

## The difference in Lexical Meaning of ‘Female’ in the Malay Language: ‘Perempuan’ against ‘Wanita’: A Corpus-Driven Analysis

MUHAMMAD FARRIS IMADI FUZE [0000-0003-4253-373X]

*Center for Research in Language and Linguistics*

*Universiti Kebangsaan Malaysia*

*Selangor, Malaysia*

*p127290@siswa.ukm.edu.my*

### ABSTRACT

*Cognitive linguistics is an interdisciplinary branch of linguistics, combining knowledge and research from cognitive science, cognitive psychology, neuropsychology, and linguistics. One of the subjects often being studied under this discipline includes a phenomenon in which a lexical tends to be followed by other lexical contiguously with the word. Hence, to analyze this phenomenon, Sinclair (1991) and Louw (1993) introduced a linguistics approach known as semantic prosody. The usage of semantic preference will form clusters of meanings that will describe the common meanings in certain lexical or phrases. Meanings that were inherited from these semantic preferences will evoke cultural connotations that can reflect the social reality of the speaker. This article will discuss the collocation importance of ‘perempuan’ and ‘wanita’, two lexical that can be translated into the exact word in English: female. A collocation analysis was conducted on a text corpus in order to see which lexical often collocates with these two lexical. The collocation meaning of ‘perempuan’ and ‘wanita’ was evaluated based on the lexical list collocated with these two lexical. The result shows that ‘perempuan’ is often used to describe the gender or sexuality of a person, while ‘wanita’ is mainly used to describe something religious. ‘Perempuan’ also has a negative value but only when collocated with negative value pronouns. This research is hoped to add some value to the previous studies of collocation meaning as well as become a pioneer for upcoming studies.*

*Keywords: collocation meaning, corpus text, perempuan, semantic preference, Wanita*

To cite this article: Fuze, M. F. I. (2023). The difference in Lexical Meaning of ‘Female’ in the Malay Language: ‘Perempuan’ against ‘Wanita’: A Corpus-Driven Analysis. *International Journal of Language, Literacy and Translation* 6(2), 14-28. <https://doi.org/10.36777/ijollt2023.6.2.072>

To link to this article: <https://doi.org/10.36777/ijollt2023.6.2.072>

Received: 15 February 2023

Accepted: 1 August 2023

Published Online: 30 September 2023

This work is licensed under the terms of the Creative Commons Attribution (CC BY) (<https://creativecommons.org/licenses/by/4.0/>).

## INTRODUCTION

### SEMANTIC PROSODY

According to Zethsen (2021), the meaning of a context contributes to the meaning of a word in which certain words presuppose a certain context to such an extent that this context can be said to form part of the lexical meaning of that particular word. As a result, another approach to studying language from a broad perspective has emerged. The specific approach used to examine the usage of lexical items, known as “semantic prosody,” was introduced by Sinclair (1991) and Louw (1993). Louw defined this term, “semantic prosody” as a consistent aura of meaning with which a form is imbued by its collocates. Semantic prosodies have been largely inaccessible to human intuition about language and cannot be retrieved reliably through introspection (Louw, 1993).

The usage of lexical items phenomenon has already been detected by Firth (1957) since early 1950. His observation of several lexical items often used consecutively in sentences has led to the discovery of this phenomenon. According to him, if a specific lexical group is used together in a text frequently, then it is not impossible to find that same lexical group in other texts. Each lexical item has its tendency to be followed by certain words, which often go together with that lexical item, either being combined, being nearby, or being collocated.

### MEANING AND COLLOCATION

In conducting a study of meaning in a broad sense, Leech (1981) has proposed seven types of meaning, one of which is known as collocative meaning. Collocative meaning consists of the associations a word acquires on account of the meanings of words which tend to occur in its environment (Leech, 1981, p. 17). The main principle of collocative meaning is that the whole meaning of a lexical depends on the context. When a particular lexical collocates with another lexical, this situation could lead to a semantic prosody evaluation of positive or negative. According to Cheng (2013), Sinclair (1987) used to describe a word form that was likely to be followed by something positive or negative as “good/positive” or “bad/negative” semantic profiles.

However, do keeps in mind that this evaluation could also be in the neutral category. One of the fundamental tenets of linguistics is that in order to speak about a category, there must, at the very least, be a twofold opposition (dichotomy) of meaning and form. But according to Çepik (2006), there is a triple opposition (positive, neutral, negative) in the evaluative type of connotation, though examples of connotatively different synonyms with all three members of the opposition are difficult to find. Thus, a word can not only have a positive or negative value but also be in a neutral position.

Cepik's (2006) statement is also similar to Traugott and Dashers' (2002) statement. According to them, the collocative meaning is part of semantic change, a form of language change regarding the evolution of word usage, usually to the point that the modern meaning is radically different from the original usage. There are a few common types of semantic change, including amelioration and pejoration. Amelioration is a process whereby a positive meaning is associated with a word (Traugott and Dasher, 2002). On the other hand, pejoration is a process whereby a negative meaning is associated with a word (*ibid.*). A lexical item is in a neutral state when it does not show any potential to be in a certain collocation situation or is limited to a certain situation only, whether positive or negative

(ibid.). Only words that exhibit significant behaviour can be used to describe the phenomenon of prosody. This group of words is known as the dominant lexical unit, including nouns, verbs, and adjectives.

#### CORE MEANING OF 'PEREMPUAN' AND 'WANITA'

The elaboration of the core meaning of these two lexical items is based on two different dictionaries, *Kamus Pelajar Bahasa Malaysia* and *Kamus Dewan*. Both of these dictionaries were published by *Dewan Bahasa dan Pustaka*, in which *Kamus Pelajar Bahasa Malaysia* was published in the year of 1992 (2<sup>nd</sup> edition) while *Kamus Dewan* was published in 2010 (4<sup>th</sup> edition).

According to *Kamus Pelajar Bahasa Malaysia* the 7<sup>th</sup> edition, the word 'perempuan' has two meanings. The first meaning is a type of human that acts as partner for man or a type of human that generally will or able to give birth (*sejenis manusia sebagai pasangan kepada lelaki atau sejenis manusia yang umumnya akan atau dapat melahirkan anak*). 'Perempuan' also is defined as live partner for man or in other word, wife (*teman hidup bagi lelaki, isteri*). The word 'wanita' is used to refer a woman in a more subtle way (*merujuk kepada orang perempuan dengan cara lebih halus*). This word is also being defined as an antonym for man (*lawan bagi lelaki*).

According to *Kamus Dewan* the 4<sup>th</sup> edition, the word 'perempuan' has the meaning of woman (*wanita*) and antonym for man (*lawan bagi laki-laki*). 'Perempuan' is also defined as a wife (*isteri*). Additionally, this dictionary defines 'wanita' as a subtle form of the word woman (*bahasa halus bagi orang perempuan*). The antonym for this word is man (*lelaki*). Both dictionaries agreed that the word 'wanita' has more positive value than 'perempuan' as both referred to this word as a subtle form of 'perempuan'. In addition, both dictionaries also agreed that the word 'perempuan' is defined as a partner or wife for a man.

#### RESEARCH QUESTIONS

This research aims to answer the following questions:

- (1) What types of words are usually collocated with the words 'perempuan' and 'wanita'?
- (2) How can we define the collocative meaning of the words 'perempuan' and 'wanita'?

#### RESEARCH OBJECTIVES

The objective of this research is:

- (1) To determine the words that frequently collocate with the words 'perempuan' and 'wanita'.
- (2) To analyse the collocative meaning of the words 'perempuan' and 'wanita' by using the "Keyword in Context" (KWIC) approach.

#### LITERATURE REVIEWS

The ability to call up word combinations rather than individual words and the existence of tools, such as mutual information for establishing relationships between co-occurring words, means that it is now feasible to treat phrases and collocations more systematically than was previously possible (McEnery & Wilson, 1996). Collocation analysis has become an important tool in various linguistics fields, and the combination of such analysis with a computational corpus-driven approach will produce dynamical linguistics research.

A certain word in a language could possess positive or negative semantic prosody according to the collocational situation or word sharing (Abdul Rahim, 2005). For example, Stubbs (1995) found that the word 'cause' has a negative semantic prosody while the word 'provide' has a positive semantic prosody. This conclusion was made based on a collocation analysis using the Cobuild corpus, which has 120 million words. Collocation is not limited to the field of semantics. According to Galingging and M.Hum (2021), understanding collocations in the translation process is part of discovering meaning, and failure to understand collocations may result in a bad or nonsense translation. However, this research will solely focus on collocation analysis in lexical semantics.

The collocative meaning of 'perempuan' and 'wanita' has been studied before by some scholars. While each scholar has their own approach to explaining this problem, all the research focused on the collocation meaning produced when both lexical items collocate with others. According to Abdul Ghani and Jabbar (1993), both 'wanita' and 'perempuan' should not be synonymized with each other in terms of strict synonymy. This is due to the fact that each word has its own domain of usage according to its linguistic context and situation. According to them, it is only in cases of loose synonym that we can accept that both 'perempuan' and 'wanita' have the same meaning.

Another piece of research confirmed the difference between these two words. Abdul Rahim (2005) has concluded that the word 'wanita' is usually used in a formal context, in which this word is more likely to be collocated with words related to economy, politics, stature, career, and so on. This indirectly causes the word "wanita" to acquire cultural connotations that are generally more positive, preventing it from collocating with pejorative or negative words. On the other hand, the word 'perempuan' seems to have a more basic and important function, which is to determine gender. This supposedly will put the word 'wanita' in a neutral state looking at the inherited connotation. However, it was found that this neutral connotation enables the word to collocate with pejorative words, which indirectly causes a negative impact on the word.

Both 'perempuan' and 'wanita' does not only exist in the Malay language but also in the Indonesian language. Even in the Indonesian language, scholars have a hard time determining the differences. According to Yuliawati (2018), the word 'wanita' is most likely to have been used in the past since it has a more honourable meaning, but nowadays, the word 'perempuan' is frequently being used as women activists popularized it. Her research has also shown that there are few semantic categories that enable us to differentiate these two words, including issues of equality, rights, and organization, as well as sexual relations.

## RESEARCH METHODOLOGY

This study has been using qualitative data in the form of a corpus. A corpus-driven analysis was conducted in order to fulfill the research's objectives. A corpus known as *Korpus Buku* (Book Corpus) was used as research material. This corpus contains more than 1 million words and is made up of various texts from various sources written in Malay. The reason why this corpus was used in this research is that, rather than focusing on one type of text only, it contains a wide variety of texts ranging from expository texts (texts that are intended to identify and characterise experiences, facts, situations, and actions), narrative texts (texts that are intended to entertain, instruct, or inform readers by telling a story), and argumentative texts (texts that are aimed at changing the readers' beliefs). By looking at these diverse texts, we will be able to see the difference in lexical meaning from a wider perspective.

The "Keyword in Context" (KWIC) approach was used in this study. Originally named by Luhn (1958), keyword in context (KWIC) is an analysis method that looks at a word and its every utterance in a text to find out its meaning and usage (Williams, 2010). This approach involves the use of concordance lines for the keywords or words studied, namely 'perempuan' and 'wanita'. Each concordance line used in this study has four or five words to the left and right of the keywords that make up the intended context. The concordance line used to analyse the two words is generated by a computer from the corpus studied. During the data collection process, a table was built that contained four parts: collocated words, left frequency, right frequency, and total frequency. These sections will be filled with data obtained through a computerised corpus-driven analysis method.

Antconc, a freeware concordance programme developed by Laurence Anthony, was used to analyse this corpus, and the data was displayed in the form of concordance lines. For the purpose of research, the version of Antconc 3.5.9 is used, which has a minor upgrade compared to the previous version of Antconc. According to Laurence Anthony's report, the 3.5.9 version had fixed a bug in the File View tool that caused highlighting and results to be mismatched when the context search option was used. Several guidelines were followed in the process of gathering data for this study, which is:

- (1) Only words within the noun and adjective categories were taken as the data for this research. Words belonging to categories other than these two will not be taken into account.
- (2) Only words that are collocated more than once either on the left or right of the Keyword in Context (KWIC) were taken into account. In other words, words that are collocated only once will not be taken as research data.
- (3) As long as the word meets criteria (1) and (2), then it will be used as research data despite carrying a negative or sensitive connotation.

Once the data collection process was complete, it was analysed further. Words that are collocated more than once were categorised into several subcategories. This categorization aims to narrow the scope of the data and make it easier for researchers to see what types of words are often collocated with both, 'perempuan' and 'wanita'.

## RESULTS AND DISCUSSION

### DATA TABULATION

Table 1 below lists the collocated words to the right and left of the words ‘perempuan’ and ‘wanita’ to discuss both keywords' collocation behavior and semantic prosody. As mentioned in the research methodology section, only words collocated twice or more with the keywords will be listed. However, it should also be clear that not all the words have been taken into account, even if the collocation occurs more than once. To further detail this research, only words from the noun and adjective groups will be accounted for in the analysis. According to Nitsaisook and Dhanesschaiyakupa (2019), nouns and adjectives in the structure of lexical collocations are content words that contain lexical meanings while function words such as preposition or auxiliary verbs are just for grammatical correctness only.

In addition, these groups of words could carry denotative and connotative meanings, in which both exist together and cannot be separated. Denotation is the precise, literal definition of a word that might be found in a dictionary, whereas connotation refers to the wide array of positive and negative associations that most words naturally carry with them. For each word in the noun and adjective groups, we are able to assign either a positive or negative value according to its connotative meaning. The list of words from both groups is shown in Table 1 below:

**Table 1**

*Frequency of Collocated Words with ‘Perempuan’ and ‘Wanita’*

Perempuan				Wanita			
Collocated Word	Left Frequency	Right Frequency	Total Frequency	Collocated Word	Left Frequency	Right Frequency	Total Frequency
saudara	26	13	39	Melayu	1	11	12
mu'min	10	5	15	muslimah	0	3	3
khusyu	2	2	4	musyrik	3	3	6
ketaatannya	2	2	4	keji	9	9	18
pengantin	11	4	15	tangis	2	0	2
nur	4	0	4	solehah	1	2	3
penuntut	3	0	3	ulama	1	3	4
soleh	2	0	2	moden	1	5	6
ahli	0	3	3	haid	0	3	3
kafir	0	3	3	ayu	2	0	2
umur	0	2	2	tudung	1	2	3
muslim	2	2	4	Allah	7	5	12
perkahwinan	8	2	10	wangi	2	1	3
lif	2	0	2	aurat	5	2	7
citarasa	0	2	2	pemimpin	11	3	14
motif	1	2	3	suami	2	8	10
hantu	2	0	2	lemah	2	2	4
sabar	2	2	4	Rasulullah	3	0	3
sikit	2	0	2	akhlak	0	2	2
suami	0	3	3	baik	19	19	39
bayi	2	0	2	cantik	4	12	16

tua	2	15	17	ratu	2	2	4
syurga	0	3	3	baju	3	2	5
budak	5	0	5	rumah tangga	0	2	2
aurat	2	0	2	halal	2	0	2
janda	0	2	2	mulia	1	2	3
tali	2	0	2	kampung	2	0	2
putera	2	0	2	lelaki	40	19	59
engkau	5	4	9	taat	1	2	3
anak	69	14	83	pakaian	3	2	5
mak	1	2	3	maruah	2	0	2
pintu	3	1	4	hartu	4	3	7
muka	3	1	4	dosa	2	1	3
kain	0	2	2	isteri	2	7	9
bijak	0	2	2	merdeka	0	4	4
kanak	6	4	10	segan	2	0	2
kaki	3	1	4	akal	1	2	3
kamu	7	9	16	penzina	2	2	4
kaum	7	0	7	mengahwini	8	2	10
muda	0	3	3	pemurah	0	3	3
penduduk	4	2	6	busana	6	0	6
makhhluk	0	2	2	beriman	4	9	13
nabi	1	2	3	malang	1	5	6
benda	0	2	2	muslim	2	4	6
orang	30	12	42	kehormatan	1	2	3
pelajar	3	0	3	budak	0	9	9
biasa	1	2	3	kaum	46	3	49
benar	3	2	5	drama	0	5	5
Allah	0	5	5	Seks	2	0	2
ibu	2	5	7	hak	7	7	14
baik	8	0	8	kebajikan	3	0	3
dia	13	11	24	adat	3	0	3
mereka	10	29	39	fatwa	3	1	4
air	2	0	2	perkahwinan	2	1	3
saya	4	3	7	karya	0	8	8
keadaan	3	0	3	seorang	32	6	38
aku	2	1	3	jiwa	1	2	3
kerja	2	0	2	makhhluk	0	4	4
strategi	0	2	2	polis	2	1	3
rumah	1	2	3	gambar	6	1	7
tahun	1	2	3	Islam	11	29	40
baru	2	0	2	ketua	10	3	13
				hukum	1	2	3
				kamu	14	9	23
				rupa	2	1	3

saudara	3	1	4
tubuh	3	0	3
jabatan	1	2	3
tenaga	2	1	3
makanan	3	1	4
ruang	0	3	3
hidup	6	3	9
mereka	29	39	68
orang	27	14	41
ilmu	1	3	4
aku	5	3	8
dunia	1	5	6
diri	9	2	7
sikap	2	1	3
rumah	3	6	9
anak	7	11	18
negara	3	4	7

□ nouns      ■ adjectives

The list of words collocated with the keyword will be followed by the total frequency of collocation either to the left or right of the keyword. The number of collocation frequencies will provide information about the collocation behaviour of the keywords that are being studied. This also reflects the tendency of Malay language users to use keywords with collocated words.

#### COLLOCATIVE MEANING OF ‘PEREMPUAN’

Based on the results of the corpus-driven analysis, it was found that the word 'perempuan' is more likely to collocate with nouns than adjectives. Based on the results of the corpus-driven analysis, it was found that the word 'perempuan' is more likely to collocate with nouns than adjectives. This lexical item also seems to collocate with individual related nouns, such as those listed in Table 2 below:

**Table 2**

*List of Individual Related Nouns that Collocated with ‘Perempuan’*

Individual Related Noun	
saudara (cousin/ related person)	anak (child)
mu'min (believer)	mak (mother)
pengantin (bride)	kanak (kid)
penuntut (student)	penduduk (citizen)
suami (husband)	nabi (prophet)
bayi (baby)	orang (people)



budak (kid)	pelajar (student)
janda (widow)	ibu (mother)
putera (prince)	ahli (member)

From the listed words above, it was found that the word 'perempuan' collocated the most with the words 'anak' (69), 'orang' (30), and 'saudara' (26) on the left side. On the right side, the words most often collocated with the keyword are 'anak' (14), 'saudara' (13), and 'orang' (12). This category of individual-related nouns can also be broken down into 3 subcategories: family, status and position, and religion. Table 3 below indicates the subcategories of the individual related nouns:

**Table 3**

*Subcategories of Individual-Related Nouns*

Subcategory	List of Words
Family	saudara, suami, bayi, budak, janda, anak, mak, kanak, ibu
Status and Position	pengantin, penuntut, putera, penduduk, orang, pelajar, ahli
Religion	mu'min, nabi

Based on Table 3 above, nouns with the family category seem to collocate more with the word 'perempuan' followed by the status and position categories. Nouns related to religion also seem to collocate with the keyword but in a lesser quantity compared to the other two subcategories. Other than individual related nouns, the keyword also seems to collocate with concrete nouns and abstract nouns, as listed in Table 4 below:

**Table 4**

*List of Concrete and Abstract Nouns that Collocated with 'Perempuan'*

Noun	List of Words
Concrete	lif (elevator) aurat (private part) tali (rope) pintu (door) muka (face) kain (cloth) kaki (feet) mahluk (creature) benda (thing) air (water) kerja (occupation) tahun (year)
Abstract	ketaatannya (obidience) nur (light) umur (age) perkahwinan (marriage) citarasa (taste) motif (motive)

---

hantu (ghost)  
 syurga (heaven)  
 Allah (god)  
 keadaan (condition)  
 strategi (strategy)

---

The collocation between the word ‘perempuan’ and concrete or abstract nouns seems to have the same value. In this situation, it can be said that ‘perempuan’ is in a neutral state in both of these noun categories. According to the analysis, the number of concrete and abstract nouns is the same, indicating that the word ‘perempuan’ can be paired with both types of nouns. Through the collocation analysis of the word ‘perempuan’ with various types of nouns, it can be concluded that the word ‘perempuan’ is used to describe the gender or gender of a particular noun as shown in Table 2.

This word also undergoes a process of amelioration, which refers to developing a meaning with a positive value. This can be proven through the lexical position of ‘perempuan’, which often collocates with lexical nouns that also have a positive value, such as ‘mu’min’, ‘ketaatannya’, ‘nur’, and ‘syurga’. Malay and Islam are two interconnected elements. In fact, this Malay-Islam relationship has been given special attention in Article 160(2) of the Federal Constitution of Malaysia, which details the definition of a Malay as a Muslim (Mohammad Ahmad, 2019). Hence, when we talk about Malay culture, we will also talk about Islamic culture. In Islamic culture, these words are considered to have a positive value and indicate a good aspect of a person who is identified as ‘perempuan’. Each word has a negative counterpart, like ‘mu’min’ vs. ‘mushrik’ (believer vs. idolater), ‘ketaatannya’ vs. ‘kemungkarannya’ (obedience vs. rebellion), ‘nur’ vs. ‘kegelapan’ (light vs. dark), and ‘syurga’ vs. ‘neraka’ (heaven vs. hell).

However, the word ‘perempuan’ also undergoes pejoration, a process whereby a negative meaning is associated with a word. This can be seen when it is paired with the word ‘hantu’. In Malay culture, ‘hantu’ or ghost is considered taboo and often synonymized with bad things. According to Saparudin, Md Sham, and Hamjah (2014), Malay people often considered hysteria to be some kind of disorder that was caused by the ghosts that have long existed in society. Since the word ‘hantu’ itself has a negative value, it will also have a negative collocative meaning of ‘perempuan’. Although the word ‘perempuan’ can undergo the pejoration process, the frequency is lesser than the amelioration process. The word ‘perempuan’ is also collocated with adjectives. However, the amount of collocation is not as much as nouns. Table 5 below indicates the words from the adjective group that collocated with ‘perempuan’:

**Table 5**

*List of Adjectives that Collocated with ‘Perempuan’*

Adjective	
khusyu (focus)	bijak (smart)
soleh (pious)	muda (young)
kafir (infidel)	biasa (normal)
sabar (patient)	benar (true)
sikit (little)	baik (nice)

tua (old)

baru (new)

Based on Table 5 above, only the words ‘kafir’ and ‘sikit’ seem to have a negative value, while other adjectives have a positive value. The collocative meaning of the word ‘perempuan’ is more likely to be towards the positive value than the negative value. This is contrary to most traditional opinions, which say this lexical item has a negative value. ‘Perempuan’ is used to describe the gender of a certain noun and has a positive connotation. However, when this word is used in everyday conversation, a negative connotation emerges. The presence of a pronoun in a sentence usually indicates that the sentence is conversational. Table 6 below indicates the list of pronouns that are collocated with ‘perempuan’:

**Table 6**

*List of Pronouns that Collocated with ‘Perempuan’*

<b>Pronoun</b>	<b>Left Frequency</b>	<b>Right Frequency</b>	<b>Total Frequency</b>
engkau (you)	5	4	9
kamu (you)	7	9	16
dia (she)	13	11	24
mereka (they)	10	29	39
saya (I)	4	3	7
aku (I)	2	1	3

The presence of a pronoun preceding or proceeding the word ‘perempuan’ will bring the meaning of a negative collocation, as some of them are considered impolite, like ‘engkau’ and ‘aku’. Some may argue that such pronouns may also indicate a form of intimacy. This can be seen through a study conducted by Dawi, Zainon Hamzah, and Hj. Mohd Kiram (2016) stated that the use of ‘aku’ among the community in Lahad Datu has its own peculiarity because there is an element of politeness, power, intimacy, warmth, and belonging in its use. However, for most people, the words ‘aku’ and ‘engkau’ are considered impolite. In fact, ‘aku’ is defined in Kamus Dewan as a rude word. Once these pronouns are collocated with ‘perempuan’, they will have a negative connotation. Although some positive pronouns are collocated with this word, the word ‘perempuan’ is considered to have a negative value in this context. This is because, compared to ‘wanita’, the term ‘perempuan’ accepts the presence of pronouns with a negative value to be collocated.

#### COLLOCATIVE MEANING OF ‘WANITA’

"Wanita," like "perempuan," appears to be more likely to collocate with nouns than adjectives. Nonetheless, the figure is less than "perempuan." Unlike ‘perempuan’, the word ‘wanita’ tends to collocate more with the word within the religion-related nouns. Table 7 shows the list of religion-related nouns that are collocated with the word ‘wanita’:

**Table 7**

*List of Religion-Related Nouns that Collocated with ‘Wanita’*

<b>Religion-Related Noun</b>	
muslimah (female followers of Islam))	musyrik (polytheist)
ulama (person expert in Islam)	tudung (hijab)
aurat (private part)	pemimpin (leader)
Rasulullah (God’s prophet)	akhlak (morality)
maruah (dignity)	dosa (sin)
penzina (adulterer)	muslim (male followers of Islam)
kehormatan (honour)	kebajikan (charity)
fatwa (ruling on a point of Islamic law)	Islam

Most religion-related nouns collocated with ‘wanita’ have positive value except for ‘musyrik’, ‘penzina’, and ‘dosa’. In this category, the words that appeared the most with "wanita" were "Islam" (40), "pemimpin" (14), and "Allah" (12). This indicates that the lexical item ‘wanita’ is often used to describe something related to religion, especially Islam. ‘Wanita’ also collocates with concrete and abstract nouns, as shown in Table 8 below:

**Table 8**

*List of Concrete and Abstract Nouns that Collocated with ‘Wanita’*

<b>Noun</b>	<b>List of Words</b>		
Concrete	tangis (tear)	makhluk (creature)	
	haid (period)	gambar (picture)	
	tudung (hijab)	rupa (appearance)	
	aurat (private part)	tubuh (body)	
	baju (shirt)	jabatan (department)	
	kampung (village)	tenaga (energy)	
	pakaian (cloth)	makanan (food)	
	harta (treasure)	ruang (space)	
	busana (fashion)	rumah (house)	
	drama (drama)	negara (country)	
	seks (sex)	karya (works)	
	Abstract	akhlak (morality)	kebajikan (charity)
		rumahtangga (household)	adat (tradition)
maruah (dignity)		fatwa (ruling on a point of Islamic law)	
dosa (sin)		jiwa (soul)	
merdeka (freedom)		hukum (law)	
akal (mind)		ilmu (knowledge)	
kehormatan (honour)			
sikap (attitude)			

The word ‘wanita’ seems more likely to collocate with a concrete noun than an abstract noun. A simple explanation that can be made from this situation is that when the word ‘wanita’ collocates with either a concrete or abstract noun, it will indicate that a female-entitled individual owns something. ‘Wanita’ also seems to collocate more with adjectives compared to ‘perempuan’ as shown in Table 9 below:

**Table 9**

*List of Adjectives that Collocated with 'Wanita'*

Adjective	
keji (vile)	cantik (beautiful)
solehah (pious)	halal
moden (modern)	mulia (noble)
ayu (pretty)	beriman (believe)
wangi (fragrant)	malang (unfortunate)
baik (nice)	hidup (alive)
taat (obedient)	

Based on Table 9 above, only the word ‘keji’ and ‘malang’ seems to have negative value while other adjectives have positive value. If the results are examined closely, the collocative meaning of ‘wanita’ is more positive than negative. This is in line with the traditional opinion stating that the word 'wanita' is more delicate compared to ‘perempuan’. In fact, Kamus Pelajar Bahasa Malaysia the 7th edition also defines the word 'wanita' as a reference to women in a more subtle way. A lexical item's meaning value, whether positive or negative, can be determined by observing its context in conversation. In most cases, pronouns are used in conversation. As a result, by comparing the value of collocation between the pronoun and the words "wanita" and "perempuan," we can determine which word has a positive and negative value. Table 10 below shows the list of pronouns that are collocated with ‘wanita’:

**Table 10**

*List of Pronouns that Collocated with 'Wanita'*

Pronoun	Left Frequency	Right Frequency	Total Frequency
kamu (you)	14	9	23

The lexical 'wanita' does not collocate much with pronouns when compared to ‘perempuan’. In fact, the only pronoun collocated with it is ‘kamu’, which happens to have a positive value, and this situation causes the lexical 'wanita' also to receive a positive value.

## CONCLUSION

The aim of this study is to determine the difference in meaning between the words 'perempuan' and 'wanita'. By adopting the Keyword in Context (KWIC) approach through a corpus-driven analysis, it was found that there is indeed some difference in meaning between these two lexical items based on the context of their use. The word 'perempuan' is more likely to be used to describe the gender of something. Meanwhile, the lexical word 'wanita' is widely used to describe something related to religion. The value of each word, whether positive, neutral, or negative, can be determined by collating these words with

pronouns. The word ‘perempuan’ collocated more with pronouns than ‘wanita’. Some pronouns collocated with ‘perempuan’ have a negative value, such as ‘engkau’ and ‘aku’. Hence, it will also receive a negative value in this collocative situation. On the other hand, the pronoun collocated with ‘wanita’ has a positive value, resulting in the word also receiving a positive value. On the other hand, it does not mean that the word ‘perempuan’ carries a negative meaning in all contexts; instead, the negative meaning is produced only when it collocates with these negative-value pronouns. In the previous studies, ‘wanita’ was assumed to be used in a formal and honourable context, as shown by Abdul Rahim (2005) and Yuliawati (2018). Through this research, it is found that the usage of the word ‘wanita’ is also being used in a religious context, particularly in an Islamic context.

## REFERENCES

- Abdul Ghani, R. & Jabbar, S. (1994). Perempuan, wanita, dan ...: satu kajian hubungan leksikal berdasarkan korpus. In: *Seminar Perkamusan Melayu*. Malaysia: Kuala Lumpur.
- Abdul Rahim, H. (2005). Impak konotasi budaya terhadap leksis: Satu kajian semantik berasaskan korpus ke atas perkataan ‘Perempuan’ dan ‘Wanita’. *Jurnal Bahasa*, 5(1), 83-111.
- Baharom, H.N. (2010). *Kamus Dewan edisi keempat*. Kuala Lumpur : Dewan Bahasa dan Pustaka.
- Çepik, S. (2006). Positive Neutral Negative Evaluation in Connotation. *IBSU International Refereed Multi-disciplinary Scientific Journal*, 1, 144-147.
- Cheng, W. (2013). Semantic Prosody. In book: *The encyclopedia of applied linguistics*. Oxford, England: Wiley-Blackwell Publishing Ltd. <https://doi.org/10.1002/9781405198431.wbeal1062>
- Dawi, N.F., Zainon Hamzah, Z.A. & Hj. Mohd. Kiram, N. (2019). ‘Aku’ dalam Norma Interpretasi Masyarakat di Lahad Datu, Sabah. *MANU Jurnal Pusat Penataran Ilmu & Bahasa*, 30, 151-168. <https://doi.org/10.51200/manu.v0i0.1941>
- Firth, J.R. (1957). *Papers in Linguistics, 1931-1951*. Oxford: Oxford University Press.
- Galingging, Y. & M.Hum. (2021). Kolokasi dalam Penerjemahan. *Dialektika: Jurnal Bahasa, Sastra dan Budaya*, 8(2), 98-111. <https://doi.org/10.33541/dia.v8i2.3725>
- Haji Omar, A. (2015). *Nahu Melayu Mutahir edisi kelima*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Husin, A., Tajudin, A., Jabbar, S., & Asraf. (1992). *Kamus Pelajar Bahasa Malaysia*. Kuala Lumpur : Dewan Bahasa dan Pustaka Kementerian Pendidikan Malaysia.
- Karim, N.S., M. Onn, F., & Haji Musa, H. (2015). *Tatabahasa Dewan edisi ketiga*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Leech, G. (1981). *Semantics The Study of Meaning*. Harmondsworth, Middlesex, England: Penguin Books Ltd.
- Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In: M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair*. Amsterdam, Netherlands: John Benjamins. <https://doi.org/10.1075/z.64.11lou>

- McEnery, T., & Wilson, A. (1996). *Corpus Linguistics an Introduction*. Edinburgh: Edinburgh University Press.
- Mohammad Ahmad, N. (2019). Definisi Melayu dalam Perlembagaan Malaysia dan Kesannya Terhadap Kefahaman dan Amalan Ajaran Islam di Malaysia. In: *Seminar Tamadun, Etnik dan Turath Islami (i-STET)*, 49-56. <https://doi.org/10.33102/sainsinsani.vol5no2.174>
- Nitsaisook, N. & Dhanesschaiyakupa, U. (2019). An Analysis of Collocations and Semantic Prosody in Thai Political News Headlines in the Bangkok Post. *International Journal of Language, Literacy and Translation*, 2(1), 26-38.
- Saparudin, I.f., Md Sham, F. & Hamjah, S.H. 2014. Pemahaman Konsep Histeria dalam Masyarakat Melayu. In: Puteh, A. & Mokhtar, A.I. (Eds.), *Seminar Antarabangsa Dakwah & Etnik 2014; Dakwah & Etnik: Perspektif Multi-disiplin*. Sabah, Malaysia: Universiti Malaysia Sabah.
- Sinclair, J.M. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Stubbs, M. (1995) Corpus Evidence for Norms of Lexical Collocation. In Cook, G. & Seidlhofer, B. (Eds.), *Principle & Practice in Applied Linguistics*. Oxford: Oxford University Press.
- Traugott, E.C., & Dasher R.B. (2002). *Regularity in Semantic Change*. United Kingdom: Cambridge University Press. <https://doi.org/10.1017/CBO9780511486500>
- Williams, R. (2010). Hans Peter Luhn and Herbert M. Ohlman: Their Roles in the Origins of Keyword-in-Context/Permutation Automatic Indexing. *Journal of the American Society for Information Science and Technology*, 61(4), 835-849. <https://doi.org/10.1002/asi.21265>
- Yuliawati, S. (2018). Perempuan atau Wanita? Perbandingan Berbasis Korpus tentang Leksikon Berbias Gender. *Paradigma Jurnal Kajian Budaya*, 8(1), 53-70. <https://doi.org/10.17510/paradigma.v8i1.227>
- Zethsen, K.K. (2021). Corpus-based cognitive semantics: Extended units of meaning and their implications for Translation Studies. *Linguistica Antverpiensia, New Series – Themes in Translation Studies*, 7. <https://doi.org/10.52034/lanstts.v7i.218>