

Lexical Understanding of Native Bahasa Indonesia Speakers through Word Association to Improve Dictionary Definitions*

Ami P. Jewalani, *Department of Linguistics, Faculty of Humanities,
Universitas Indonesia, West Java, Depok, Indonesia (ami.pramesti@ui.ac.id)*,

Myrna Laksman-Huntley, *Department of Linguistics, Faculty of
Humanities, Universitas Indonesia, West Java, Depok, Indonesia
(laksman@ui.ac.id)*

and

Harwintha Y. Anjarningsih, *Department of Linguistics, Faculty of
Humanities, Universitas Indonesia, West Java, Depok, Indonesia
(harwintha@ui.ac.id)*

Abstract: This study examined lexical cognitive word associations of Bahasa Indonesia native speakers. A word association task using 30 cue nouns was conducted with 45 educated adult native speakers of Bahasa Indonesia aged 20–29, after which the generated data was classified based on the extensive semantic taxonomy. It was found that most responses related to the cue words were associated with lexical features, followed by entity features, situation features, taxonomic category, and introspective features, all of which suggested that this group of Bahasa Indonesia speakers related words to other words that shared similar lexical features, and especially with words that usually come after the target words. It was also found that the participants rarely associated feelings to the cue words as there were very few introspective feature associations. While this was a limited study focused on a specific population and only used nouns, the results could be of assistance in developing dictionaries and thesauri, or could be used as preliminary data to build databases, such as WordNet in Bahasa Indonesia. As there have been few studies focused on Bahasa Indonesia word associations, this study could also be used for future comparative word association studies.

Keywords: WORD ASSOCIATION, SEMANTIC PROPERTIES, CONCEPTUAL PROCESSING, SEMANTIC TAXONOMIC CODING, BAHASA INDONESIA, COGNITION, DICTIONARY DEFINITIONS, WORD RETRIEVAL, LEXICAL ACCESS

* A shortened version of this article was presented as a paper at the Third Asia-Pacific Research in Social Sciences and Humanities Conference 2018 (APRISH 2018), which was hosted by the Faculty of Social and Political Sciences, Universitas Indonesia, in Jakarta, Indonesia, 13–15 August 2018.

Résumé: Compréhension lexicale des locuteurs natifs de Bahasa Indonesia à travers d'une association de mots pour améliorer les définitions de dictionnaires.

Cette étude a examiné les associations de mots lexicaux cognitifs en Bahasa Indonesia de ses locuteurs natifs. Une tâche d'association de mots utilisant 30 noms de mots-clés a été réalisée en impliquant 45 locuteurs indonésiens bien instruits âgés de 20 à 29 ans à la suite de quoi les données générées ont été classées selon la vaste taxonomie sémantique. Cette étude a révélé que la plupart des réponses liées aux mots-clés étaient associées à des caractéristiques lexicales, suivies des entités, des caractéristiques de situation, de la catégorie taxonomique et des caractéristiques introspectives, ce qui suggère que ce groupe de locuteurs indonésiens associe des mots avec d'autres mots partageant des caractéristiques lexicales similaires, en particulier avec les mots qui viennent habituellement après les mots cibles. Par ailleurs, il est à constater que les participants associaient rarement les sentiments aux mots-clés car il y avait très peu d'associations présentant des caractéristiques introspectives. Bien qu'il s'agisse d'une étude limitée centrée sur une population spécifique et utilisant uniquement des noms, les résultats pourraient être indispensables pour le développement de dictionnaires et de thésaurus, ou pourraient s'utiliser comme données préliminaires pour créer des bases de données, telles que Wordnet en Bahasa Indonesia. Puisque peu d'études ont été consacrées aux associations de mots en Bahasa Indonesia, cette étude pourrait également servir à de futures études comparatives sur les associations de mots.

Mots-clés: ASSOCIATION DE MOTS, PROPRIÉTÉS SÉMANTIQUES, TRAITEMENT CONCEPTUEL, CODAGE TAXONOMIQUE SÉMANTIQUE, BAHASA INDONESIA, COGNITION, DÉFINITIONS DE DICTIONNAIRE, RECHERCHE DE MOTS, ACCÈS LEXICAL

1. Word relations for concept understanding

In all languages, words are used to label concepts that are understood by others who speak that language. Generally, concepts are mentally connected to other concepts, such as the concept of food to the concept of eating, or food to hunger, or food to dining room. Language and situated simulation (LASS) theory states that the "conceptual system is organized around situated action" (Barsalou 2003: 522); in other words, accessing one concept brings to mind other associated concepts and accessing a word automatically brings to mind other associated words.

Language and cognition studies have found that understanding these concepts and word associations can give some indication as to how these are stored in the memory, and sociolinguistic studies have examined these word and concept associations to understand how lexical items are understood in different cultures and languages, including Bahasa Indonesia. Therefore, this study examines how Indonesians understand their language through a word association task, which "is the simplest possible form of the linguistic processing that occurs during conceptual processing" (Barsalou, Santos, Simmons and Wilson 2008: 249).

A word association test was used by Bahar, Johnstone and Sutcliffe (1999) to investigate student cognitive structures in a genetics context, for which Eng-

lish genetics cue words, such as gene, chromosome, and mutation were used to generate other cue words to understand how the students linked the words to others. De Deyne and Storms (2008) also used a Dutch word association test to explore semantic networks and their properties, from which they found that the associates were invariably nouns regardless of the word category cues, and most of the elicited words were related to a certain situation or in taxonomic ways to the cue words. The second responses also revealed that entity features became more available indicating that while the first words were taxonomically related, the following associations were conceptual (De Deyne and Storms 2008: 223).

Pranoto and Afrilita (2018) explained the manner in which the network of mental lexicon models in someone's cognition structures can be seen through a lexical relation in psycholinguistics. Therefore, word association tasks can reveal how certain concepts are stored in a certain language community as they require participants to produce words they feel are closely connected to a particular word. Word associations are subject to the culture of the speaker of a certain language since, as Pranoto and Afrilita (2018: 30) stated, "the mental lexicon develops according to the events experienced by the language learner." For example, the word *candle* for Bahasa Indonesia would not be connected to *church* as most Indonesians are Moslem, so this word would be more likely associated with *power failure* — a situation when Indonesians use candles to light the room.

Geeraerts (1987) classified dictionaries into three types based on the theories used in compiling them and their functional intentions. One of them is the large-scale dictionary that is designed to scientifically describe linguistic variations and spread knowledge of the words which uses prototypical concepts as the definitions. The *Kamus Besar Bahasa Indonesia* (KBBI) is one of the examples. In the same way, the associated words from native Bahasa Indonesia speakers could be seen as prototype definitions and could be used by lexicographers to assess a word's meaning in the speakers' minds based on prototype theory. Van Sterkenburg (2003) claimed that while the definitions in dictionaries are in both whole sentences and phrases, the sentence format was more natural and spontaneous in defining words and concepts. There are several headwords with short definitions in the KBBI. For example, the lemma *jawab* 'answer', in the KBBI V offline, is defined as *sahut, balas* 'response, reply'; therefore, more information is needed for this to be fully understandable. Adding information about the context in which *jawab* is defined as *balas* would assist learners to understand better when and how to use the word. Therefore, it is intended that the current study could assist lexicographers develop better definitions by revealing the associated concepts in the minds of Bahasa Indonesia speakers.

This study is also expected to assist in developing a thesaurus of Bahasa Indonesia. In one Bahasa Indonesia thesaurus, Endarmoko (2007), the headword synonyms are written in numerical order. For example, the synonyms for the lemma *bulat* 'round' are written as follows:

bulat *a* 1 bundar; 2 a cukup, genap integral, jangkak (*Jw*), kafi (*Ar*), komplet, lengkap menyeluruh, padu, penuh, sempurna, solid, tamam (*kl*), utuh; tunggal

However, as the difference between 1 and 2 is not obvious, using examples in sentences or from the domain could be added. Therefore, the associated words and conceptual relationships elicited in this study could be a reference for determining an example or the domain in which the words are used.

As words can be related to other words in many ways, to understand the relationships between the cue and the participant associated words, the semantic taxonomic coding adapted by De Deyne and Storms (2008) was used.

2. Semantic taxonomic coding

Depending on culture, language, background and experience, people connect words in many ways. *Sky*, for example, can be related to *cloud*, *stars*, and *moon* or *blue* or even to the word *romantic*. De Deyne and Storms (2008) developed a semantic taxonomy that had a wide range of categories that can be used to determine how one word is linked to another word by meaning and external factors, such as the location where the word is usually found. There were five main categories in De Deyne and Storms' semantic taxonomy: (1) entity features, (2) situation features, (3) taxonomic categories, (4) lexical features, and (5) introspective features as well as several subtypes, all of which are helpful in understanding the cognitive processes associated with the way words are used. Below is the elaboration of De Deyne and Storms' semantic taxonomy categories and their subclasses that is exemplified using Bahasa Indonesia.

2.1 Entity features

De Deyne and Storms (2008) explained that a word association is put into the entity features category if a word is linked to its associate based on their specific concrete features regardless of the situation, for example, *kapur* 'chalk' and *putih* 'white'. They further classified this category into ten subclasses that explain the specific features shared with the intended word: (1) external component, (2) external surface feature, (3) internal component, (4) internal surface feature, (5) behavior, (6) material, (7) quantity, (8) associated abstract entity, (9) systemic feature, and (10) larger whole.

The external feature is related to the outer part of the concept labeled by the main word while the internal feature is related to the inner part (De Deyne and Storms 2008). In addition, the surface feature is regarding the entity that exists in the concept of the word itself or is not its component. For example, with regard to the cue word *cabai* 'chili', *merah* 'red' is categorized as an external surface feature whereas *pedas* 'spicy' is classified as an internal surface feature.

Additionally, the component is associated with the words denoting composition. For instance, *roda* 'tire' is the external component of *mobil* 'car' whereas *bensin* 'fuel' is the internal component.

Moreover, according to De Deyne and Storms (2008), the behavior feature is related to words denoting natural characteristics, such as *bumi* 'earth' that is associated with *berputar* 'rotate' while the material feature is used to classify the relation of two words when one of the words is the material of the object denoted by the other word such as *karung* 'sack' is made of *plastik* 'plastic'. Moreover, the quantity feature designates the amount or number of a particular concept that usually exists, such as *tangan* 'hands' are *dua* 'two'. Words can also have abstract associations with other words, such as *rumah* 'house' and *keluarga* 'family'; systemic characteristics association, such as *asap* 'smoke' and *uap* 'steam', or larger whole type associations, such as a *rumah* 'house' being the larger part of a *lantai* 'floor'.

2.2 Situation features

This type of class, which is divided into thirteen subtypes, defines words of which the properties are linked by situations (De Deyne and Storms 2008). (1) The function subtypes indicate words that are related by a function, for example *berjalan* 'to walk' is a function of the word *kaki* 'leg'. Words can also be linked by (2) action properties, such as a *bola* 'ball' is *ditendang* 'kicked' or (3) object properties, such as *kaki* 'feet' and *sepatu* 'shoes'. (4) The person property can also be associated with a word in a particular situation, such as *anak-anak* 'children' that is associated with the word *gambar* 'picture'. (5) Living thing properties link living things to other words, such as *kucing* 'cat' and *pasir* 'sand', (6) social organization properties relate words, such as *murid* 'students' to *sekolah* 'school', and (7) social artifact properties link words, for instance *gambar* 'picture' with the artifact *lukisan* 'painting'.

In addition to the previous features, De Deyne and Storms (2008) stated that (8) location features, (9) time features, and (10) events features connect words through the situational properties; *cabai* 'chili' that can be found in the *pasar* 'market', a *karung* 'sack' that is used to celebrate *Hari Kemerdekaan* 'Independence Day', and a *stadion* 'stadium' connected to *piala dunia* 'world cup', are all examples of the location, time, and event subtypes, respectively. Words are also connected by (11) manner properties — how an event or action is performed or to what an entity is transformed, (12) physical states of a situation, and (13) quantities; for example, the association between *jawab* 'answer' and *jelas* 'clear' (manner), *lilin* 'candle' and *panas* 'hot' (physical state), and *balap* 'race' and *banyak mobil* 'many cars' (quantity).

2.3 Taxonomic categories

Another category used to display the manner in which words are connected to

each other, as defined by De Deyne and Storms (2008), is taxonomic categories in which words are connected taxonomically in more general terms (superordinate), more specific terms (subordinate), or have a shared superordinate meaning (coordinate). Words are also linked as a synonym or antonym of the target word, or as a specific example of it. *Mobil* 'car', for instance, has *sedan* 'sedan' as a subordinate and *sepeda motor* 'motorcycle' as a coordinate as these are linked through the superordinate *kendaraan* 'vehicle'. An example of words related to the antonym and synonym, respectively, is *tanya* 'ask' and *balas* 'response' for the word *jawab* 'answer', and the word *Toyota* is an individual example of the word *mobil* 'car'.

2.4 Introspective features

The feelings of speakers toward certain words labeling a concept can be a reason for an association. People's positive or negative evaluations of a word and their emotions regarding a concept are two of the several types of introspective features in De Deyne and Storms' (2008) classification. *Nyaman* 'comfortable' and *damai* 'peaceful' are emotions speakers may feel when thinking about the word *house*. Meanwhile, *ramah* 'friendly' could be an evaluation of the *sopir* 'driver' concept. Representation, one of the introspective feature subtypes, refers to the representational state in the mind of a situational participant, such as their beliefs, goals, desires, ideas, and perceptions (De Deyne and Storms 2008). The word *patah* 'broken' associated with the word *sayap* 'wings' is an example of representation. The possibilities of how a concept denoted by a word can have happened (cause and effect) can be grouped as contingency, such as the *asap* 'smoke' that is caused by an *api* 'fire'. Words can relate to other words through comparisons that require cognitive operation, such as defining a *kuda* 'horse' that is similar to a *keledai* 'donkey'. Introspective features also include word associations that negate a property, such as a *burung unta* 'ostrich' that *tidak bisa terbang* 'cannot fly'. Finally, another subcategory of introspective features is quantity, in terms of numerosity, frequency, intensity, or typicality of an introspection or one of its properties (De Deyne and Storms 2008), such as the *matahari* 'sun' that elicits the word *sangat panas* 'very hot'.

2.5 Lexical features

De Deyne and Storms (2008) posited another feature of how words are associated with one another: lexical feature. It explains that words are connected to each other by their lexical features, such as the words that are located after or before (forward and backward completion) the word, its fragment, or orthographic similarity; for example, *boneka* 'doll' is a forward completion of the word *rumah* 'house', *susun* 'stack' is a backward completion of the word *rumah*

'house', *mata* 'eye' is a fragment of the word *kacamata* 'glasses', and *baku* 'basic' and *bahu* 'shoulder' are related as they share similar orthographic forms.

Furthermore, De Deyne and Storms (2008) added that words can also relate to one another through mediation or because they are an expression used in language; for example, *kapur* 'chalk' and *semen* 'cement' are mediated by the concept *pembangunan* 'construction', and *bunga* 'flower' and *desa* 'village' are related through an expression in Bahasa Indonesia (i.e., to denote the most beautiful lady in the region — *bunga desa*). Comments about the characteristics of words, called metacomments, are also considered word associations; for example, the word *duplikasi* 'duplication' comes to mind for the word *kupu-kupu* 'butterfly'.

3. Method

3.1 Participants

Forty-five adult native speakers of Bahasa Indonesia (20–29 years old) participated in the study selected by a purposive sampling technique. As all were Universitas Indonesia students, it was assumed they had a wide Bahasa Indonesia vocabulary. Participants in this study joined the experiment on their own accord after viewing the advertisement broadcast through WhatsApp. Nevertheless, before the experiment began, they were additionally asked if they were willing to be involved in the experiment. No participants were language impaired or hearing impaired and all had normal or corrected-to-normal vision. Moreover, since this study was conducted in Universitas Indonesia, the participants' ethnicity varied. Most participants were Javanese, some others were Sundanese, Betawi, and Minang. This research was approved by DRPM (Directorate of Research and Community Service) of Universitas Indonesia.

3.2 Materials

This study used 30 nouns as cue words taken from previous research done by Sari (2019). The words were ascertained as nouns based on the dictionary *Kamus Besar Bahasa Indonesia V* (KBBI V). Brysbaert, Wijnendaele and De Deyne (2000) noted that the earlier individuals acquire words, the easier they are to access from the lexicon; therefore, with this in mind, this study selected words that are acquired earlier in life. The age of acquisition of the cue words is five years, or in other words, they have already been acquired by five-year-old children. It was confirmed by five kindergarten teachers who teach 34 children aged five years, that those cue words have been acquired by their students. The cue words were selected based on nouns that exist in close proximity to everyday life. Table 1 displays all the cue words used in the study.

Cue words		
<i>bola</i> (ball)	<i>lantai</i> (floor)	<i>lilin</i> (candle)
<i>pasir</i> (sand)	<i>asap</i> (smoke)	<i>paku</i> (nail)
<i>tempat</i> (place)	<i>besi</i> (iron)	<i>sopir</i> (driver)
<i>rumah</i> (house)	<i>roda</i> (tire)	<i>balap</i> (race)
<i>jawab</i> (answer)	<i>nilai</i> (score)	<i>karpét</i> (carpet)
<i>kaki</i> (foot)	<i>kapur</i> (chalk)	<i>kaus</i> (t-shirt)
<i>bunga</i> (flower)	<i>susun</i> (stack)	<i>sayap</i> (wing)
<i>mobil</i> (car)	<i>karung</i> (sack)	<i>setir</i> (steering wheel)
<i>gambar</i> (picture)	<i>cabai</i> (chili)	<i>badut</i> (clown)
<i>tangan</i> (hand)	<i>senam</i> (gymnastics)	<i>bantal</i> (pillow)

Table 1: Nouns that are in close proximity to everyday life as cues

3.3 Procedure

The research took place in a classroom. The participants were shown the cue word on a screen for fifteen seconds and asked to write down as many words or concepts as they could think of after reading the cue. However, for this study, only the first responses were used to ensure the analyzed associated words were not influenced by words other than the cues or other words accessed after the cues. The cue words were presented randomly.

3.4 Data processing and analysis

The collected data — the first words written by participants after viewing the cue words — were classified by the first rater (i.e., the first author) based on semantic taxonomic coding adapted by De Deyne and Storms (2008). For example, it was decided that the word *cepat* 'fast' was connected to the cue word *asap* 'smoke' through its manner properties; in other words, as *smoke* moves *fast* in certain situations, the word *cepat* 'fast' was classified as situation-manner. Once the category for each response was determined, the classification was subsequently assessed by four different assessors to ensure the validity of interpretation. All assessors were lecturers in the language and linguistics field whose

ages ranged between 30 and 58 years. Each of the four assessors was given a questionnaire in the form of tables containing responses of participants that had already been categorized. They were required to assess whether the categorization in the tables was correct or not. If the latter is applied, then the assessors were asked to classify the response based on De Deyne and Storms' (2008) classification. The classification that has been completed was determined as valid if three out of four assessors confirmed them as correct.

4. Word association categories of responses

The response received from participants was around 1,350 responses. However, two responses were excluded as one participant did not give any response to one word and the other response was not related to the cue word at all — the word *dari* 'from' for cue word *senam* 'gymnastics'. Another 33 remaining categorized responses could not be counted as valid as the number of raters agreeing to a category was fewer than three people.

Based on the categorization adapted by De Deyne and Storms (2008), most of the 1,315 responses that were confirmed as valid by the assessors were related to the cue word by their entity properties (25.9%). For example, *tangan* 'hand' and *jempol* 'thumb' were followed by responses related to their lexical features (23.7%), such as responses *tanggal lahir* 'date of birth' for cue word *tempat* 'place'. From Figure 1, it can additionally be noted that situation feature was the category to which a large number of responses were related (21.3%) after lexical features. The word *cincin* 'ring' was an example of a response that was related by situation feature to the cue word *tangan* 'hand'. As much as 18% of the responses were categorized to be related to the cue words according to their taxonomical categories, such as *permadani* 'rug' for the cue word *karpet* 'carpet'. The remaining responses (10.8%) were associated with the cue words by the mental state attached to the cue words, named by De Deyne and Storms (2008) as introspective features — *nilai* 'score' and *jelek* 'bad'.

As the majority of cue words were associated with entity and lexical features, the results indicated that the participants tended to first associate the cue words, nouns in this case, with the information related to the concept characteristics as an entity than with their application in a phrase, which suggested that younger Indonesians may store words close to words that share the same properties or entities connected to the cues or their collocations. The relatively few words associated with introspective properties might have been an indication that these words were not associated with any particular mental state in their lexical experiences.

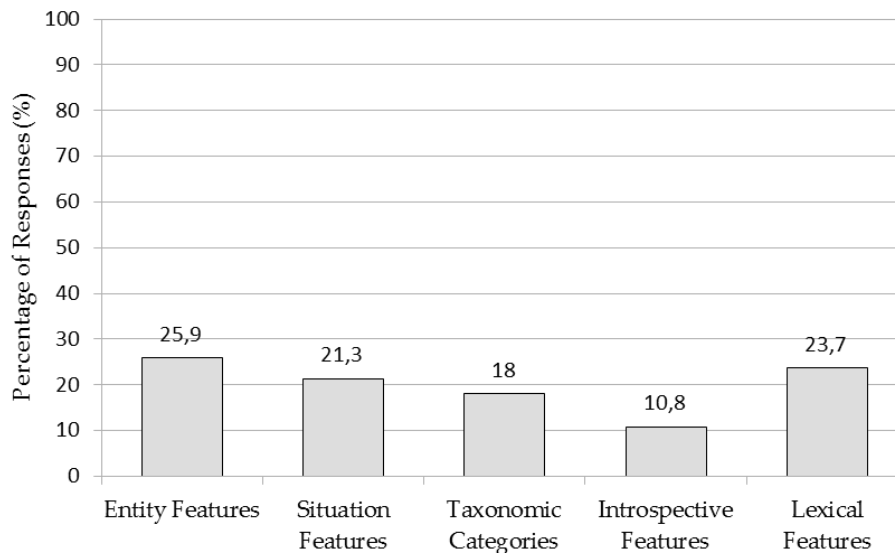


Figure 1: Elicited words sorted into main classes

4.1 Entity features

As mentioned before, most words elicited from the participants were associated with the cue words by their entity features. Around 41.3% of the responses to the cues that were related to entity was classified as being linked by the external surface features of the cue words; for example, the words *karpét* 'carpet' elicited the words *berbulu* 'furry', *kasar* 'rugged', *keras* 'hard', *merah* 'red', and *tebal* 'thick', all of which were related to the surface characteristics — the texture, color, and the size — of the carpet. The second largest entity feature responses (11.7%) were associated with parts of the cue words; for example, the word *kendaraan* 'vehicle' and *mobil* 'car' were evoked by *roda* 'tire'. A further 11.1% of responses were associated with materials; for instance, the word *lantai* 'floor' evoked *marmèr* 'marble' and *ubin* 'tile'.

Two subtypes in this category had similar proportions, namely systemic features and internal surface features, at 8.5 % and 8.2%, respectively. A systemic feature example was the cue word *meleleh* 'melt' being associated with *lilin* 'candle', and an example of an internal surface feature was the word *pedas* 'spicy' being associated with *cabai* 'chili'. The external component was the feature by which 7.6% responses are related to the cue words. For instance, the word *kelopak* 'petal' is the external component of *bunga* 'flower'.

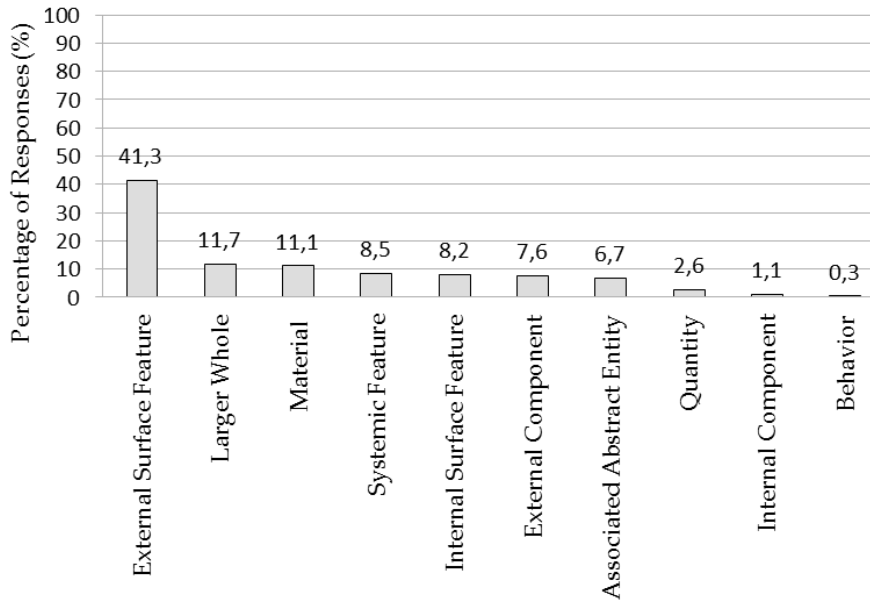


Figure 2: Percentage of responses classified in the entity features category

Entity features, which are abstract association properties, such as *kuntilanak* (a ghost that comes from an Indonesian myth who is deterred by sticking a nail on its head) for the cue word *paku* 'nail' and *karung* 'sack' with *lomba* 'contest' (sacks are often used in contests in Indonesia, especially while celebrating Independence Day), were found in 6.7% of responses. Words denoting quantity were found in 2.6% of cue word responses, such as *dua* 'two' and *sepasang* 'pairs' from the cue word *kaki* 'foot'. Internal component connections were only found in 28 out of 341 responses (9.8%), such as *putik* 'pistil' from the cue word *bunga* 'flower'. There were very few behavioral entity responses (0.3%); for example, the term *dapat berputar* 'able to spin' was given in response as a natural characteristic of *roda* 'tire'.

These results gave some indications regarding the thinking of young Indonesians on the identification of entity features. As there was a higher proportion of external concept responses and less focus on the internal aspects of the entity, it could be discerned that these educated young Indonesians may have stored/defined words related to an entity from the outside to the inside and then to the natural behavior. However, more detailed research would be necessary to confirm this supposition because the very little published research in Bahasa Indonesia discusses word storage from different perspectives not directly comparable to the one employed here (as in Pranoto and Afrilita 2018).

4.2 Lexical features

As can be seen in Figure 3, the majority of the words associated with cue words were forward completion (72.1%), followed by backward completion (16.7%). Almost all the cue words that evoked forward or backward completion suggested that these young educated Indonesian participants were able to easily access the associated collocates. Furthermore, as most responses were associated with collocates located to the right of the cues, this might have been because of the Bahasa Indonesia phrase pattern *Diterangkan Menerangkan* or Head + Modifier (Suharianto in Dewi 2013), which locates the head to the left of the modifier. Nouns in Bahasa Indonesia can be used as both a head and a modifier in a phrase. Nevertheless, the result showing that most of the participants think of words that modify the cue words indicates that most participants consider the cue words as the headword of a phrase. When the cue word *rumah* 'house' was given, for example, participants were more likely to mention other words whose role is as the modifier of the cue words, such as *makan* 'eating' – restaurant; *sakit* 'sick' – hospital; or *duka* 'sad' – funeral home, rather than convert the cue words to the modifier by mentioning *atap* 'top'.

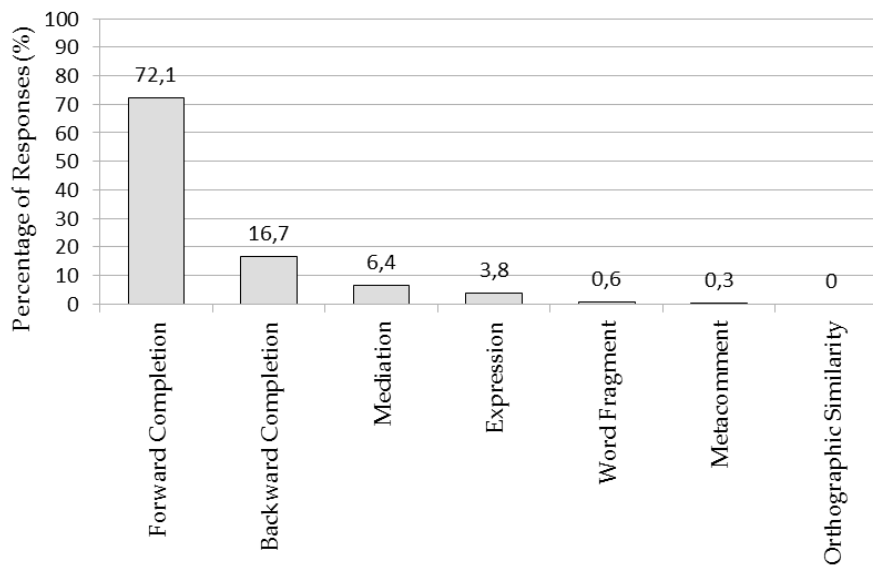


Figure 3: Percentage of first word association connected to the cues by lexical features

Around 6.4% of the responses in the lexical features categories were connected to the cues through situation or entity mediation; for example, the word *asap* 'smoke' induced the word *awan* 'cloud' as both words were mediated through

their shapes — the shape of smoke is similar to clouds. A few words (3.8%) stimulated by the cues were categorized as expressions since the cues were used in idiomatic phrases to express concepts apart from the literal meaning; for example, the cue *tangan* 'hand' evoked the word *panjang* 'long', which in combination is an expression in Bahasa Indonesia for a thief, thereby indicating that these expressions are closely stored. However, there were fewer words that were arising from the cue words associated with fragments (0.6%) and metacomments (0.3%), while there were no responses that were related due to orthographic similarity to the cue words, although the stimuli were presented in their written forms.

4.3 Situation features

As mentioned, 20.51% of the words related to the cue words had situation properties. As Figure 4 shows, most of these responses (30.4%) were related to another object in a situation; for example, the word *bola* 'ball' evoked the word *gawang* 'goal post'; or location (60/283 responses or 21.2%); for example, *pantai* 'beach' and *laut* 'sea' were associated with the cue word *pasir* 'sand'. There were 15.5% of words denoting function from the cue words, such as *menulis* 'to write' from the cue word *tangan* 'hand'.

Around 6.4% of responses identified people closely related to the cue words, such as *sopir* 'driver' from the cue word *setir* 'steering wheel' — named as a person. As with the person category, 18 responses were related to the cue words as they were the physical state of the cue words. Meanwhile, 6% of responses denoted actions, such as *mekar* 'bloom' from the cue word *bunga* 'flower.' There were few manner-related responses, namely *berurutan* 'in sequence' from the cue word *susun* 'stack' and time *mati lampu* 'power failure' from the cue word *lilin* 'candle' words, at only 4.2% and 3.5%.

The social artifacts and event classes appeared in even fewer situation feature responses (2.1% and 2.8%); for example, the word *karya* 'creation' was related to the cue word *gambar* 'picture' as a social artifact, and 17 *Agustus* (Indonesian Independence Day) was associated with the cue word *karung* 'sack'.

The classes with the fewest situation feature responses were living things, and buildings. The word *kucing* 'cat' was linked to the word *pasir* 'sand' as cats use sand as their toilet, and only one cue word elicited a building in their relation; *rumah* 'house' elicited by *paku* 'nail'.

The data analysis revealed that participants most often associated the words based on the situation with other objects in a certain situation, the location, and the function, which suggested that other features were not closely attached to the word in the mental lexicon.

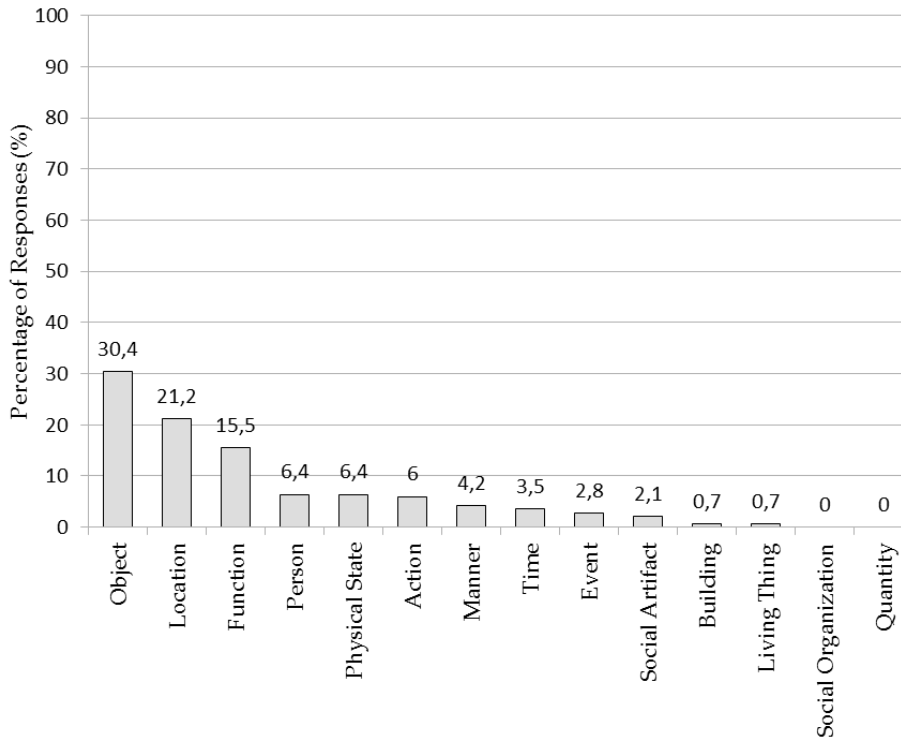


Figure 4: Responses related to the cue words by situation

4.4 Taxonomic categories

Besides associating words with their lexical features, entity, and situation properties, people often recalled other words that had taxonomic connections. The majority of responses (27.4%) that connected the cue words by the taxonomic category were subordinate to the cue words (Figure 5); for example, *anggrek* 'orchid', *cemara* 'pine', *matahari* 'sun', *mawar* 'rose' and *melati* 'jasmine' for the cue word *bunga* 'flower'. Superordinates, such as *logam* 'metal' from cue word *besi* 'iron', were elicited by 23.2% of responses in this category, and coordinates (e.g. *debu* 'dust' and *tanah* 'soil' from the cue word *pasir* 'sand') were elicited in 16.9% of the responses.

Cue word antonyms emerged in the taxonomic category responses more often than coordinates — 43 times out of 237, for example, *acak* 'random' was given as the opposite of the cue word *susun* 'stack'. Synonyms and specific cue examples were given in 7.09% of responses in this category, such as *kemudi* that had the same meaning as *setir* 'steering wheel' and *pengemudi* 'driver' that had a similar meaning to *sopir* 'driver'. Examples of words given to designate a spe-

cific example of the cue words were Dufan, McD, and Joker for the cue word *badut* 'clown'.

The result shows that it could be possible that the participant group stored words in more taxonomic ways than in meaningful relationships with other words, except for antonyms, since superordinates, subordinates, and coordinates of the cue words emerged far more often than their synonym and the specific examples. Moreover, the small number of specific examples suggested that these may be stored at quite a distance from the intended word and not common to anyone since, as mentioned earlier, mental lexicons are dependent on what has been experienced by a person.

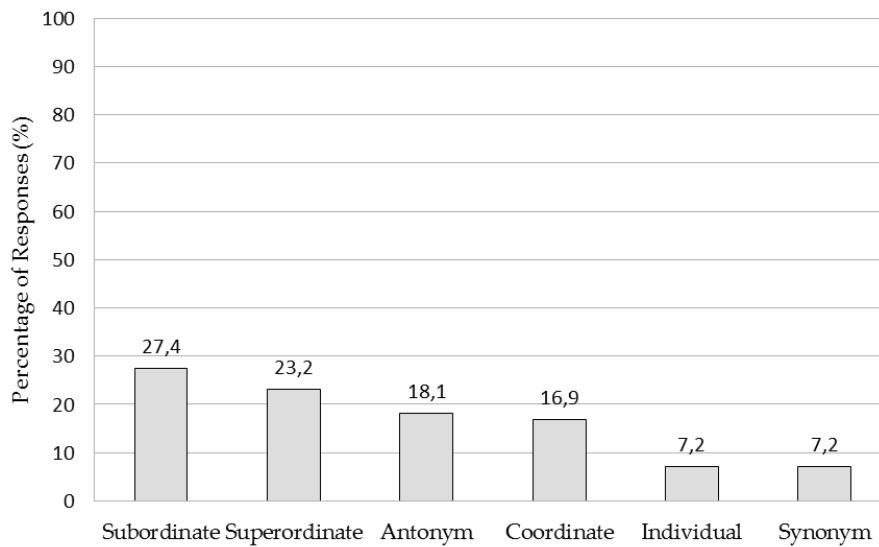


Figure 5: Proportion of the responses linked by the taxonomic categories

4.5 Introspective features

Introspective word association represents the mental state of the participants toward the cue words (De Deyne and Storms 2008). As Figure 6 shows, the majority (64.8%) of the responses in this category that had introspective features were related to the participants' positive or negative evaluations of the cues; for example, the word *mobil* 'car' elicited *keren* 'cool', *hedon* 'hedonic', *mahal* 'expensive', and *mewah* 'luxurious'. Cause and effect features — labeled as contingency features — were found in 21.1% of responses; for example *bugar* 'fit', *capek* 'exhausted', *keringat* 'sweat', and *sehat* 'healthy' from the cue word *senam* 'gymnastics'. Emotional features were found in 12.5% of responses; for example, *takut* 'afraid', *seram* 'scary', and *lucu* 'funny' were responses to the cue

word *badut* 'clown'. The representation feature was found in only a small number of responses (4/144 or 2.2%); for example, the cue word *sayap* 'wing' elicited *patah* 'broken' and *enak* 'tasty'.

These results indicated that the introspective responses were mainly in the form of participants' judgement, positive or negative, and that words related to the cue words for causal effect were stored closer than those related to emotion. The absence of responses that are related to the cue words by cognitive operation, negation or quantity may suggest that when thinking of nouns, Bahasa Indonesia speakers rarely associate them with something that does not exist within them or state their comparison to other nouns.

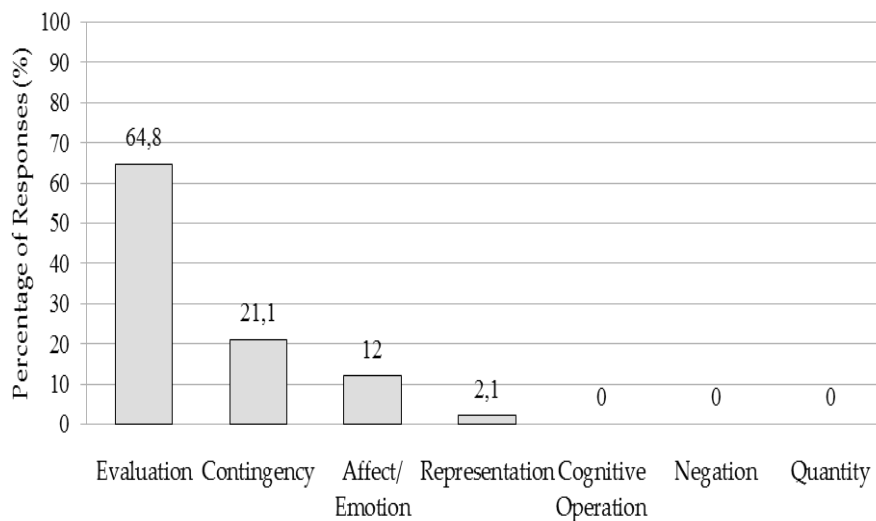


Figure 6: Introspective feature responses

5. Conclusion

Observing the way speakers of a certain language think about words and concepts can reveal the ways that these concepts are stored in the mental lexicon. This research has revealed information about how young native speakers of Bahasa Indonesia process words in their lexicon by connecting them with words that share lexical features or share one or more entity feature aspects. Generally, it was found that the Bahasa Indonesia participants, with reference to the entity features, perceive nouns based on their outside characteristics first. There was also a tendency for them to access nouns according to how they were used in speech, such as in phrases.

Therefore, the physical properties, situation, taxonomic categories, and feelings influenced how the words were processed in the mental lexicon. The appearance of these features in the retrieved words indicates the contexts attached to the words as conceptualized by native speakers of Bahasa Indonesia.

In conclusion, knowing how the native speakers of a certain language understand a word in their language can assist in formulating dictionary definitions so that language learners and native speakers can understand the best possible way that new words can be learned and stored. Lexicographers could use this type of research information to include more outside characteristics for the words listed in the dictionary, such as defining the meaning of a noun to ease the understanding of those particular words. Further, the words associated with the lemma in the dictionary entry could be used as the best example of how the lemma is used in a sentence and as a prototype of the language users' cognition.

Knowing how native speakers of Bahasa Indonesia understand a word in their language and its characteristics could also assist in developing a Bahasa Indonesia thesaurus to assist writers or learners understand how words in Bahasa Indonesia are used. The words generated in this study could also be used as examples of the domains in which the synonyms belong. Knowing how words are related to another in the minds of certain cultures could be a base to create a tool that relates words, such as WordNet, which has not yet been developed for Bahasa Indonesia.

However, as the cue words used in this study were only nouns, different results were found from the extensive study done by De Deyne and Storms (2008). Further, although the cue words were all nouns, it is undeniable that there might be a possibility that the participant consider some of those as verbs since the classification of word categories according to the speakers of Bahasa Indonesia is vague. This is corroborated in Jewalani's study (2019), which found that many of the participants — Bahasa Indonesia speakers — mistakenly classified the category of target words, or in other words, different from what is stated in the KBBI V. According to Jewalani (2019: 71-72), this phenomenon demonstrated that the theoretical system of Bahasa Indonesia, especially regarding word class category, is different from its implementation in the Bahasa Indonesia speech community. Thus, the upcoming studies should take this possible phenomenon into consideration.

In addition, as the participants were educated young people aged 20–29, there is a wide scope for future studies to add to these results and develop a database of Bahasa Indonesia lexical associations. Future research could focus on wider age and socio-economic groups in Indonesia and also assess whether there are any word association differences between the many ethnic Indonesian groups, all of which would deepen understanding of the way that Bahasa Indonesia is lexically stored.

Acknowledgements

The authors are very thankful for the financial support given by Direktorat Riset dan Pengabdian kepada Masyarakat (DRPM) Universitas Indonesia through Publikasi Internasional Terindeks untuk Tugas Akhir Mahasiswa (PITTA) grant 2018.

We also would like to express our special thanks to Sansiviera M. Sari for her major contribution to data collection.

References

Dictionaries

- Badan Pengembangan dan Pembinaan Bahasa.** 2016. *Kamus Besar Bahasa Indonesia (KBBI) V* (Version 0.2.1 Beta (21)) [Mobile application software]. Retrieved from <http://play.google.com/store/apps/details?id=yuku.kbbi5>.
- Endarmoko, E.** 2007. *Tesaurus Bahasa Indonesia* [Thesaurus of Bahasa Indonesia]. Jakarta: Gramedia Pustaka Utama.

Other References

- Bahar, M., A.H. Johnstone and R.G. Sutcliffe.** 1999. Investigation of Students' Cognitive Structure in Elementary Genetics through Word Association Tests. *Journal of Biological Education* 33(3): 134-141.
- Barsalou, L.** 2003. Situated Simulation in the Human Conceptual System. *Language and Cognitive Processes* 18(5-6): 513-562.
- Barsalou, L.W., A. Santos, W.K. Simmons and C.D. Wilson.** 2008. Language and Simulation in Conceptual Processing. De Vega, M., A.M. Glenberg and A.C. Graesser (Eds.). 2008. *Symbols and Embodiment: Debates on meaning and cognition*: 345-283. Oxford: Oxford University Press.
- Brybaert, M., I. van Wijnendaele and S. de Deyne.** 2000. Age-of-acquisition Effects in Semantic Processing Tasks. *Acta Psychologica* 104(2): 215-226.
- De Deyne, S. and G. Storms.** 2008. Word Associations: Network and Semantic Properties. *Behavior Research Methods* 40(1): 213-231.
- Dewi, J.** 2013. Wujud Kesalahan Bahasa Guru Bahasa Indonesia SMP Laboratorium Undiksha: Tinjauan Aspek Struktur dan Diksi. [Form of Errors of Bahasa Indonesian Teachers of Undiksha Laboratory Junior High School: Overview of Structural and Dictional Aspects]. *Jurnal Pendidikan Bahasa dan Sastra Indonesia Undiksha* 1(6): 1-15. Retrieved from <https://ejournal.undiksha.ac.id/index.php/JJPBS/article/view/991/859>.
- Geeraerts, D.** 1987. Types of Semantic Information in Dictionaries. Ilson, R. (Ed.). 1987. *A Spectrum of Lexicography: Papers from AILA, Brussels 1984*: 1-10. Amsterdam: John Benjamins.
- Jewalani, A.P.** 2019. *Pengaruh Kelas Kata pada Akses Leksikal Penutur Bahasa Indonesia dalam Kasus Tip-of-the-tongue* [The Influence of Word Categories on Lexical Access of Indonesian Speakers in Tip-of-the-tongue Cases]. Unpublished Thesis. Depok: Universitas Indonesia.

- Pranoto, B.E. and L.K. Afrilita.** 2018. The Organization of Words in Mental Lexicon: Evidence from Word Association Test. *TEKNOSASTIK* 16(1): 26-33.
- Sari, S.M.** 2019. *Defisit Semantik Leksikal pada Anak Autism Spectrum Disorder* [Lexical Semantics Deficit in Children with Autism Spectrum Disorder]. Unpublished Thesis. Depok: Universitas Indonesia.
- Van Sterkenburg, P. (Ed.).** 2003. *A Practical Guide to Lexicography*. Terminology and Lexicography Research and Practice 6. Amsterdam/Philadelphia: John Benjamins.