTIONS, STATES, AND PROCESSES (related to abilities).

Disabilitas

After obtaining the collocates and semantic preferences of the words *cacat* and *difabel*, the next step was to trace the collocates and semantic preferences of the word *disabilitas*. The same thing was done, i.e tracing the top collocates of the word *disabilitas* through the *ind_mixed_2023* and *Koin*.

There are 15 categories that show the regions or domains of words that collocate with the word disability. Among these 15 categories, there are 4 categories that only contain 1 word that collocates with the word disability. Thus, 11 domains remain, i.e. 1) GENERAL AND ABSTRACT TERMS (especially regarding judgment/evaluation, degree, and opportunity); 2) BODY (regarding anatomy and physiology); 3) EMOTIONS (emotional states and actions); 4) GOVERNMENT AND PUBLIC DOMAIN (government institutions and officials); 5) ARCHITECTURE, HOUSING AND THE HOME (about home and facilities); 6) MOVEMENT AND TRAVEL; 7) NUMBER (about size, and level); 8) LINGUISTIC ACTIONS AND PROCESSES (concerning communication, language and language acts); 9) SOCIAL ACTIONS, STATES, AND PROCESSES (concerning needs, concerns, participation, power, and power relations); 10) TIME in general and age; 11) PSYCHOLOGICAL ACTIONS, STATES, AND PROCESSES (i.e. concerning mental acts, abilities, intelligence, attention, and sensory (taste and vision)).

Existence of the words *cacat*, *difabel*, & *disabilitas*

Based on the two corpora whose data are shown in Figures 2 & 3, there is a clear difference. The words *cacat* & *disabilitas* have significant differences, while the word *difabel* tends to have no difference. Given that the *ind_mixed_2013* corpus data was collected up to 2013 and the *Koin* data was collected up to October 2022, the difference in the frequency of word usage can be attributed to real changes that occur diachronically, for example due to socio-cultural developments in society that can determine the level of iconicity of a word over time (Monaghan & Roberts, 2021).

The results of tracking the popularity of these three words reinforce Maftuhin's (2016) statement that in the future, optimistically, *disabilitas* will replace *cacat*. Maftuhin's research used data in the form of texts from Google and was published in 2016. However, the phrase "replace" is not entirely true. In reality, although the word *cacat* loses its position to the word *disabilitas*, *cacat* still has a large frequency. This means that the word has not disappeared or been marginalized. Through this corpus-based research, the word disabled appears to have decreased in frequency. From 34,01 per million words (up to 2013) to 28,56 per million words (up to 2022). This decrease proves that the popularity of *cacat* in society is decreasing. On the other hand, *disabilitas* experienced a

significant increase, from 4,27 occurrences (up to 2013) to 51,58 occurrences per one million words. This increase indicates that the popularity of the word *disabilitas* is getting higher in society. In contrast to the other two words, *difabel* did not experience any significant change, from 4,11 (up to 2013) to 4,57.

Through the data above, it appears that there is a change in the level of popularity of the three synonyms. In the older corpus data (*ind_mixed_2013*) *cacat* is still much more popular than the word *disabilitas*, while in the newer corpus data--*Koin*--*disabilitas* has increased beyond *cacat*. In contrast to the two words that alternate places, the word *difabel* still remains in the third position, both in *ind_mixed_2013* and *Koin*. This is in line with what Maftuhin (2016) said that the term *difabel* will remain the second alternative term.

The meaning of *cacat*, *difabel*, & *disabilitas*

In the previous section, through an exploration of the collocates and semantic preferences of the words *cacat*, *difabel*, & *disabilitas*, it can be seen how each of these words is used in the language activities of the community. Looking back, in many contexts, the three words are indeed interchangeable. However, there are also differences that make them not interchangeable. To explore both the similarities and differences between the three words, concordance or KWIC (Key Word in Context) analysis is used. Concordance is an important analysis technique in corpus studies because it allows a large number of instances of word usage to be brought together in one place, in their original context (Evison, 2010). The concordance feature in the corpora application makes it easier for users to distinguish between lemma and sublemma words (Almos, Pramono, Seswita, Asma, & Putri, 2023). In addition to using concordance, n-Gram is also used. n-Gram is used to obtain statistical data on the clusterization of the three words under study. n-Gram presents a collection of word groups detected in a text. N-Grams are capable of displaying various word clusters ranging in length from 2 words (2-Grams) to 4 words (4-Grams). n-Grams are valuable for various purposes, such as improving the accuracy of speech recognition, spell checking, or machine translation systems (Roziewski & Kozłowski, 2021). Besides presenting a list of word groups, the n-Gram table also provides information on the frequency of occurrence of the word group in the corpus as well as the number of texts containing the word group.

***Cacat*, *difabel*, & *disabilitas* are interchangeable when referring to human**

Based on collocation and semantic preference analysis, it is known that people use the three words interchangeably as synonymous words in general. However, not all semantic categories found in each word are interchangeable, the similarity in the use of the three words lies in 8 semantic categories, namely 1) GENERAL AND ABSTRACT TERMS; 2) body and individual; 3) GOVERNMENT AND PUBLIC DOMAIN; 4) ARCHITECTURE, BUILDING, AND HOUSE; 5) NUMBERS AND MEASUREMENTS; 6) SOCIAL ACTIONS, STATES, AND PROCESSES, 7) TIME; and 8) PSYCHOLOGICAL ACTIONS, STATES, AND PROCESSES.

For example, the words *cacat*, *difabel*, & *disabilitas* are interchangeable when talking about an individual's body. The following concordance is obtained from the *Koin*.

<i>Memperoleh layanan khusus bagi penyandang</i>	<i>cacat</i>	<i>sesuai dengan sarana dan prasarana yang</i>
(Obtaining special services for people with	handicapped	in accordance with the facilities and infrastructure that)
<i>.... seperti bayi, anak kecil atau penyandang</i>	<i>difabel</i>	<i>.</i>
(.... such as infants, young children, or people with	difable	.)
<i>.... lebih cepat mengasosiasikan penyandang</i>	<i>disabilitas</i>	<i>dengan atribut negatif dibandingkan dengan ...</i>
(.... is quicker to associate people with	disability	with negative attributes compared to ...)

The concordance above shows the use of the words *cacat*, *difabel*, & *disabilitas* to talk about individuals, especially about human health, anatomy, or physiology. The three data displayed above show that both *cacat*, *difabel*, & *disabilitas* collocate with the word *penyandang* (people with). However, through the n-Gram (2-Gram), the words *cacat* and *disabilitas* are mostly found paired with the word *penyandang*, but not with the word *difabel*. Of the 95 words that accompany the word *penyandang*, the phrase *penyandang disabilitas* is in first place with a frequency of occurrence of 109, the phrase *penyandang cacat* is in second place with a frequency of occurrence of 32, while the phrase *penyandang difabel* is in 75th place with a frequency of occurrence of 1. In addition, through further investigation, it is evident that the words *cacat* and *disabilitas* are substituted more often in many contexts that talk about individuals than *difabel*. Despite this condition, it is evident that society uses all three in the context of THE BODY AND THE INDIVIDUAL (human).

The next high level of similarity is found in the semantic category of social actions, states and processes, such as talk about rights. The following concordance data shows these similarities.

<i>4 pemilih tunanetra serta penyandang</i>	<i>cacat</i>	<i>lainnya dalam memberikan hak pilihnya....</i>
(4 blind voters and other people with	handicapped	in exercising their right to vote....)
<i>.... ketentuan ini dimaksudkan agar penyandang</i>	<i>cacat</i>	<i>anak memperoleh tiga hak berikut ini.</i>
(... This provision is intended to give children with	handicapped	disabilities the following three rights .)
<i>....bekasi mulai 2021 diberi tunai penuhi hak</i>	<i>difabel</i>	<i>"serta digali juga mengenai besaran nilai</i>
(....bekasi from 2021 is given cash to fulfill the rights of the	<i>difable</i>	"and also explored the value of)
<i>....Kaum</i>	<i>difabel</i>	<i>mempunyai hak yang sama dalam hal</i>
(.... People with	Difable	have equal rights in terms of)
<i>Melindungi hak penyandang</i>	<i>disabilitas</i>	<i>untuk mencalonkan diri dalam pemilihan,</i>

(Protecting the right of people with	disability	to stand for election,)
.... kerja wajib menjamin agar penyandang	disabilitas	dapat melaksanakan hak berserikat dan
(.... work must ensure that persons with	disabilitas	can exercise their rights of association and)

In discussing *hak* (rights), *cacat*, *difabel*, & *disabilitas* are interchangeable. However, the occurrence of the word *hak* as a collocate of the word *cacat* is not significant. This statement is based on the difference between the words *cacat*, *difabel*, & *disabilitas* in the semantic category of *social actions, states, and processes*. This illustrates that in social conversations, the word *cacat* is used less frequently than the other two words.

CACAT	DIFABEL	DISABILITAS
<i>santunan, rehabilitasi, orang</i> (compensation, rehabilitation, people)	<i>ramah, orang, kaum, hak, hak-hak, diskriminasi, kelompok, masyarakat, kekerasan, komunitas, perempuan</i> (friendly, people, people group, right, rights, discrimination, group, society, violence, community, women)	<i>rehabilitasi, pemberdayaan, perlindungan, intervensi, hak-hak, stigma, perlindungan, kepedulian, sosial, kuat, diskriminasi, masyarakat, partisipan, partisipasi, pelayanan, kesetaraan, kontak, kebutuhan khusus</i> (rehabilitation, empowerment, protection, intervention, rights, stigma, protection, awareness, social, strong, discrimination, community, participant, participation, services, equality, contact, special needs)

Table 1: Wordlist in the Semantic Categories of SOCIAL ACTIONS, STATES, AND PROCESSES

Table 1 provides information that the semantic preference of the word *cacat* with the semantic category of SOCIAL ACTIONS, STATES, AND PROCESSES is not as high as the other two words. There are only 3 collocates of the word *cacat* in this category, 11 collocates of the word *difabel*, and 18 collocates of the word *disabilitas*.

Discussions about *cacat*, *difabel*, & *disabilitas* will not be far from rights, law, protection, services, and discrimination. The small number of collocates of the word *cacat* in this context compared to the other 2 words could be because people are more careful or consider the sensitivity of the word chosen or what kind of impact occurs due to the choice of words used. In other words, people think about the feelings of people with disabilities in choosing terms. This is in line with what Pepper (2016) said that changes in terms occur because they are based on painful feelings felt by people with disabilities towards existing terms. With these results, Ashadi & Premasari's (2020) statement that people do not understand the difference between *cacat* & *disabilitas* cannot be proven. Empirical data suggests that the word disability for social matters tends to be avoided by the community.

Thus, the use of the words *cacat*, *difabel*, and *disabilitas* have many similarities in meaning that appear in semantic preferences. The 8 semantic categories indicate that the three words are interchangeable. However, people's choices are not always neutral. In certain speech contexts,

one word tends to be avoided over the other two, and vice versa, one word tends to be chosen over the other two.

***Cacat* is irreplaceable when used metaphorically**

Through an examination of collocations and semantic preferences, differences were found between the words *cacat*, *difabel*, and *disabilitas*. The differences appear as follows.

KATEGORI SEMANTIK	CACAT	DIFABEL	DISABILITAS
EMOTION	<i>diproses, disebabkan, mendapatkan, menyebabkan, dikhawatirkan, mental</i> (processed, caused, get, cause, worry, mental)	-	<i>perasaan, empati</i> (feelings, empathy)
MONEY AND COMMERCE	kerja, produk, jual, beli	-	-
LIFE AND LIVING THINGS	<i>pertumbuhan, bakteri</i> (growth, bacteria)	-	-
SUBSTANCES, MATERIALS, OBJECTS AND EQUIPMENT	<i>batu, kayu, folat, kayunya</i> (stone, wood, folate, the wood)	-	-
ENTERTAINMENT, SPORTS AND GAMES	-	<i>atlet, olahraga</i> (atlet, sport)	-
MOVEMENT, LOCATION, TRAVEL AND TRANSPORTATION:	-	<i>akses, solo, roda, keterbatasan</i> (access, solo, wheels, limitations)	<i>aksesibilitas, jarak, perjalanan</i> (accessibility, distance, travel)
EDUCATION	-	<i>pendidikan, SLB, PPRBM, UB, FKDB, mahasiswa</i> (education, SLB, PPRBM, UB, FKDB, university students)	-
LANGUAGE AND COMMUNICATION	-	-	<i>Komunikasi, Bahasa, Berbicara, Wicara</i> (communication, language, speaking, speech)

Table 2: Differences in semantic preferences of *cacat*, *difabel*, and *disabilitas*

Table 2 provides information on the differences in the semantic categories of the words *cacat*, *difabel*, dan *disabilitas*. In the category of EMOTION, *cacat* appears to be used more often than *disabilitas*, while *difabel* never collocates with words in this semantic category. In the MONEY AND

COMMERCE category, only the word *cacat* is used by the community, while the other 2 words are not. Similarly, in the LIFE AND LIVING THINGS category, *cacat* is used, while the other 2 words are not. In the next category, SUBSTANCES, MATERIALS, OBJECTS, AND EQUIPMENT, only the word *cacat* is significantly collocated with the words in that category, while the other two words are not. Differences are found in the categories of ENTERTAINMENT, SPORTS AND GAMES and EDUCATION. Only the word *difabel* collocated with words in this domain, while the words *cacat* and *disabilitas* did not. In the MOVEMENT, LOCATION, TRAVEL, AND TRANSPORTATION category, the words *difabel* and *disabilitas* are found, but not the word *cacat*. Then, finally, the category of LANGUAGE AND COMMUNICATION. In this category, only the word *disabilitas* has semantic preference.

Based on the data in Table 8, there are three semantic domains or categories that are only related to the word *cacat*, namely MONEY AND COMMERCE; LIFE AND LIVING THINGS; and SUBSTANCES, MATERIALS, OBJECTS AND EQUIPMENT. From this, the question arises why the other two words do not relate to these three categories. From the description of the data, it can be concluded that the reason is that the words *difabel* and *disabilitas* only relate to human-related entities, while the word *cacat* is broader, appearing not only in discussions about humans, but other living things, including inanimate objects. Therefore, in this context, *cacat* is no longer translated as handicapped in English, but defect, defective, partially damaged, and so on. The following concordance data can serve as an explanation of the semantic preference of the word *cacat* in the categories of MONEY AND COMMERCE and SUBSTANCES, MATERIALS, OBJECTS, AND EQUIPMENT.

....dalam hal si kreditur membeli barang yang	cacat	tersembunyi, maka ia mempunyai hak untuk
(.... In the event that the creditor purchases goods with hidden	defects	, he has the right to)
....dunia ini memiliki nilai jual , termasuk barang	cacat	yang justru disebut istimewa karena tidak ada
(.... this world has a selling point, including	defective	goods that are actually called special because there is no ...)
....baret, cat kusam, lampu mati dan kondisi	cacat	sangat menurunkan harga jual .
(.... scratches, dull paint, dead lights and	defective	conditions greatly reduce the selling price .)
....hal barang pesanan yang dikirim salah atau	cacat	maka bprs atau pemasok harus bertanggung
(.... if the ordered goods are sent wrong or	defective	then the bprs or supplier must be responsible)
.... kerugian yang disebabkan karena adanya produk	cacat	maupun karena kerja produk diproses ulang....
.... losses caused by	defective	products or due to product reprocessing work....)

The concordance data above illustrates that the use of the word *cacat* in those sentences cannot be replaced by *difabel* or *disabilitas*. In the context of trade, sellers or traders often use the word

cacat to explain that the goods being sold have flaws. *Cacat* in this case can be interpreted as an *imperfect condition*.

In addition—in relation to the semantic category of LIFE AND LIVING THINGS—*Cacat* can also be interpreted as an *imperfect condition* not only in humans and inanimate substances/objects, but also in life or living things in general. This means that in society, the expression *cacat* used in the domain of animals or plants is still acceptable, for example the expression *hewan cacat* (defective animal) or *cacat pada tanaman* (defect in plants). Of course, this does not apply to the words *difabel* and *disabilitas*.

In ind_mixed_2013 (LCC), the combined words *barang cacat* (defective goods) appear 1745 times and the combined words *cacat produksi* (defective production) appear 1012 times. This large frequency provides evidence that the word defect will survive because its use is needed by the community.

Within the framework of Cognitive Semantic Theory, *cacat* used in expressions such as *barang cacat* (defective goods), *cacat produktif* (defective production), *hewan cacat* (defective animals), & *tanaman cacat* (defective plants) are metaphors. Based on its cognitive function, these expressions are included as ontological metaphors. Yolanda (2022) says that simply put, ontological metaphors are used by people to concretize abstract entities by borrowing other more concrete entities. The word *produk* (product) or *produksi* (production) is an abstract entity so people try to explain it by utilizing a more concrete entity, i.e. *cacat* (defects). Nonmetaphorically, *cacat* is a concrete entity because it relates to the condition of the human body (which is concrete). In addition, expressions that use the word *cacat*, but do not refer to humans, are personifications. Lakoff & Johnson (2003) explain that personification is part of ontological metaphor. Metaphorically, the words *produk* (product), *produksi* (production), *hewan* (animal), and *tumbuhan* (plant) in society's cognition system are regarded as human beings (who can literally be handicapped).

Through the description above, the word *cacat*, which Maftuhin (2016) says will in the future be replaced by the word *disabilitas*, may only apply in general, in overall frequency, but not in certain domains where *cacat* is the preferred choice over the other two words. In fact, in the context of commerce, the word *cacat* shows a high frequency of use to indicate that a substance or object is in a less-than-perfect condition.

Society uses the expression manufacturing *cacat* more often than its literal form, such as, for example, the expression *produksi yang kurang sempurna* (*less than perfect production*). In conceptual metaphor theory, *cacat produksi* (production defects) are categorized as conventional metaphors. This is in line with what Lakoff & Johnson (2003) said, a metaphor that is overused by society will lose its poetic or artistic element so that it is considered literal or not a metaphor. Thus, it can be said that metaphors are a place for words with negative connotations, such as *cacat*, to exist in society.

Indonesia is a big country. 737 regional languages are still actively used by their speakers (Yamin, Setiawan, Anam, & Kurnia, 2020). Language is a reflection of culture (Aitchison & Wardaugh, 1987) and it has been discussed in the introduction that a word can have different connotations in one culture from another. The word *cacat* and other words with negative connotations actually still

exist and are needed by society. Therefore, the term absorption requirement stated in the Pedoman Umum Pembentukan Istilah (General Guidelines for Term Formation) (PUI) (Pusat Bahasa, 2007) that terms absorbed in Indonesian must not contain negative connotations needs to be reviewed because PUI is the main reference for Badan Bahasa in forming or absorbing words in Indonesian. Restrictions on the formation and absorption of words with negative connotations will be a problem in the future. In a very diverse society, with a background of 737 regional languages, it is difficult to agree on what kind of terms really have negative connotations for the whole community. In addition, words with negative connotations are not always used for impolite language purposes. In certain contexts, they are necessary because they cannot be replaced by other terms.

Conclusion

Through this research, it can be concluded that diachronically, the word *cacat* has decreased in frequency of use, the word *disabilitas* has increased in frequency of use, and the word *difabel* has not changed significantly. However, the decrease in the frequency of the word *cacat* only occurs in its literal use, which refers to humans, while in its metaphorical meaning the word *cacat* actually increases in frequency of use, which is in the context of trade, living things (other than humans), and inanimate substances or objects. In fact, the word *cacat* becomes people's first choice when talking about substances or objects that are in a less-than-perfect condition in a trading context. From this fact, the statement that in the future *cacat* will be abandoned by society cannot be justified. *Cacat* will still exist, in its metaphorical use.

The recommendation that can be given based on the results of this research is a review by the government of the term absorption rules in the Pedoman Umum Pembentukan Istilah (General Guidelines for Term Formation) (PUI) which states that terms formed or absorbed in Indonesian must not contain negative connotations. This rule needs to be revised. Indonesia is a vast country with various cultural backgrounds, a word that has a negative connotation in one culture is not necessarily negative in another culture. In addition, both negative and positive connotation terms can be equally needed by society.

Declaration of Conflicts of Interests

The authors declared no potential conflicts of interest.

Funding Disclosure/Acknowledgement

The authors would like to thank Universitas Jember for the funding this research and publication.

References

- Aitchison, J., & Wardaugh, R. (1987). An Introduction to Sociolinguistics. *The British Journal of Sociology*, 38(3).
<https://doi.org/10.2307/590702>

- Ajmal, M., Kumar, T., Ritonga, M., & Nukapangu, V. (2022). A Corpus-Based Analysis of the Adjectives and Synonyms -Beautiful, Handsome, and Pretty. *World Journal of English Language*, 12(2). <https://doi.org/10.5430/wjel.v12n2p159>
- Almos, R., Pramono, P., Seswita, S., Asma, R. A., & Putri, N. O. (2023). Linguistik Korpus: Sarana dan Media Pembelajaran pada Mata Kuliah Leksikologi dan Leksikografi di Perguruan Tinggi. *Lectura: Jurnal Pendidikan*, 14(1), 45–59. <https://doi.org/10.31849/lectura.v14i1.11705>
- Ashadi, & Premasari, A. A. (2020). Handicap and Disability: What is the Difference? *International Journal of Linguistics: Literature and Translation (IJLLT)*, 3(1), 163–167.
- Bernardini, S., & Aston, G. (2002). Review of Stubbs ((2001)): Words and Phrases. Corpus Studies of Lexical Semantics. *International Journal of Corpus Linguistics*, 7(2). <https://doi.org/10.1075/ijcl.7.2.07ber>
- Bolt, D. (2005). From blindness to visual impairment: Terminological typology and the Social Model of Disability. *Disability and Society*, Vol. 20. <https://doi.org/10.1080/09687590500156246>
- Candra, E. N., & Martiarini, E. (2018). Analysis of Connotation Words in The Jakarta Post Online Headline News. *DEIKSIS*, 10(01). <https://doi.org/10.30998/deiksis.v10i01.2194>
- Chaengchenkit, R. (2023). A Corpus-based study of the synonyms cease, halt, and stop. *LEARN Journal: Language Education and Acquisition Research Network*, 16(1).
- Cheng, W. (2011). Exploring Corpus Linguistics. In *Exploring Corpus Linguistics*. <https://doi.org/10.4324/9780203802632>
- Creswell, J. W., & Creswell, J. D. (2018). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches - John W. Creswell, J. David Creswell - Google Books. In *SAGE Publications, Inc.*
- Devlieger, P. J. (1999, July). From handicap to disability: Language use and cultural meaning in the United States. *Disability and Rehabilitation*, Vol. 21, pp. 346–354. <https://doi.org/10.1080/096382899297594>
- Devlieger, P. J. (2009). From handicap to disability: language use and cultural meaning in the United States. <https://doi.org/10.1080/096382899297594>, 21(7), 346–354. <https://doi.org/10.1080/096382899297594>
- Evans, V., & Green, M. (2018). Cognitive linguistics: An introduction. In *Cognitive Linguistics: An Introduction*. <https://doi.org/10.4324/9781315864327>
- Evison, J. (2010). What are the Basics of Analysing a Corpus? In A. O’Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 122–135). New York: Routledge, Taylor & Francis Group.
- Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in Corpus-Based Language Learning Research: Identifying, Comparing, and Interpreting the Evidence. *Language Learning*, Vol. 67. <https://doi.org/10.1111/lang.12225>
- Goyak, F., Muhammad, M. M., Khaja, F. N. M., Zaini, M. F., & Mohammad, G. (2021). Conversational Mental Verbs in English Song Lyrics: A Corpus-Driven Analysis. *Asian Journal of University Education*, 17(1). <https://doi.org/10.24191/ajue.v17i1.12619>
- Haller, B., Dorries, B., & Rahn, J. (2006). Media labeling versus the US disability community identity: A study of shifting cultural language. *Disability and Society*, Vol. 21. <https://doi.org/10.1080/09687590500375416>
- Herianah, H., Garing, J., Agus, N., Jerniati, J., Hastianah, H., & Ridwan, M. (2022). Denotative and Connotative Meanings of Panasuan Language: Minor Language Conservation Efforts in West Sulawesi. <https://doi.org/10.4108/eai.15-9-2021.2315603>
- Imran, H.-A. (2011). Analisis kognitif semantik peribahasa melayu bersumberkan anjing (canis familiaris). *GEMA Online Journal of Language Studies*, 11(1).
- Iskandar, M. M. (2018). Analisis Makna Kolokasi Kata Karada dalam Textbook Shin Chugaku Hoken Taiiku. Retrieved from <http://lib.unair.ac.id>

- Jafarpour, A. A., Hashemian, M., & Alipour, S. (2013). A corpus-based approach toward teaching collocation of synonyms. *Theory and Practice in Language Studies*, 3(1). <https://doi.org/10.4304/tpls.3.1.51-60>
- Ji, M. (2018). A quantitative semantic analysis of Chinese environmental media discourse. *Corpus Linguistics and Linguistic Theory*, 14(2). <https://doi.org/10.1515/cllt-2016-0040>
- Kamil, I. (2020, October 7). Istilah Penyandang Cacat dalam UU Cipta Kerja Menyakiti Perasaan. <https://Nasional.Kompas.Com/Read/2020/10/07/18402651/Istilah-Penyandang-Cacat-Dalam-Uu-Cipta-Kerja-Menyakiti-Perasaan>.
- Kovecses, Z. (2010). *Metaphor: A Practical Introduction, Second Edition*. New York.
- Kruawong, T., & Phoocharoensil, S. (2022). A Genre and Collocational Analysis of the Near-Synonyms Teach, Educate, and Instruct: A Corpus-Based Approach. *Teflin Journal*, 33(1). <https://doi.org/10.15639/teflinjournal.v33i1/75-97>
- Lakoff, G., & Turner, M. (1989). *More than Cool Reason: A Field Guide to Poetic Metaphor* (Later Printing; -, Ed.). The University of Chicago Press: The University Of Chicago Press.
- Lakoff, George., & Johnson, M. (2003). *Metaphors we live by*. University of Chicago Press.
- Lindquist, H. (2009). Corpus linguistics and the description of English. In *Corpus Linguistics and the Description of English*. <https://doi.org/10.2478/icame-2020-0006>
- Maftuhin, A. (2016). Mengikat Makna Diskriminasi: Penyandang Cacat, Difabel, dan Penyandang Disabilitas. *INKLUSI*, 3(2). <https://doi.org/10.14421/ijds.030201>
- Masrurin, A. (2021). ANALISIS TINDAK TUTUR BAHASA BLANTIK SAPI DI PASAR SAPI KABUPATEN TUBAN (PENDEKATAN PRAGMATIK). *EDU-KATA*, 7(2). <https://doi.org/10.52166/kata.v7i2.2718>
- McArthur, T. (1982). *Longman Lexicon of Contemporary English (ELT)*. London: Greener Books.
- McEnergy, T., & Hardie, A. (2011). Corpus linguistics: Method, theory and practice. In *Corpus Linguistics: Method, Theory and Practice*. <https://doi.org/10.1017/CBO9780511981395>
- Miles, M., & Huberman, A. (2004). *Qualitative Data Analysis: An Expanded Sourcebook*, Second Edition. In *Qualitative Data Analysis: An expanded sourcebook*. London: Sage Publications.
- Monaghan, P., & Roberts, S. G. (2021). Iconicity and Diachronic Language Change. *Cognitive Science*, 45(4). <https://doi.org/10.1111/cogs.12968>
- Pepper, P. (2016, November 22). We've had all the insults. Now we're reclaiming the language of disability | Penny Pepper | The Guardian. Retrieved January 7, 2023, from theguardian.com website: <https://www.theguardian.com/commentisfree/2016/nov/22/language-of-disability-stereotypes-disabled-people>
- Phoocharoensil, S. (2021). Semantic prosody and collocation: A corpus study of the near-synonyms persist and persevere. *Eurasian Journal of Applied Linguistics*, 7(1). <https://doi.org/10.32601/ejal.911269>
- Pusat Bahasa. (2007). Pedoman Umum Pembentukan Istilah. In *Jakarta: Konsil Kedokteran Indonesia, 1992* (3rd ed.). Jakarta: Departemen Pendidikan Nasional.
- Rayson, P., Archer, D., Piao, S., & McEnergy, T. (2004). The UCREL semantic analysis system. *Proceedings of the Beyond Named Entity Recognition Semantic Labelling for NLP Tasks Workshop*.
- Roziewski, S., & Kozłowski, M. (2021). LanguageCrawl: a generic tool for building language models upon common Crawl. *Language Resources and Evaluation*, 55(4). <https://doi.org/10.1007/s10579-021-09551-7>
- Sameer, I. H. (2016). A cognitive study of certain Animals in English and Arabic Proverbs: A Comparative Study. *International Journal of Language and Linguistics*, 3(5).
- Septiana, R. E., & Rahmawati, L. E. (2021). BENTUK EUFEMISME DALAM BERITA UTAMA SURAT KABAR TEMPO. *PARAFRASE: Jurnal Kajian Kebahasaan & Kesastraan*, 21(1). <https://doi.org/10.30996/parafrase.v21i1.4538>

- Sinambela, S. I., & Mulyadi, M. (2020). ANALISIS EUFEMISME DAN DISFEMISME PADA DETIK.COM DI TWITTER. *BAHAS*, 30(1). <https://doi.org/10.24114/bhs.v30i1.16669>
- Sinclair, J. (1991). *Corpus, concordance, collocation*. USA: Oxford University Press.
- Yamin, M., Setiawan, S., Anam, S., & Kurnia, F. D. (2020). Critical thinking in language planning for local languages maintenance: Evidence from Banjarese, Indonesia. *Language Related Research*, 11(5). <https://doi.org/10.21859/LRR.11.5.179>
- Yolanda, Y. (2022). *Metafora konseptual: memahami bahasa dari sudut pandang semantik kognitif* (S. T. Purnani, Ed.). Jember: UPT Penerbitan Universitas Jember.
- Yuliawati, S. (2016). Profil Semantis Nomina Perempuan dalam Korpus Majalah Berbahasa Sunda (Manglè, 1958-2013). *Kongres Internasional Masyarakat Linguistik Indonesia*.
- Yuliawati, S. (2018). Perempuan atau Wanita? Perbandingan Berbasis Korpus tentang Leksikon Berbias Gender. *Paradigma, Jurnal Kajian Budaya*, 8(1), 53. <https://doi.org/10.17510/paradigma.v8i1.227>

Yoga Yolanda is an Assistant Professor of Study Program of Indonesian Language and Literature Education, in Universitas Jember. He mainly works in the field of Semantics. Some of his contributions appeared in journals such as *Indonesian Language Education and Literature Journal*. He also published several reference books, such as *Metafora Konseptual: Memahami Bahasa dari Sudut Pandang Semantik Kognitif*.

Budi Setyono is currently affiliated with the Department of English, the Faculty of Teacher Training and Education, Universitas Jember, Indonesia. In the year of 2019, he was granted a professor in English Education by the Ministry of Education and Culture of the Republic of Indonesia. His areas of interest include critical analysis of English textbooks, curriculum development, and language teacher education. Four of his articles have been published in the "Journal of Intercultural Education", "Sexuality and Culture", "The Journal of ASIA TEFL", and "Indonesian Journal of Applied Linguistics (IJAL)".
