

THE EVALUATION OF DISABILITY-RELATED TERMS IN KBBI: A CORPUS-BASED LEXICOGRAPHY STUDY

Dewi Khairiah and Selly Rizki Yanita

The Agency of National Research and Innovation (BRIN), Tangerang
dewi040@brin.go.id; selly.rizki.yanita@brin.go.id

Abstract: As a descriptive-prescriptive monolingual dictionary, *Kamus Besar Bahasa Indonesia* (KBBI) or the Great Dictionary of Indonesian Language, records all kinds of linguistic phenomena, including slang, jargon, etc., found in Indonesian language. At the same time, it functions as the source of Indonesian language learning by providing standard and non-standard vocabularies. Therefore, it should adhere to Indonesian language corpora in the entry selection as well as in defining each concept. Unfortunately, the disability-related entries and their definitions in KBBI do not appropriately represent disability concept and become the incessant issues, mainly among people with disability. In this article, we try to solve these issues by doing evaluation of three disability-related entries in KBBI. For the purpose of this study, we use the Indonesian corpus selected in the period of 2010—2020, taken from the Leipzig Corpora Collection. Then, we analyze the corpus data to find the word frequency, collocations, and concordances of the words *cacat* ('disabled; disability'), *disabilitas* ('disability'), and *difabel* ('disability'). The semantic fields of the collocations are determined based on the USAS Semantic System to find out the senses of the words. Next, based on the analysis of collocations and concordances, we decide which terms need evaluation. This evaluation shows that the definitions of the words *cacat*, *penyandang cacat*, and *difabel* are not adhered to the corpus evidence. Finally, we give appropriate recommendations for some improvements.

Keywords: *disability, monolingual dictionary, lexicography, corpus linguistics*

INTRODUCTION

Law on Disabilities No. 8 of 2016 says that people with disability are those having long term physical, mental, intellectual, or sensory limitations which restricts their social interaction and participation. In Indonesia, there are various terms related to disability, accounted as *bercacat* ('disabled'), *tuna* (deficit/loss), *penderita* or *penyandang cacat* ('people with disability'), *anak berkebutuhan khusus* ('child with special need'), *difabel*, *disabilitas* and *penyandang disabilitas* ('people with disability'). The term *difabel* emerged in the 1990s as the acronym for 'differently able' to counter the stigma that disability is always associated to 'having no ability', otherwise, people with disability have ability different from common people. For example, a crippled person is still able to move by using the wheelchair (Syam, 2020). Nowadays, there is a tendency, if we can say pressure, to eliminate the use of the term *penyandang cacat* from the regulation texts since it is regarded offensive and discriminative (Kamil, 2020).

The disability terms are recorded in KBBI, reflecting the linguistic perceptions of the Indonesian speakers from time to time, as KBBI is a diachronic as well as descriptive dictionary. The term *penyandang cacat* is recorded in all editions of KBBI, while the term *difabel* are only found in the fourth and fifth edition of KBBI. The term *disabilitas* comes later in the fifth edition of KBBI. From our perspective, dictionary is significant in the study of how a society perceive

concepts of things, mainly the conceptualization of thought and ideas about disability. The inaccurate defined concepts in dictionary will lead to the misconception and bad stigma, whereas dictionary should be a reliable reference. Therefore, lexicographers should treat the entries in KBBI appropriately to fulfil the needs of their users for relevant and precise concepts.

Unfortunately, as far as we know, there is no academic study conducted yet on the discussion of disability-related terminologies and how they are treated as entries in dictionary. The topic of disability-related terminologies in Indonesia still evolves in the study of the historical background of the terms and their usage from time to time, as conducted by Suharto et al. (2016) and Widinarsih (2019). Suharto (2016) classifies the changes of disability terminologies into two periods, the Old Order Era (1945—1966) and the New Order Era (1966—1998 and afterwards). He finds that the terms used in the New Order Era are subtler than the ones in the Old Order Era, while after reformation era of 1998 the more inclusive terms are born. Furthermore, he also highlights the use of *difabel* in the last decade as part of the struggle of disability community in Indonesia to conceptualize themselves in the right way, since this term is considered more positive than others. Similar to Suharto, Widinarsih (2019) records the changes of disability terminologies in Indonesia. She also highlights that disability awareness in Indonesia is mostly negative, due to the normality concept perceived by society.

To fill in the gaps in the previous studies, we conduct a mixed-method study of corpus linguistics and lexicography to analyse and evaluate the definitions of disability-related terms in KBBI. Our research only focuses on three terms, they are *cacat*, *disabilitas*, and *difabel* because the three terminologies are mainly found in the government regulations and legal documents. We analyse the collocations and concordances to grasp the appropriate sense or meanings of each word. Then, we evaluate the definitions provided by KBBI for the words *cacat*, *penyandang cacat*, and *difabel* based on our analysis of collocations and concordances. The result of this evaluation will be the basis for our suggestions to revise the entries *cacat* and *penyandang cacat* in KBBI.

Literature Review

Dictionary Evaluation

Evaluation is commonly known as review or critique, so evaluation for dictionary means a review or critique to dictionary. However, most reviewers often only describe the design features of dictionary, and the choice of lemmas becomes a favourite topic. They do not explicate their evaluation and leave the readers with no solutions. Thus, they do not perform the functions of evaluation: 1) to provide a sharp analysis of the quality of dictionary that can assist the dictionary users in choosing the best dictionary based on their usage needs, and 2) to help lexicographers in strengthening the function of their dictionary. In lexicography, the evaluation on dictionary is commonly known as dictionary criticism. Gouws, et al. (2018: 36) argues that dictionary criticism not only gives a review of a lexicographic product but also suggests improvements to enhance the quality of dictionary. In this article, the evaluation of disability terms in KBBI can be categorized into dictionary criticism as it is aimed to make recommendation for KBBI's definitions based on semantic and corpus analysis.

Dictionary criticism concentrates on the properties of dictionary based on the criteria that are appropriate for the type of dictionary and its target users. There have been various theories built on the parameters for dictionary evaluation. In general, the dictionary evaluation criteria can be divided into dictionary type (all types or specific types), dictionary features, and information covered (scope of evaluation) (Swanepoel, 2008: 214). Furthermore, some evaluation criteria are applicable to all types of dictionaries, specifically the Renkemas's (1996) 3C model which involves correspondence, consistency, and correctness.

In terms of correspondence criterion, the criticism can possibly focus on the deviations in the ordering of lexical items in the wordlist of a dictionary, the microstructural treatment, or

the definition strategy. Meanwhile, consistency is closely related to the design of a dictionary with which users get acquainted. The lexicographer's consistency to apply the guidelines of dictionary design leads to the dictionary's reliability. Correctness corresponds to "the truth", meaning that dictionary provides users with the factual evidence of language use (Swanepoel, 2008: 210-211).

Definition in Monolingual Dictionary

In lexicography, meaning is commonly known as "sense" or "word sense" in which one lexical unit may have one or more senses (polysemes). The ability of lexicographer to express the meaning of lexical item by means of paraphrase is fundamental in lexicographic practice, whether it is conducted in the same language (monolingual dictionary) or in two different languages (bilingual dictionary). It is often a challenging task for most lexicographers to organize sense or senses into definition because they should adapt the design and style of definition to some key factors, such as the size, purpose, media (print, digital, or online), and target users of the dictionary. For bilingual dictionary, this challenge is quite easier to face (but not at all) because its users only need equivalence for a lexical unit. However, monolingual dictionary sometimes finds it tricky to describe meaning to its users who come from different background of knowledge and ability to understand vocabularies (Haas, 2014).

The construction of sense or meaning of a lexeme in definition should distinguishes the meaning of the lexeme to the meaning of other lexemes in similar semantic field. Unfortunately, there is no standard theory or method of lexicography regarding how to construct a definition. Nevertheless, Haas (2014) says that definition in monolingual dictionary generally comprises one or more of these features, they are descriptive phrases, full-sentence definitions, synonyms, combination of paraphrase and synonym, pragmatic information or situational context, and encyclopaedic information. Moreover, the description of meaning in definition is based not only on the lexical semantic categories but also the different usage of the word, depending on the needs of the users.

According to Svensén (2009), definition can be classified into two, they are intentional and extensional.

a. Intentional definition

This type of definition relates to the system of conceptual hierarchy, describing the generic conceptual relation of superordinate, subordinate, and coordinate, in which the concepts are categorized based on the similarities and the differences. Although this definition is commonly found in technical dictionaries, it is also used to describe terms in general dictionaries.

b. Extensional definition

This definition describes the meaning in partitive conceptual relation in which partial concepts are constructed to define a lexical unit. For example, 'chair, table, bed, cupboard, etc.' are the subordinates while 'furniture' is the superordinate.

METHODS

In finding sense variations of *cacat*, *difabel*, and *disabilitas* in KBBI, we apply the approach of corpus linguistic. Corpus linguistics is an approach that enables researchers to investigate language variation and use empirically which results general and valid research findings (Biber and Reppen, 2015). To analyse semantic change or variation, corpus linguistics is significant (1) to trace how one word is differently used across the time, (2) to give authentic evidence of changes, and (3) to identify the meaning development of a word (Puspita and Kamal, 2020). The corpus we use in this research is taken from Corpora Collection of Leipzig University (Leipzig Corpora) from 2010 to 2020.

The features in corpus linguistics are useful in investigating the change of meaning, specifically the concordance and collocation (Sinclair, 1991). Concordance enables researchers to look at how one word takes position in various contexts. Meanwhile, collocation provides the information about the surrounding words at the left and the right of the investigated word. These features are applied in diachronic order to see the behavioural change of the word. For the sake of this research, the corpus used is arranged diachronically as the behaviour change of word might be seen obviously in diachronic corpus (Kučera and Stluka, 2010).

FINDINGS AND DISCUSSION

The data in this research is taken from various online articles in Indonesian Corpora Collection of Leipzig University (Leipzig Corpora) with 110,858,505 tokens, ranging from the year 2010 to 2020. The Antconc corpus software is utilized to investigate the collocations and concordances of *cacat*, *difabel*, and *disabilitas*.

Word Frequency Analysis

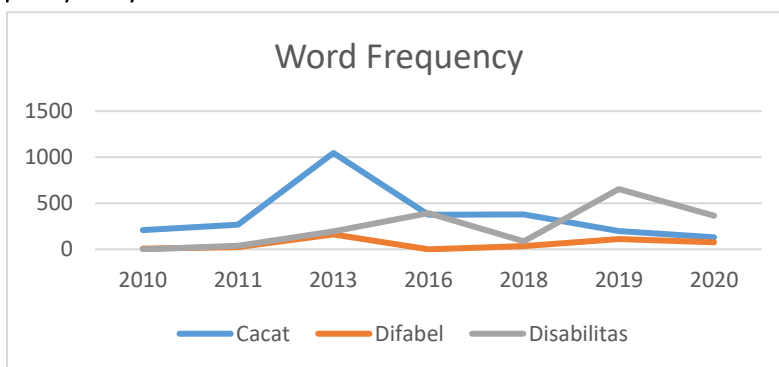


Figure 1. Word Frequency of cacat, difabel, and disabilitas

According to the result of Antconc word frequency, the word *disabilitas* is not found in the data of 2010, but it is found in the data of 2011. This is possibly related to the use of this term in the Law No. 19 of 2011 concerning the rights of persons with disabilities. Meanwhile, in 2010, the frequency of *cacat* is higher than *difabel*. The word *cacat* is frequently used in the variety of semantic fields, both formal and informal. Not only common people and government, the groups of people with disabilities use this word in naming their organizations and activities, like *Yayasan Pembinaan Anak Cacat (YPAC)*. Although the word *difabel* is found in 2010, its frequency is very small.

In 2013, there is a significant uplift in the presence of the words *cacat*, *difabel*, and *disabilitas*, in which *cacat* is the most escalated word. The semantic field of *cacat* at that time was more expanded, it was not exclusively used to refer to person and thing, but also other abstract materials, such as law, administration, mind, etc. The overarching discussion on the contexts of *cacat* is discussed in the collocation and concordance section. Concurrently, the frequency of the word *disabilitas* started to escalate since 2013. This phenomenon is in reference to the issues of local government regulations concerning on promoting the rights of people with disabilities, including the signing of MoU between the government of Yogyakarta City and UNESCO in the building of inclusive cities.

Contrary to the data of 2013, the data of 2016 shows the significant fall of the frequency word of *cacat* below the frequency word of *disabilitas*. This fact is probably provoked by the emergence of the Law No. 8 of 2016 on People with Disability which dominated the news topic in Indonesia at that time. In contrast, the frequency of the word *difabel* declines.

The data at the end of the period investigated reveals the fact that there is a trend for the word *disabilitas* to increase during 2018—2019, yet it declines in 2020. This situation is different with the word *cacat* in which it tends to drop consistently during 2018—2020. The massive campaigns of equality for people with disability contribute to this phenomenon. At the same time, the frequency of the word *difabel* slightly rises from 2018 until 2019, before it is stagnant by 2020. This data obviously shows that the position of the word *disabilitas* among the Indonesian speakers is firmer than the word *difabel* and *cacat*.

Collocation and Concordance Analysis

After the investigation of the word frequency of *cacat*, *difabel*, and *disabilitas*, the next step is to look for the collocations of the three words. Since the Indonesian tagger application is not provided yet, the tagging is processed through two steps. First, the word is translated into English. Second, the English translation is tagged by using USAS English tagger. Following is the findings and discussion about collocation and concordance of each word based on USAS semantic fields.

Collocations related to person and group of persons.

Penyandang, *penderita*, *orang*, and *teman* are the collocations related to disabled individuals and *kaum*, *komunitas*, and *kelompok* are the collocations for group of disabled individuals. The collocation with the highest significancy value for the word *cacat* and *disabilitas* is *penyandang*, it means that Indonesian speakers commonly call the disabled people as *penyandang cacat* or *penyandang disabilitas*. Moreover, we find the essential fact that the word *penderita* is mostly found in the data from the year of 2010—2015, yet it appears in small number in the period of 2016—2020. It also only comes up in the concordance of the word *cacat* and *disabilitas* but not with *difabel*. From the data, we can assume that *penyandang cacat*, *penyandang disabilitas*, and *penderita cacat* are commonly used instead of *penderita disabilitas* and *penyandang difabel*.

It should be noted that *penyandang* and *penderita* are associated with health and disease, while *teman* and *orang* are included in the semantic field of people, of which the former indicates relationship. In general, it reflects the fact that people with disability is lately considered as part of society, not “the others”. The word *disabilitas* and *difabel* collocate with *teman* as found in our corpus data from 2016 to 2020.

Collocation related to ethic and respect.

Ethic is about norms and standard of moral principles in society, while respect relates to admiration or esteem. The collocations related to ethic and respect are *memalukan*, *cela*, and *aib* (‘shameful; disgrace’). These can refer to the character of person when *cacat* is related to the disgraceful action, manner, or attitude of someone, as in the context of *cacat moral* (‘immoral’); or to the situation in which ethic is abandoned, as in the context of *perkawinan yang cacat* (‘a flawed marriage’).

Collocation related to evaluation.

The collocation related to evaluation is *sempurna* (‘perfect’) which only occurs in the corpus of *cacat*. This evaluation of good or bad is implied in the context of *cacat*. It is interesting that this collocation depicts the quality of something based on the structure (body) condition and the functions of parts of structure. The term *cacat* is the evaluation result of the perfectness degree of characteristics, both in human and things, in which it refers to the imperfection, for example *cacat moral* (‘immoral’), *cacat hukum* (‘legal defect’), *cacat fisik* (‘physical impairment’). However, when discussing about *cacat* as the evaluation concept of human body, we are not talking about the bad or good “quality” of a human body, but about the incompleteness (*kekurangan*) or abnormality (*kelainan*) in human body which limits the functions of mental, physic, or both (Centre for Disease Control and Prevention, n.d.).

Collocation related to damaging and destroying.

From the corpus, we collect two collocations categorized into the semantic field of damaging and destroying, they are *rusak* ('broken') and *retak* ('cracked'). The collocation of *rusak* implies the sense of physical and functional damage, as in the term *produk cacat* ('defect product'). This is also related to quality, alike the collocation of *sempurna*. This collocation is mostly found in the context of products, only once appear regarding damage happened to part of human body.

Collocations related to metaphorical expression.

Considering some collocations to which the word *cacat* collocates possibly construct non-literal sense, we propose the semantic field in which the collocations are related to the forming of metaphorical expression. The collocations in the form of abstract things are *hukum* ('legal'), *moral*, *prosedur* ('procedure'), and *prosedural* ('procedural'). The concept of defect or *cacat* in relation to abstract things describes the imperfection and insufficiency of product, process, or legal document. Defect in a product is closely connected to its quality, but defect in a process (like administrative process) or in a legal document is related to the procedure and requirement. For example, the term *cacat hukum* ('legal defect') is in relation to the failure of a case or legal document to meet administrative requirements or required procedure.

Definitional Evaluation of Disability-Related Terms in KBBI

After analysing the collocations and the concordances of the words *cacat*, *disabilitas*, and *difabel*, we conclude that the word *cacat* has a wider range of potential meanings (or polysemes) than *disabilitas* and *difabel*, as it refers not only to human but also to products or things. On the contrary, the words *disabilitas* and *difabel* exclusively refer to person (human) and his or her condition. Based on this analysis, we evaluate the definitions of the entries *cacat*, *penyandang cacat*, and *difabel* in KBBI.

Definitional Evaluation of *cacat*

The entry *cacat* in KBBI comprises four polysemes as follows.

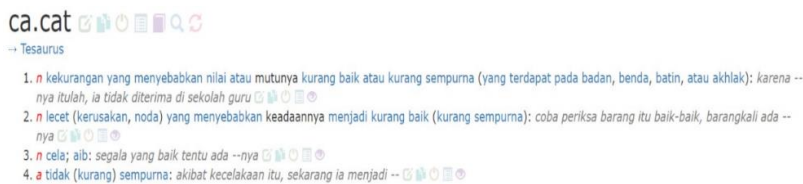


Figure 2. The Definitions of *cacat* in KBBI

As seen from the figure above, polyseme 1 comprises the concepts of evaluation of body, thing, mental, and moral, indicated by the words *nilai* (value) and *mutu* (quality). Value, according to Cambridge Dictionary (n.d.) is 'the importance or worth of something for someone; how useful or important something is' and quality is defined as 'how good or bad something is'.

If we refer to the result of collocation analysis related to evaluation, the evaluation concept of value and quality implied in the word *cacat*, both referred to human and products or things, is definitely different. *Cacat* in a product or thing is associated with the degradation of its value and quality as a whole entity, but *cacat* in human body does not degrade him or her as a human being. Furthermore, the imperfection in product is related to the quality evaluation while the imperfection in a baby is related to appearance. Similarly, it is not common to say a disabled baby is not qualified, yet we can say a defect product is unqualified.

We conclude that the senses of value and quality as the concept of evaluation implied in polyseme 1 of the entry *cacat* is not appropriate to human body. As we look at the definition in the second polyseme, it seems that it implies similar sense with the one in the first polyseme,

indicated by the word *kerusakan* ('damage; impairment'), *lecet* ('scratch'), *noda* ('stain') which all refer to the same semantic field, that is damage. Therefore, the senses in polyseme 1 and polyseme 2 are possibly merged into one sense.

Next, we investigate the polyseme 3 of the entry *cacat*. It defines *cacat* as *cela; aib* ('dishonour; disgrace'). Based on our collocation and concordance analysis, this sense is implied when *cacat* collocates with abstract things, such as in the phrase *cacat hukum, cacat moral*, etc, or when *cacat* is used in the context of abstract things, such as in the expression of *prestasi pegawai itu sangat sempurna, tidak ada cacatnya* ('the achievement of that employee is excellent, with no disgrace'). This sense of the entry *cacat* is categorized into metaphor and needs to be labelled as *kiasan (ki)* or 'metaphoric'.

Furthermore, the polyseme 4 of the entry *cacat* contains the contextual use of this word. However, we consider that the sentence example given should be changed since the "perfectness" concept referring to *cacat* in human is totally different to the one in things or products. A disabled person is still able to perform limited physical activities with his or her "imperfect condition". In contrast, a defect or broken product is considered imperfect, and its price is lower than the perfect ones. The corpus data we found supports this opinion about "perfectness". It shows the tendency of the use of *sempurna* regarding with product context, as this term is considered inappropriate to be labelled to disabled people. Therefore, the context use for polyseme 4 should be for product.

Definitional Evaluations of *penyandang cacat* and *difabel*

KBBI defines *penyandang cacat* as *penderita cacat*, while *difabel* is defined as *penyandang cacat*. This kind of circular definition is common in lexicography practice, in which lexicographers decides to define a term by using another term that is assumed known by the users. Thus, the lexeme (s) being used in definition is or are synonymous lexeme (s) as shown in the figure below.



Figure 3. The definitions of *penyandang cacat* and *difabel* in KBBI

Based on the collocations and concordances analysis, *penderita* ('sufferer') is associated with health and disease issue. In fact, not all disability results from disease. Some are due to accident, in which parts of body are missing or impaired. Otherwise, this cause-effect relationship conceptualizes disability as seen in the definition of *disabilitas* below.

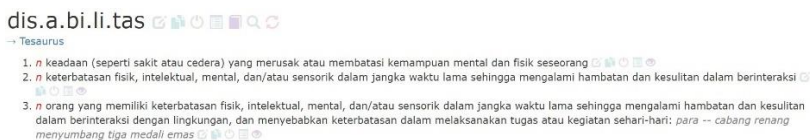


Figure 4. The definitions of *disabilitas* in KBBI

The concepts implied in the definition of *disabilitas* cover both the senses of *penyandang cacat* and *difabel*. Consequently, *disabilitas* can be used properly to define these terms.

CONCLUSIONS

Based on our evaluation to the definition of the entry *cacat*, we propose the improvement as the following:

- ca.cat** 1. *n* kekurangan (leceh, noda, kerusakan) pada benda yang menyebabkan nilai atau mutunya kurang baik atau kurang sempurna
2. *n* kelainan atau kekurangan pada anggota tubuh yang membatasi fungsi tubuh
3. *n* *ki* cela; aib: *prestasinya sempurna tanpa --*
4. *a* tidak atau kurang sempurna: *produk ini – sehingga tidak dapat dipasarkan*

Besides, we propose an improvement to define the entry *penyandang cacat* with the synonym definition, that is *disabilitas*. Therefore, the entry *penyandang cacat* and *difabel* are defined as below.

penyandang cacat disabilitas

di.fa.bel *n* disabilitas

As a reference used to understand the concepts in Indonesian language, KBI should appropriately record the senses of words adhered to the available corpus. Besides, the lexicographers should select carefully which lexical units are suitable to represent the concepts that construct a definition. Otherwise, the users will be provided misleading concepts.

At last, the evaluation we do and the recommendations we make for the improvements of the disability terms in KBI are based on the corpus evidence of the actual linguistic usage found in Indonesian speakers. As a result, this reflects how Indonesian society perceive and construct the concepts of disability.

REFERENCES

- Biber, D., & Reppen, R. (2015). Introduction. In Biber, D. and Reppen, R. (ed). *The Cambridge handbook of English corpus linguistics*. Cambridge University Press.
- Cambridge Dictionary. (n.d.). quality. In *dictionary.cambridge.or*. Retrieved May 28, 2023, from <https://dictionary.cambridge.org/dictionary/english/quality>
- _____. (n.d.). value. In *dictionary.cambridge.or*. Retrieved May 28, 2023, from <https://dictionary.cambridge.org/dictionary/english/value>
- Centers for Disease Control and Prevention. (n.d.). *Disability and health overview*. <https://www.cdc.gov/ncbddd/disabilityandhealth/disability.html>
- Gouws, R.H., et al. (2018). Dictionary critics and dictionary criticism. In Bielinska, M. & Schierholz, S.J. (eds.), *Dictionary Criticism*. Berlin: De Gruyter, 33–49.
- Haas, C. (2014). *The different ways of describing meaning in monolingual dictionaries*. Munich: GRIN Verlag.
- Kamil, I. (2020, October 7). *Istilah penyandang cacat dalam UU Cipta Kerja menyakiti perasaan*. Kompas. <https://nasional.kompas.com/read/2020/10/07/18402651/istilah-penyandang-cacat-dalam-uu-cipta-kerja-menyakiti-perasaan>
- Kučera, K. & Stluka, M. (2010). Diachronic corpora: Seeing histories of words from another angle. In Considine, J. (Ed.). *Web of words: New studies in historical lexicography* (h.1-7). Newcastle upon Tyne, England: Cambridge Scholar Publishing, 1–7.
- Puspita, D. & Kamal, Y. (2020). Sketching the semantic change of *jahanam* and *hijrah*: A corpus-based approach to manuscripts of Arabic-Indonesian lexicon. *Journal of Arabic Studies*. 5. 1. 10.24865/ajas.v5i1.246.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Suharto, Pim, K., & Dorsett, P. (2016). Disability terminology and the emergence of ‘diffability’ in Indonesia. *Disability & Society*, 31(5), 693–712.
- Svensén, B. (2009). *A handbook of lexicography: The theory and practice of dictionary*. Cambridge: Cambridge University Press.

- Swanepoel, P. (2008). Towards a framework for the description and evaluation of dictionary evaluation criteria. *Lexicos*, 18, 207–231.
- Syam, A.Z. (2020, July 21st). *Catatan perkembangan istilah difabel dalam kebijakan Indonesia*. Solider. <https://www.solider.id/baca/6135-catatan-perkembangan-istilah-difabel-dalam-kebijakan-indonesia>
- Widinarsih, D. (2019). Penyandang disabilitas di Indonesia: Perkembangan istilah dan definisi. *Jurnal Ilmu Kesejahteraan Sosial*, 20(2), 127–142.