

Exploring Gender Issues Associated With *wanita*/woman and *perempuan*/woman in Malaysian Parliamentary Debates: A Culturomic Approach

Jamaluddin Aziz

jaywalk@ukm.edu.my

Media and Communication Research Center,
Faculty of Social Sciences & Humanities,
Universiti Kebangsaan Malaysia

ABSTRACT

Language use and function can reveal tremendously about the society. In the context of Malaysian Hansard, parliamentary discourse is useful in understanding how certain issues are discussed, understood and resolved. For this paper, two Malay words *wanita*/woman and *perempuan*/woman, both refer to the English equivalent 'woman', are used to identify some Malaysian women issues as found in the Malaysian Hansard Corpus (MHC). The objectives of this research are: 1) To identify the frequency of both lexicals *wanita*/woman and *perempuan*/woman; 2) To identify the unique trend or pattern of the usage of these two lexical items; 3) To identify some issues related to women based on the trend identified. For that reason, Malaysian Hansard Corpus from Parliament 1 (P1) (year 1959) to Parliament 13 (P13) (year 2018) involving over 157 million words is used. The research employs both quantitative and qualitative research approaches. Culturomics approach and tool (AntCon software), are employed to identify the frequency and the trend by generating n-grams. The qualitative approach involves a discourse analysis of MHC based on the trend shown in the n-grams. The research finds that the nodes *wanita*/woman and *perempuan*/woman collocate rather consistently with the word *lelaki*/man. This means women issues are often compared to men's conditions and situations. There is also a shift in trend in terms of how *wanita*/woman and *perempuan*/woman have been discussed, that is, from women functioning in private sphere (collocates like *anak*/child and *mengandung*/pregnant to a more public one (collocates like *guru*/teacher and *gaji*/salary). However, there is no apparent shift in terms of women being defined as a biological category (sex) or a cultural category (gender). Cultural connotations of both words influenced the choice of words employed in the discourse that contribute towards our understanding of the changing trend of the way women issues have been discussed.

Keywords: binary structure; *wanita*/woman; *perempuan*/woman; Culturomics; Malaysian Hansards Corpus

INTRODUCTION

On September 17, 2019, a report on global gender equality produced by Bill and Melinda Gates Foundation was announced, reaffirming the problematic and unequal treatment of women worldwide. The report found that "about half a billion people worldwide still don't get basic health and education, and girls everywhere suffer disadvantage" (Channelnewsasia.com). The fact that this report highlights the vulnerable condition of girls globally should raise alarm on our own shore. While annual reports on gender equality have consistently been produced by other

global organizations including the United Nation Economic Forums, these global reports seem to show fluctuation; this directly points to the continuous vulnerability that women and girls have to endure every day.

Indeed, gender equality index on Malaysia echoes this global phenomenon. According to Malaysian Gender Gap Index 2017, Malaysia has improved its position to 70th of 144 countries, an improvement from 75th place in 2016 (nst.com.my). Of the four gender gap indicators, which are Women in Health, Women in Education, Women in Politics and Women in Economy, Malaysia is struggling to reach gender parity on Women and Politics. This, according to the Inter-Parliamentary Union, resulted in Malaysia being ranked “155 out of 188 nations in terms of women’s representation in national legislatures, below less developed Southeast Asian nations such as East Timor, Vietnam and Laos” (Lih Yi, 2018 (<https://www.reuters.com/article/us-malaysia-women-politics/women-sidelined-in-new-malaysian-government-despite-campaign-promises-idUSKCN1IM11U>)). And this situation continues, despite the historic change of government of Malaysia. Malaysia still fails to create 30 percent of women in decision making positions, including as policy makers.

Despite the fact that Malaysian women’s involvement in the country’s development has been well-documented, the trend seems to be imperviously stagnant and the causes for this stagnancy vary. Going against the popular belief that cultural practices have shackled Malaysian women, Charles Hirschman asserts that “the claim that Malaysian women lag behind men primarily because of traditional cultural values is challenged, however, by writings of area studies specialists and social historians about gender roles and equality in South-east Asia” (2016, p. 34). This is proven by women pioneering the political landscape of Malaysia since the pre-Independence period that is in 1955 with Halimahton Abdul Majid as the first elected woman Member of Parliament in the first general election, and the immediate post-Independence period, that is, with the appointment of the first woman senator, Aishah Ghani, in 1962, and in 1969, the first woman minister, Fatimah Hashim (Ministry of Women, Family and Community Development, 2012). Sadly, even with the purported new Malaysia, the number of women MPs in Malaysia is still below the 30% global inspiration.

The low representation of women in public spheres raises the question of the visibility of women issues and a logical place to understand this is to look at the parliament record. This paper suggests that one way of looking at this is by investigating the words *wanita*/woman and *perempuan*/woman – both are the equivalence of the word woman. The main reason why both words are used as the nodes for this research is because they are often used interchangeably in social and political contexts. For instance, a corpus study by Hajar Abdul Rahim (2005) that looks at the impact of connotative meanings of the semantic prosody on both words *wanita*/woman and *perempuan*/woman reveals that while the first is often associated with positive connotations, the latter is more neutral. Nonetheless, the neutrality of the word *perempuan*/woman exposes it to more pejorative terms, which “indirectly have negative effects on the word” (Hajar Abdul Rahim, 2005, p. 102). As the present study takes this information of semantic prosody into account, it rationalizes that the use of the nodes *wanita*/woman and *perempuan*/woman in this corpus-driven research is therefore important to reveal more about women issues and how they are represented in MHC.

PARLIAMENTARY DEBATES AND CORPUS RESEARCH

The interest in studying parliamentary discourse is increasing as parliament is a place where national and international issues are debated; subsequently, policies and laws are passed; thus, marking the vibrancy of democracy. Malaysian Hansard is an official document of Malaysian Parliament, which under “Section 81 of the Evidence Act (EA)” 1950 states that it is “presumed to be genuine” (Minderjeet Kaur, 2017). Ilie observes that “[d]iscourses enacted in parliament not only reflect political, social, and cultural configurations in an ever-changing world, but they also contribute to shaping these configurations discursively, cross-rhetorically, and cross-culturally” (2015, p. 1). Similarly, MHC is a rich discourse repository that provides researchers with a virtually limitless reservoir of data to mine from. As the background of Malaysian Hansard can be found elsewhere (see Kim Hua Tan et al., 2017; Norsimah Mat Awal et al., 2019), I therefore focus on some past studies that have been carried out on parliamentary debates and corpus studies.

While earlier parliamentary discourse studies were the major interests among political scientists and sociologists; latterly, interdisciplinary and cross-cultural approaches became more prevalent especially with contributions from linguists (Ilie, 2015). Indeed, there is an increase in the interest in doing research on parliamentary discourse especially within the domain of political discourse. This inevitably creates an intersection between language studies with other areas such as social and economy that allows more ideologically-driven analysis to be carried out. For example, employing gender perspective, Marion Loffler analyses “two debates on neutrality act in the Austrian parliament in 1955” to reveal that as “neutrality was metaphorically associated with femininity and infancy [...] neutrality could finally serve as a remasculinizing feature of the Austrian nation” (2019, p. 446). Another function of Hansard document is as a historical data especially employed by historians. For example, Sahul Hamid Mohammed Maidin (2017) studies the establishment of the propaganda agencies in Malaysia during Malaysia-Indonesia confrontation period from 1948 to 1965 using the Hansard data as a one of the main historical records. The study finds that the Malaysian government had successfully countered Indonesian propaganda messages and avoided war by strategically winning the hearts and minds of Malaysians with messages that appeal to their patriotism.

By the same token, as language elements are central to Parliament Hansard, parliament discourse can reveal the emotive aspect of the speech community. For instance, a study done by Rheault and et al. (2016) measures emotion in Parliamentary debates by using automated textual analysis. This study focuses on the emotions of politicians in British Parliament in response to economic recessions using Hansard data from 1909 to 2013, arguing that “politicians react emotionally to national and world events in a manner that is predictable” (p. 1) and that their emotions “affect their decisions” (p. 2). Rheault and his team were able to study the discourse of over 100 year’s period of the British House of Commons.

Likewise, technological tool makes it possible for larger data to be accumulated, constructed and analysed. Indeed, the affinity that language studies and corpus research tool has created more possibilities for research related to a large discourse data to be carried out. Hajar Abdul Rahim observes that “[i]n the last 30 years, the number of corpus-based studies has progressively increased in tandem with the development of language corpora (Malay and English) and also the accessibility of attested language materials in digitized form via online and internet sources” (2014, p. 7). As many of these studies are done by scholars in the areas of language and linguistics in Malaysia, it is not surprising that most focus on language elements

such as lexical (Hajar Abdul Rahim 2005; Noorli Khamis et al., 2018) and phrasal features (Kim Hua Tan et al., 2017). Technology therefore is central to the development of corpus studies.

A research team known as Digital Humanities group at the Universiti Kebangsaan Malaysia, in which I am part of, has come up with several studies utilizing the MHC. These studies (for example Kim Hua Tan et al., 2017; Norsimah Mat Awal et al., 2019) have been carried out on parliamentary discourses and they usually focus on the language aspect. For example, Kim Hua Tan et al. (2017) look at adjunctive and disjunctive phrases, using corpus-based methodology to find the usage patterns and trends of both phrases. Their research finds that Malaysian Hansard data has displayed a gradual increase in the use of adjunctive and disjunctive, revealing a certain discourse strategy employed by Malaysian politicians. The study of the *air/water* issue in Selangor state by Norsimah Mat Awal et al. (2019) reveals that with the word *air/water* consistently appears as the collocate for the word ‘Selangor’, water related issues such as water treatment and water catchment area are recurring issues. Norsimah Mat Awal et al.’s study is relevant to my own study as it can also reveal some issues related to the nodes *wanita/woman* and *perempuan/woman*.

Several corpus-based studies have been carried out that look at gender issues and representation (Kesumawati, 2013; Gralinski et al., 2016; Bouferrouk & Dendane, 2018); however, there is a dearth of studies that utilize culturomics as its approach. For instance, Kesumawati’s (2013) study explores the issues of methodology in demonstrating the ways gender-related themes are represented in a personal advertisements corpus. Almost in the same vein, Gralinski et al. (2016) proposed a new method of creating a corpus based on male/female first person expressions. Bouferrouk and Dendane (2018) use corpus to examine how a female muslim attire, the burka, is represented in one of the UK’s newspapers to expose the negative semantic prosody often associated with the attire. Coupled with the fact that the present study is also a corpus-driven research, this paper attempts to identify cultural patterns related to the nodes *wanita/woman* and *perempuan/woman* that can reveal how women or gender issues are discussed in Malaysian Parliament from 1959 (P1) to 2018 (p13). The objectives of this paper are:

- 1) To identify the frequency of both lexicals *wanita/woman* and *perempuan/woman*;
- 2) To identify the trend of the usage of these two lexical items;
- 3) To identify issues related to women based on the trend identified.

METHODOLOGY

This research employs both quantitative and qualitative approaches. Culturomics is an approach that, according to Jean-Baptiste Michel et al., enables the investigation into cultural trends quantitatively; this can be done by examining “both linguistic changes, such as changes in the lexicon and grammar; and cultural phenomena, such as how we remember people and events” (2011, p. 3). In this study, culturomics is employed to quantify the lexicons *wanita/woman* and *perempuan/woman* so that a cultural trend can be identified. Using AntConc software, the frequencies of each lexical item are recorded. The frequencies are generated to show how often the lexicals are used in Malaysian Parliament from September 1959 (Parliament 1 (P1)) to April 2018 (Parliament 13 (P13)) with 157 million words as the data of the corpus. From the frequencies, an n-gram is then generated that provides a visual representation of the cultural trend related to both *wanita/woman* and *perempuan/woman*.

With AntConc software, the concordances of *wanita*/woman and *perempuan*/woman can be generated. The concordances provide qualitative data that is further analysed using discourse analysis sing discourse analysis of the Hansard texts. These texts are selected based on the pattern of occurrence produced in the n-gram. Discourse analysis is a crucial part of the culturomic approach this research adopts as it provides a critical appraisal of the quantitative data. This is in line with Jean-Baptiste Michel’s argument that “the challenge of culturomics lies in the interpretation of this evidence” (2011, p. 7). Therefore, the concordances are useful as they help retrieve the context of the discussion or debate; thus, relating the collocate with larger social contexts. The importance of context in the study of language and gender is undeniable as argued by Victoria DeFrancisco: “to study gender and language we must at a minimum study it in context” (1997, p. 46).

The procedures of data mining follow the 4 step procedures proposed by Norsimah M. Awal et al. (2019) in order to meet the objectives of this study:



Step 1: This step involves generating the frequency of the lexical *wanita*/woman and *perempuan*/woman in each parliamentary sitting i.e from P1 through to P13 based on mutual information (MI) score of 6. The raw frequency for *wanita*/woman’ and *perempuan*/woman in each parliament was determined using AntConc software and the frequency presented for this study was then normalized. Evinson (in Norsimah Mat Awal et al, 2019, p. 105) asserts that “[n]ormalization is a process to standardise word count which is done to accurately compare corpora (or subcorpora) of different sizes.” The normalized frequencies are then plotted in an n-gram to identify the trend of occurrence for the word *wanita*/woman and *perempuan*/woman from P1 to P13.

Step 2: The next step is to generate the collocates of *wanita*/woman and *perempuan*/woman in each parliament sitting. Collocation is “a co-occurrence relationship between words or phrases. Words are said to collocate with one another if one is more likely to occur in the presence of the other than elsewhere” (McEnery, 2013 in Norsimah Mat Awal, 2019, p. 105). MI as a measure of collocation strength is employed in this study. A higher MI score indicates significant relationship between words. For instance, MI score of 6 is higher than MI score of 3, reflecting “a stronger association and thus a more coherent relationship between words” (Salazar, 2014 in Norsimah Mat Awal et al., 2019, p. 105).

As the steps are a duplicate of Norsimah Mat Awal et al.’s (2019) method, the study adopts the idea that the MI score is suitable with content words rather than function words. MI score of 3 or higher shows the frequency of the co-occurrence of the two lexical items. This study sets the minimum MI Score to 6, and item must have a minimum co-occurrence frequency of 10, within a 5 (left) - 5 (right) window span to figure out the collocate of nodes given. Using this setting, greater statistically significant differences of the collocations can be produced.

Step 3: This step strengthens the analysis by identifying the most consistent collocate/s of *wanita*/woman and *perempuan*/woman. The MI scores and the normalized co-occurrence

frequencies of the identified collocate/s were recorded manually. This is then plotted and presented in a n-gram graph.

Step 4: After trends are identified, using discourse analysis, certain issues are addressed based on the collocate/s and concordance/s of *wanita/woman* and *perempuan/woman*.

FINDINGS AND DISCUSSION

The research finding will be presented in three parts:

Part 1: Present the frequency of *wanita/woman* and *perempuan/woman* in the Malaysian Hansard Corpus (MHC).

Part 2: Present the collocates of *wanita/woman* and *perempuan/woman* in each parliamentary session in MHC.

Part 3: Present the discourse analysis of related concordances

PART 1: THE FREQUENCY

This section provides the data to fulfil the first research objective, that is, to identify the frequency of both lexicals *wanita/woman* and *perempuan/woman*. Information about the word frequency is important as it shows the tendency of the speech community to use the word in the discourse, reflecting the relevance of the word and its context in MHC.

Figure 1 and 2 show the trend of *wanita/woman* and *perempuan/woman* using normalized frequencies per one million of lexical items. For a guideline, the size (token of running number) of subcorpora (parliament) (Imran Ho et al., 2017) is presented in Table 1.

TABLE 1. The size (token of running number) of subcorpora (Parliament) in MHC

Parliament	No of Subcorpora
1	6060551
2	9893721
3	6264859
4	8040934
5	8691728
6	9485250
7	9106187
8	15171864
9	12919341
10	14123916
11	17047556
12	22188820
13	18517944

Based on Table 1, there is a stable increase in the number of tokens from P1 to P13. The highest number of tokens is in P12 (22188820), while P3 shows the lowest number of tokens (6264859). The rise in number of token could be influenced by the number of parliamentary proceeding of that particular parliament session. The highlighted parts are the data chosen for analysis in the present study. All the number exceed one million, thus in order to compare and plot in an n-gram graph format, all the data were standardized per one million using normalized frequency formula.

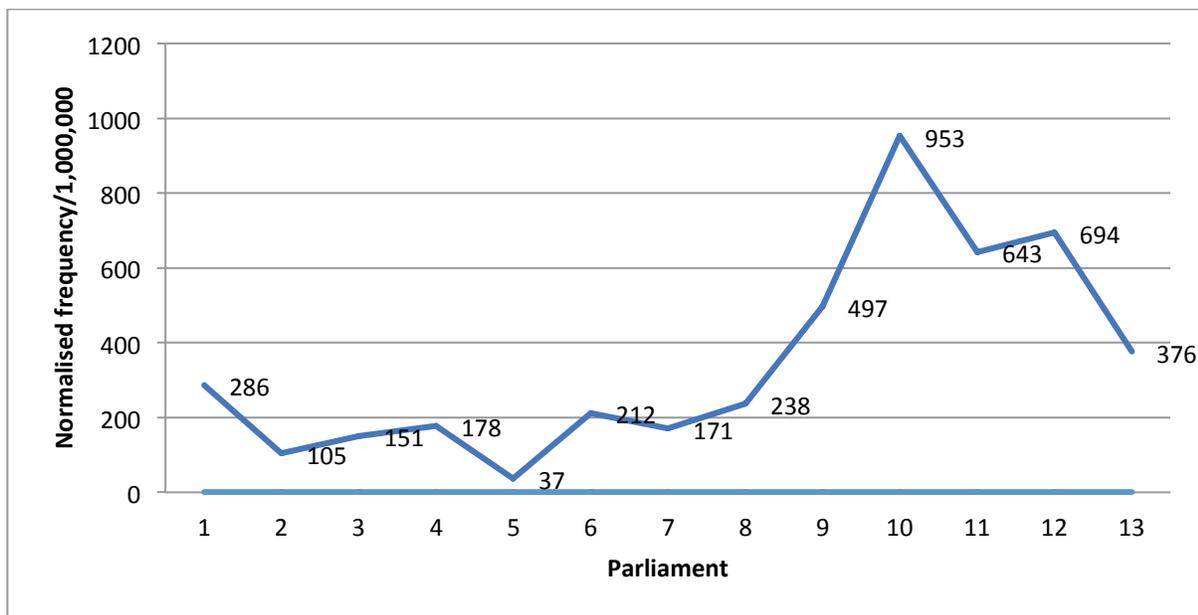


FIGURE 1. An n-gram of the trend of lexical *wanita/woman* in MHC

Figure 1 shows that the lexical *wanita/woman* has a consistent upward trend beginning from P7 (1986-1990) to P10 (1999-2003) as highlighted in Table 2, but has shown a downward trend afterward. The frequencies of *wanita/woman* are as the following.

TABLE 2. Frequencies of *wanita/woman* in P1-P13 by Year

Parliament	Freq	Year
1	286	1959-1964
2	105	1964-1969
3	151	1971-1973
4	178	1974-1978
5	37	1978-1981
6	212	1982-1986
7	171	1986-1990
8	238	1990-1994
9	497	1999-1995
10	953	1999-2003
11	643	2004-2007
12	694	2008-2012
13	376	2013-2018

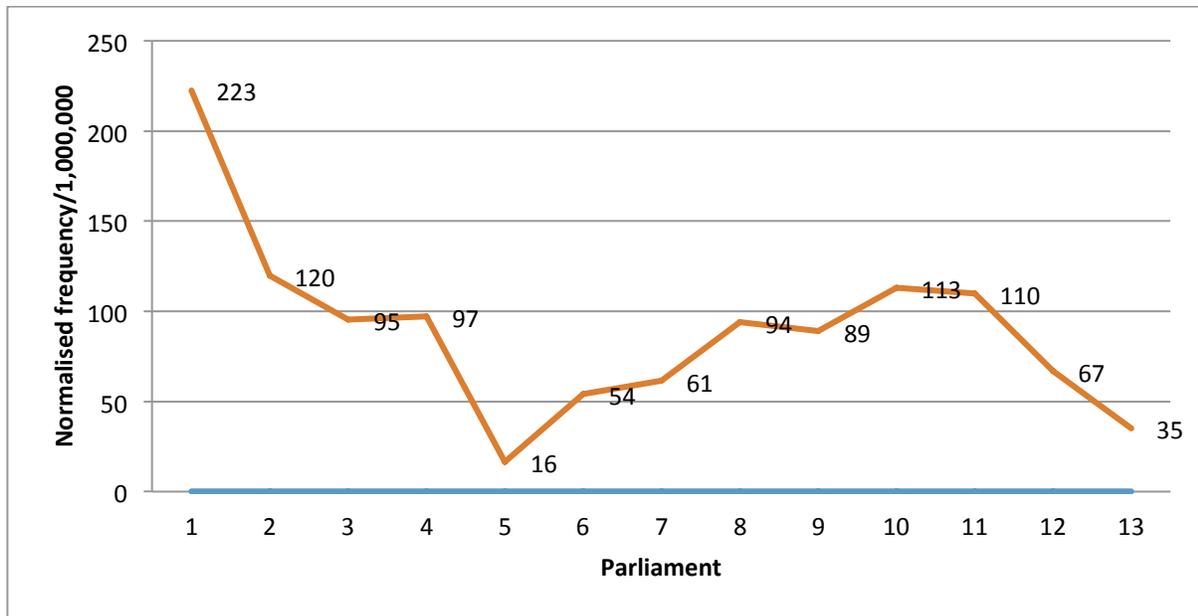


FIGURE 2. n-gram of the trend of the lexical *perempuan/woman* in MHC

Figure 2 shows that the lexical *perempuan/woman* has a downward trend beginning from parliament 1 (1959) to parliament 5 (1978). But then, an upward trend begins from P6 (1982-1986) until P10 (1999-2003) and declines once again from P11 until P13.

TABLE 3. Frequency of *perempuan/woman* (P1-P13)

Parliament	Freq	Year
1	223	1959-1964
2	120	1964-1969
3	95	1971-1973
4	97	1974-1978
5	16	1978-1981
6	54	1982-1986
7	61	1986-1990
8	94	1990-1994
9	89	1995-1999
10	113	1999-2003
11	110	2004-2007
12	67	2008-2012
13	35	2013-2018

Tables 2 and 3 show the frequency of the nodes *wanita/woman* and *perempuan/woman*. There is a gradual upward trend in the middle (P6/P7 until P10) but then show a downward trend until P13, reflecting the idea a recurring concern for women issues in Malaysia. It could represent a decrease in debates about women issues and concerns in MHC for the coming years.

This study has identified 4 parliaments for *wanita/woman* and 3 parliaments for *perempuan/woman* that represent focal point to explain certain pattern and trend that will be explained in Part 2 and Part 3. Table 4 briefly explains the reason to choose the selected parliament to be highlighted in the study.

TABLE 4. The reasons of selection of P1, P5, P10 and P13 for this study

Parliament	Reasons
1	The first discussion of <i>wanita/woman</i> and <i>perempuan/woman</i> in MHC and the highest frequencies for <i>perempuan/woman</i> .
5	The lowest frequencies of <i>wanita/woman</i> and <i>perempuan/woman</i> in MHC
10	The highest frequencies of <i>wanita/woman</i> in MHC
13	The last discussion of <i>wanita/woman</i> and <i>perempuan/woman</i> in MHC

The following are the frequencies based on collocation for P1, P5, P10 and P13 as discussed earlier. These frequencies are based on collocation and will show the relationship between the words and the nodes.

TABLE 5. P1 (The first discussion of *wanita/woman*)

#Total No. of Collocate Types: 1363		
#Total No. of Collocate Tokens: 8760		
Freq	MI Scores	Collocate
100	9.7141	<i>Kaum/group</i>
27	9.67348	<i>Laki/man</i>
12	9.6222	<i>Pemudi/young lady</i>
100	9.62055	<i>Wanita/woman</i>
10	9.20304	<i>Lelaki/man</i>
11	9.13409	<i>Lemah/weak</i>
13	7.99659	<i>Maju/advance</i>
11	7.86661	<i>Askar/soldier</i>
18	7.80246	<i>Merasa/feel</i>
23	7.50355	<i>Gaji/salary</i>
26	7.48181	<i>Polis/police</i>
14	7.21898	<i>Selalu/often</i>
16	7.12943	<i>Tangan/hand</i>
13	6.90889	<i>Latehan/training</i>
20	6.83859	<i>Berikan/to give</i>

Table 5 shows the collocates of *wanita/woman* in the first discussion on women in MHC. The most frequent collocate is *kaum/group* as in *kaum wanita/a group of women*. In the old days, the plural *wanita-wanita/women* are often referred to as *kaum wanita/a group of women*. The second most frequent collocate is *laki/man*; this reflects the binary structure of how women issues are discussed then, that is, women in relation to men.

TABLE 6. P5 (The lowest frequencies of *wanita/woman*)

#Total No. of Collocate Types: 1159		
#Total No. of Collocate Tokens: 4130		
Freq	MI Scores	Collocate
104	11.99299	<i>Wanita/woman</i>
33	11.3265	<i>Gadis/young girl</i>
19	10.39127	<i>Lelaki/man</i>
21	9.88301	<i>Pakaian/clothing</i>
11	8.09768	<i>Perlindungan/protection</i>
62	7.90469	<i>Kaum/group</i>
22	7.38398	<i>Anggota/body</i>
43	7.38367	<i>Islam/Islam</i>
42	6.83062	<i>Pekerja/worker</i>

32	6.82488	<i>Seorang/a person</i>
15	6.45794	<i>Polis/police</i>
10	6.42872	<i>Asing/alien</i>
13	6.13176	<i>Terutama/main</i>

Table 6 shows the lowest frequencies of *wanita/woman* as found in P5. This is also plotted in the n-gram that indicates the least number of times the word *wanita/woman* is mentioned. What this table also shows is that the node *wanita/woman* collocates with the word *laki/man*, reflecting the binary *wanita-woman/lelaki-man* structure in the discussion about women issues.

TABLE 7. P10 (The highest frequencies for *wanita/woman*)

#Total No. of Collocate Types: 4992		
#Total No. of Collocate Tokens: 66430		
Freq	MI Scores	Collocate
14	10.31471	<i>Bersuami/married woman</i>
11	10.24242	<i>Leiaki (typo)/man</i>
28	10.05691	<i>Ppb/An acronym for an award bestowed by the sultan or heads of states</i>
43	10.03092	<i>Mendiskriminasikan/discriminating</i>
13	9.82447	<i>Hamil/pregnant</i>
90	9.55699	<i>Gadis/young girl</i>
358	9.14623	<i>Lelaki/man</i>
187	9.08796	<i>Rajang/ a place in Sarawak (Teacher's Institute)</i>
19	9.06939	<i>Aurat/intimate body part that needs to be veiled for muslims</i>
89	8.89563	<i>Berkahwin/married</i>
16	8.84109	<i>Mengandung/pregnant</i>
1436	8.82114	<i>Wanita/woman</i>
16	8.78299	<i>Dirogol/being raped</i>
763	8.7219	<i>Keluarga/family</i>
43	8.66277	<i>Hawa/woman</i>

Table 7 shows the highest frequencies for *wanita/woman* found in MHC as indicated in P10. As shown by the other collocates, some women issues discussed are quite prevalent such as *mendiskriminasi/discriminating*, *mengandung/pregnancy*, and *dirogol/being raped*.

TABLE 8. P13 (The downward trend of frequencies of *wanita/woman*)

#Total No. of Collocate Types: 4800		
#Total No. of Collocate Tokens: 67970		
Freq	MI Scores	Collocate
120	11.03604	<i>Selangau/ a town in Sarawak</i>
20	10.89165	<i>Tuala/towel</i>
36	10.13761	<i>Hamil/pregnant</i>
18	9.84656	<i>Incubator</i>
10	9.52908	<i>Pencinta/lover</i>
51	9.47965	<i>Mengandung/pregnant</i>
10	9.45108	<i>Pampers/a brand of baby diaper</i>
14	9.42955	<i>Placement</i>
55	9.26362	<i>Pembuat/manufacturer</i>
14	9.23024	<i>Dirogol/being raped</i>
1298	9.18145	<i>Keluarga/family</i>
229	9.17325	<i>Lelaki/man</i>

		<i>Datin</i> /the wife of a dato (an award received from the head of state or the king and sultan)
59	9.1512	
80	8.98133	<i>Diskriminasi</i> /Discrimination
10	8.86612	Cedaw/acronmyn for Convention on the Elimination of All Forms of Violence against Women

Table 8 shows the downward trend of the frequencies of *wanita*/woman as indicated in P13. This downward trend may explain a new tendency to discuss women issues in a new term, that is, gender equality. This is evident with the inclusion of CEDAW, an international treaty that Malaysia is a signatory to.

The following are frequencies based on collocation for P1, P5 and P13.

TABLE 9. P1 (The first discussion and the highest frequency of *perempuan*/woman)

#Total No. of Collocate Types: 1289		
#Total No. of Collocate Tokens: 6960		
Freq	MI Scores	Collocate
39	11.49836	<i>Lelaki</i> /man
68	11.33789	<i>Laki</i> /man
21	9.45327	<i>Budak</i> /a child
18	9.14472	<i>Tangga</i> /steps
80	9.08107	<i>Murid</i> /student
11	9.04291	hostel
22	9.02705	<i>Lanjutan</i> /extension
45	8.89394	<i>Kaum</i> /group
15	8.88692	<i>Maktab</i> /institute
139	8.86855	<i>Anak</i>/child
110	8.82302	<i>Guru</i> /teacher
32	8.64038	<i>Perempuan</i> /woman
15	8.57236	<i>Terdiri</i> /consitute
38	8.55976	<i>Gaji</i> /salary
18	8.53215	<i>Penuntut</i> /student

Table 9 shows the first and the highest frequency of the node *perempuan*/woman in P1. The highest tendency to use the node *perempuan*/woman at this time is attributable to the idea that there was no clear difference between the word *perempuan*/woman and *wanita*/woman in the period immediately after the Independence in 1957. This does not mean women issues were largely ignored later. Indeed, the node *wanita*/woman becomes increasingly more popular after that. The tendency to use *wanita*/woman instead of *perempuan*/woman thereafter is due to the more positive connotation associated with the word *wanita*/woman.

TABLE 10: P5 (The lowest frequencies of *wanita*/woman)

#Total No. of Collocate Types: 936		
#Total No. of Collocate Tokens: 2920		
Freq	MI Scores	Collocate
90	13.13537	<i>Lelaki</i>/man
38	11.54082	<i>Perempuan</i> /woman
21	10.84713	<i>Laki</i> /man
13	8.25983	<i>Pesakit</i> /patient
10	7.70925	<i>Menengah</i> /secondary
29	7.47806	<i>Murid</i> /student
15	7.47224	<i>Penuntut</i> /student

34	7.41252	<i>Seorang/someone</i>
18	6.96965	Hospital
14	6.54566	<i>Pelajar/student</i>
19	6.21322	<i>Guru/teacher</i>
18	6.12982	<i>Dia/his or her</i>

Table 10 shows the lowest frequency of the node *wanita/woman* as indicated in P5. Interestingly, P5 is also the lowest frequency of *perempuan/woman* at only 16 mentions as stated in Table 3. This shows that women issues were not central in MHC during that period.

TABLE 11. P13 (Downward trend of frequencies of *wanita/woman*)

#Total No. of Collocate Types: 1426		
#Total No. of Collocate Tokens: 6350		
Freq	MI Scores	Collocate
18	14.64514	Zainab (a woman's proper name)
216	12.50901	Lelaki/man
18	12.42791	<i>Laki/man</i>
64	11.48653	<i>Perempuan/woman</i>
11	10.68672	<i>Mengandung/pregnant</i>
18	10.02749	<i>Berkahwin/married</i>
22	9.75873	<i>Remaja/teenager</i>
10	9.71393	<i>Adik/younger sibling</i>
22	9.23144	<i>Budak/child</i>
10	9.02896	<i>Kahwin/married</i>
13	8.66312	<i>Berumur/ageing</i>
27	8.65117	<i>Umur/age</i>
15	8.09486	<i>Isteri/wife</i>
78	7.80439	<i>Kanak/a child</i>
75	7.481	<i>Seorang/someone</i>

Table 11 shows the downward trend of the frequency of *wanita/woman* in P13. The most frequent collocate is *lelaki/man*; thus, showing that women issues are discussed in relation to *lelaki/man*. Despite the downward trend, the MI scores indicate that the frequency is still significant.

The frequencies of the nodes *wanita/woman* and *perempuan/woman* reflect the significance of women issues in MHC. This is the number of occurrence as empirical evidence that women issues have never been ignored, albeit their fluctuations.

PART 2: CHANGING TREND

This section presents the finding to address the second objective, that is, to identify the trend of the usage of these two nodes *wanita/woman* and *perempuan/woman*. This can be achieved by identifying the collocates and their frequency. This is in line with Hajar Abdul Rahim's assertion that "the number of the frequency of the collocations can inform about the behavior of the collocations studied" (2005, p. 96).

One crucial finding in terms of the pattern is that there exists a binary pattern that reflects the way women issues are discussed. Indeed, it is found that there is a binary pattern of collocate between *perempuan/woman* and *lelaki/man*. A binary pattern is a pattern of opposite, such as man and woman, boy and girl, masculine and feminine.

TABLE 12. Binary Pattern of *wanita-lelaki*/woman-man

Parliament	Frequency	MI Scores	Note
1	10	9.20304	Freq > 10, MI > 6
5	19	10.39127	Freq > 10, MI > 6
10	358	9.14623	Freq > 10, MI > 6
13	229	9.17325	Freq > 10, MI > 6

As shown in Table 12, P1 shows the highest collocate of *wanita-lelaki*/woman-man. Although the collocation *wanita-lelaki*/woman-man shows a decline compared to the collocation at the peak in P10, the frequency is still high. Thus, this is an indication that there are issues that need further analysis to see why the use of lexical *perempuan*/woman seems ‘less important’ compared to *wanita*/woman in parliament. One explanation is provided by Hajar Abdul Rahim’s study that shows that the lexical *wanita*/woman is often used in a more formal context while the lexical *perempuan*/woman is more frequently used “to determine gender” (2005: 102) such as *budak*/a child and *murid*/student’ (96). However, all Parliaments from P1 to P13 achieve the minimum requirement of MI scores 6 and Frequency minimum 10.

TABLE 13. The frequent collocation with *wanita*/woman for P1, P5, P10, P13

Parliament	Frequency	MI Scores	Collocate
1	100	9.7141 9.62055	<i>Kaum</i> /group <i>Wanita</i>/woman
5	104	11.99299	<i>Wanita</i>/woman
10	1436	8.82114	<i>Wanita</i>/woman
13	1298	9.18145	<i>Keluarga</i>/family

Based on Table 9, it is found that the same pattern where *wanita*/woman is frequently discussed on its own term, that is, not in relation to others like *wanita-woman/lelaki*-man or *wanita-woman/anak*-a child occurs in P1, P5 and P10 (for details refer to Table 5, 6, 7 and 8 in Part 1). However, P13 shows that the pattern changes in which *keluarga*/ family appeared more frequently in *wanita*/woman discussion in MHC. Thus, this shows that the role of *wanita*/woman is linked to *keluarga*/family and this becomes increasingly important as it shows a shift in the conception of a woman from a person (biological category) towards a woman in relation to her cultural roles (gender category). The concordance of the collocation *wanita-keluarga*/woman-family thus was analysed contextually. Detailed discussion is in Part 3 (refer file .txt concordance P13)

Similar to *wanita*/woman, there also exists a binary pattern of collocate between *perempuan*/woman and *lelaki*/man. However, based on MI scores, collocation of *perempuan*/woman with *lelaki*/men are more significant compared to *wanita*/woman for P1, P5, and P13 except P10 where there is only a discussion under the lexical *wanita*/woman.

TABLE 14. Shows the binary pattern *wanita*/woman-*lelaki*/man

Parliament	Frequency	MI Scores	Note
1	39	11.49836	Freq > 10, MI > 6
5	90	13.13537	Freq > 10, MI > 6
13	216	12.50901	Freq > 10, MI > 6

TABLE 15. Comparison of normalized frequency between *wanita/woman* and *perempuan/woman* (P1-P13)

Parliament	Freq (<i>Wanita/woman</i>)	Freq (<i>Perempuan/woman</i>)	Year
1	286	223	1959-1964
2	105	120	1964-1969
3	151	95	1971-1973
4	178	97	1974-1978
5	37	16	1978-1981
6	212	54	1982-1986
7	171	61	1986-1990
8	238	94	1990-1994
9	497	89	1999-1995
10	953	113	1999-2003
11	643	110	2004-2007
12	694	67	2008-2012
13	376	35	2013-2018

Based on Table 15, the collocation for *perempuan-lelaki/woman-man* shows an upward trend beginning from P7 and declining after reaching the second peak in P10. Overall the frequency is still below P1 (First peak) when the comparison was made throughout P1 until P13 using normalized frequency per 1 million lexical items (see Figure 2 in Part 1). Instead, the collocation for *wanita/woman* has a different trend in which the frequency stated is far greater than *perempuan/woman* (see Figure 1 Part 1). This again reaffirms the finding of Hajar Abdul Rahim's study (2002) that reveals the use of *wanita/woman* in a more formal setting. However, all the Parliaments (P1-P13) achieve the minimum requirement of MI scores 6 and the minimum frequency of 10.

TABLE 16. The frequent collocation with *perempuan/woman* for P1, P5 and P13

Parliament	Frequency	MI Scores	Collocate	Note
1	39	11.49836	anak/child	Freq > 10, MI > 6
5	90	13.13537	lelaki/man	Freq > 10, MI > 6
13	216	12.50901	lelaki/man	Freq > 10, MI > 6

P5 and P13 show the same binary pattern in which *perempuan/woman* is frequently discussed with *lelaki/man*. However, P1 shows that 'anak/child' appeared more frequently in *perempuan/woman* discussion or debate in parliament (for the details refer Table 9, 10, 11 in Part 1). Thus, this shows that the issues related to the lexical *perempuan/woman* have changed its gender grammar shifting the focus from a woman as a daughter *anak/child* as one family unit to discussions solely on woman as a woman. The concordance of the collocation *perempuan-lelaki/woman-man* was then analysed contextually to identify the gender related issues. Detailed discussion is in part 3.

PART 3: ISSUES RELATED TO WOMEN

As discussed in the Part 2, both *wanita/woman* and *perempuan/woman* have different trend of occurrences. It is found that the two lexicals consistently collocate with *lelaki/man*; this strengthens the argument that the discussion of gender in MHC follows a dichotomy of gender or binary pattern, which is *wanita/woman-lelaki/man* and *perempuan/woman-lelaki/man*. Part 3 looks at some issues related to women. There are two sub-categories: 1) Issues of *wanita/woman* based on concordance and contextual analysis of the most frequent collocate in P1, P5, P10 and

P13; 2) Issues of *perempuan*/woman based on concordance and contextual analysis of the most frequent collocate in P1, P5 and P13.

1) ISSUES OF *WANITA*/WOMAN BASED ON CONCORDANCE AND CONTEXTUAL ANALYSIS OF THE MOST FREQUENT COLLOCATE IN P1, P5, P10 AND P13.

This section provided the analysis of issues of *wanita*/woman based on the concordance to identify some contextual information for discourse analysis to be carried out. The most frequent collocates as generated in P1, P5, P10 and P13 are presented in the table below:

TABLE 17. Collocate of *wanita*/women based on the highest frequencies of P1, P5, P10 and P13

Parliament	Frequency	MI Scores	Collocate
1	100	9.62055	<i>Wanita</i> /woman
5	104	11.99299	<i>Wanita</i> /woman
10	1436	8.82114	<i>Wanita</i> /woman
13	1298	9.18145	<i>Keluarga</i> /family

**Kaum*/group is taken out from the table since it only represents a group of women.

For this part, there are four women issues identified and discussed in the section that follows:

ISSUE 1: RURAL WOMEN DEVELOPMENT

The first issue identified in P1 with the collocate *wanita-wanita*/woman-woman is concerning the development of rural women.

ara tadi. Jadi Kementerian ini hendak-lah memakai *wanita2 dan saya shorkan wanita2* ini di-ambil daripada memegang jari wanita itu. Saya tahu ia-itu *wanita dari Dungun dan juga wanita* dari sa-belah Kerajaan pada ini hari ia-lah terdiri daripada *wanita2 daripada guru2 wanita* di-dalam UMNO. Pada masa Typist\x94 di-pejabat2, mereka lebih sukakan kaum *wanita daripada kaum laki2 kerana wanita* itu dengan chara Sidek: Terima kaseh. Jadi saya minta supaya ahli *wanita Dewan ini dan juga wanita* yang lain di-usaha RIDA ini ia-lah mengadakan Taman Asohan *Wanita. Di-dalam Taman Asohan Wanita* ini, Tuan Yang 2 dengan pehak wanita Indonesia maka orang itu\x97 *wanita2 di-Indonesia attack wanita2* di-sini (Ketawa) baharu- Pada hari ini beliau mengatakan lebih2 bagus bagi *wanita2 di-Pantai Timor kerana wanita* di-Pantai Timor ini saya pandang ada-lah satu penghinaan kepada *wanita2 diPantai Timor. Bukan wanita2* di-Pantai Timor sahaja

The context of the issue is identified as the following:

21 DECEMBER 1960

[...] *Dalam soal pegawai2 ini saya tidak tahu sama ada pegawai2 ini ada terdiri dari wanita atau tidak, sa-kira-nya belum saya berharap kepada Kementerian yang berkenaan supaya mengadakan Pegawai2 Wanita dalam soal ini, kerana kemajuan di-luar bandar itu ada-lah sa-bahagian besar-nya di-pegang oleh tanggungan wanita2 di-kampong. Maka sa-kira-nya pegawai ini atau penolong pegawai ini terdiri daripada wanita tentu-lah senang dan mudah-lah bagi wanita itu memberi nasehat dan mengadakan perundingan atau memberi bermacam2 pandangan kepada wanita2 kampong.*

With regards to these officers, I am not sure if they are woman officers or not. If they are not, I hope the Ministry concern is able create a post for woman officers as the development of rural

areas depends on these rural women. If there are woman officers, it would be easier for these officers to advise and provide consultation or give all sorts of ideas to rural women.

In this discourse, the role of *wanita*/woman is seen as positive. The speaker identifies the need of creating formal employment opportunities for women in rural areas who in turn can contribute towards the development of rural communities. The idea that women can engage with women better may sound outdated now, but it is the reality during that time. This positive relationship is the reason why *wanita*/woman collocates with *wanita*/woman.

ISSUE 2: WOMEN SECURITY

The second issue identified in P5 with collocate *wanita-wanita*/woman-woman is regarding woman and security in the context of working in the private agencies.

ipu oleh agensi-agensi pekerjaan swasta tersebut. *Wanita-wanita ini dikatakan* telah terkandas apabila mereka ditas oleh kerana ianya tidak sejajar dengan pakaian *wanita-wanita Islam*, dan jika ia, sama ada Kerajaan saya yakin dan percaya apa yang diharapkan oleh *wanita-wanita Islam* khususnya dan umat Islam di negara innya yang mana hukuman hanyalah dikenakan keatas *wanita-wanita Islam* manakala pasangan yang terdiri dari lelaki kenaaan tadi bahawa tidak berlaku tegahan terhadap *wanita-wanita Islam* menutup aurat tetapi ini ada satu h Kerajaan akan berusaha supaya jangan menghalang *wanita-wanita Islam* menutup auratnya seperti mana yang dikehenc supaya ditambah clause tertentu, khususnya kepada *wanita-wanita Islam* selain daripada kebenaran Menteri atau Ketu yang ada di negara ini. Apa yang menyedihkan *wanita-wanita Islam* selama ini tidak berasa berpuas hati eh badan-badan dakwah misalnya untuk menggalakkan *wanita-wanita Islam* supaya berpakaian secara Islam yang menutup bahawa usaha-usaha untuk menggalakkan pengambilan *wanita-wanita Islam* supaya mereka memakai cara yang dibenarkar ran kecil nanti mengekori peraturan ini khas bagi *wanita-wanita Islam*. Tuan Yang di-Pertua: Adakah Yang

The context of the issue is identified as the following:

9 DISEMBER 1980

Di dalam kes Sabah dan Sarawak, Kementerian ini telah menerima beberapa aduan mengenai wanita muda yang telah tertipu oleh agensi-agensi pekerjaan swasta tersebut. Wanita-wanita ini dikatakan telah terkandas apabila mereka ditawarkan bekerja di negeri-negeri jiran dengan tidak diberi kemudahan untuk babk dan juga pekerjaan-pekerjaan yang diberikan itu tidak seperti yang telah dijanjikan. Ada kemungkinan bahawa sebilangan besar wanita-wanita tersebut telah dipaksa bekerja di tempat-tempat yang di-gunakan untuk tujuan lucuh. Dengan adanya undang-undang yang dicadangkan ini maka saya harap agensi-agensi pekerjaan swasta dapat diperiksa dan kes-kes penipuan ini dapat dikurangkan.

In the case of Sabah and Sarawak, the Ministry has received several complains regarding cases of women being cheated by private employment agencies. These women are stranded in neighbouring countries when the jobs they are being offered are not the same as what they have been promised. There is a probability that some of these women were forced in the adult industry. With the new proposed law, it is hoped that these private employment agencies are monitored so that these cases can be eradicated

The speaker in this session argues that her Ministry has been proactive in dealing with women trapped in the employment scam. What is important is that the collocation *wanita-wanita*/woman-woman' actually refers to the plural women; this highlights the large number of women involved in such cases as victims. In this case, *wanita*/woman is depicted in a negative setting in which they can be sexually exploited.

ISSUE 3: WOMEN EMPOWERMENT

The third issue identified in P10 with collocate *wanita-wanita/woman-woman* is concerning the establishment of the Ministry of Women, Family and Community Development.

i adalah kesungguhan kerajaan untuk membela nasib *wanita sehingga ada kementerian wanita*, tetapi bagaimana pula keadaan ah dengan keikhlasan untuk meningkatkan kedudukan *wanita. Sekiranya kerajaan menganggap wanita* adalah sama rata dengan

The context of the issue is identified as the following:

1.8.2001

Sebenarnya ini adalah kesungguhan kerajaan untuk membela nasib wanita sehingga ada kementerian wanita, tetapi bagaimana pula keadaan di Kelantan? Berapakah peruntukan Kerajaan Negeri Kelantan untuk membela kaum wanita? Dan juga di negeri Terengganu. Di negeri-negeri lain misalannya di Sabah kita ada Pengarah Hal Ehwal Wanita yang memastikan aktiviti wanita.

Actually, this is a proof that the government with the establishment of the Women Ministry tries to uphold women's life and integrity. What about in Kelantan? What is the allocation provided by the state government of Kelantan to protect women? What about the state of Terengganu? In other states like Sabah, there is the director of women's welfare who ensures that there are activities for women.

The speaker in this parliament session tries to establish the usefulness of the Ministry of Women in empowering women. In this case, the issue raised is whether the states held by the opposition parties (Kelantan and Terengganu) provide the same services for women they way the federal government is doing. The collocation points to the relationship between *wanita/woman* with the Women Ministry.

ISSUE 4: DISCRIMINATION AGAINST MARRIED WOMEN

The fourth issue identified in P13 with collocate *wanita-keluarga/woman-family* is concerning the discrimination against women who are married.

n 010000 iaitu di bawah Butiran Dasar Pembangunan *Wanita, Keluarga dan Masyarakat*, serta Pelaksanaan Dasar Sosial Ne rjaan 1955 yang saya kira Kementerian Pembangunan *Wanita, Keluarga dan Masyarakat* harus mengambil serius bagi mendes Jawatankuasa Butiran B.48 Kementerian Pembangunan *Wanita, Keluarga dan Masyarakat*. Tuan Pengerusi, saya ingin

Saya mengakui bahawa Kementerian Pembangunan *Wanita, Keluarga dan Masyarakat* melalui Jabatan Kebajikan Masyarakat ERTI yang dijalankan oleh Kementerian Pembangunan *Wanita, Keluarga dan Masyarakat* yang khususnya memilih beberapa bu soalan kepada Yang Berhormat Menteri Pembangunan *Wanita, Keluarga dan Masyarakat*. Bagaimanakah kita hendak mengatas n yang boleh diambil oleh Kementerian Pembangunan *Wanita, Keluarga dan Masyarakat* untuk membantu golongan wanita ini i hendak tanya, bagaimana Kementerian Pembangunan *Wanita, Keluarga dan Masyarakat* akan membantu babies yang cukup ke nyalah, apa itu rancangan Kementerian Pembangunan *Wanita, Keluarga dan Masyarakat* untuk membantu golongan ini?

The context of the issue is identified as the following:

21.11.2016

[...] saya ingin bertanya kepada kementerian tentang tindakan yang dibuat supaya tidak berlaku diskriminasi kepada kaum wanita di tempat kerja mahupun ketika mereka mengandung kerana banyak kes yang menunjukkan bahawa ibu-ibu yang melahirkan anak mungkin ibu-ibu yang mengandung ada kes yang mereka dibuang daripada tempat kerja dia dan sebagainya. Jadi saya harap based on kita punya persetujuan dalam hak membela dan menindas wanita ini,

diskriminasi terhadap wanita ini diberi perhatian juga kepada peruntukan-peruntukan di tempat kerja dan sebagainya supaya tidak berlaku diskriminasi.

[...] I would like to ask the ministry regarding the action taken so that there is no discrimination against women at workplace especially when these women are pregnant. There have been cases of women who have given birth or pregnant been given a sack. So, I hope, based on this agreement on the issue of discrimination against women, attention is also paid to workplace discrimination etc. so that no more discrimination can take place

The speaker, in this case, enquires about the actions taken by the Ministry (of Women) to ensure that women who are with babies or children are not discriminated against in workplace. Even though the issue discussed here is negative, the tone is actually positive as it refers to the need to eradicate discrimination against women at workplace. Again, this *wanita-keluarga/woman-family* collocation refers to the relationship between women and the Ministry of Women, Family and Community Development. This suggests the usefulness of the Ministry in handling women issues.

2) ISSUES OF PEREMPUAN/WOMAN BASED ON CONCORDANCE AND CONTEXTUAL ANALYSIS OF THE MOST FREQUENT COLLOCATE IN P1, P5 AND P13.

This section provided the analysis of issues of *perempuan/woman* based on the concordance to identify some contextual information for discourse analysis to be carried out. The most frequent collocates as generated in P1, P5 and P13 are presented in the table below:

TABLE 18. Collocate of perempuan/women based on the highest frequencies of P1, P5, and P13

Parliament	Frequency	MI Scores	Collocate
1	139	8.86855	Anak/child
5	90	13.13537	Lelaki/man
13	216	12.50901	Lelaki/man

ISSUE 1: THE QUALITY OF TRAINING AND EDUCATION OF ANAK PEREMPUAN/DAUGHTER'

The first issue identified in P1 with collocate *perempuan-anak/woman-child* is concerning the the quality of the training and education of daughters.

tangga 5021 6 FEBRUARY 1961 5022 ia-itu di-mana anak2 perempuan kita dapat di-beri latehan dan pendidekan ini ia-lah bagi mem-beri peluang kepada anak2 perempuan kita daripada sekolah kebangsaan yang di-kampung tidak berapa mengambil berat tentang hal kemajuan anak2 perempuan kita di-Indonesia. Beliau pergi hanya untuk mem-beri pendidekan yang sa-luas2-nya kepada anak2 perempuan kita di-dalam negeri ini. Saya merasa ini supaya membanyak-kan sekolah bagi pendidekan anak perempuan kita di-Tanah Melayu ini. Enche\x92nya datang-lah register maka terpaksa-lah tangan anak perempuan kita di-sana di-pegang oleh Postman, ada satu pusat latehan di-mana wanita2 atau anak2 perempuan kita di-lateh dan di-didek sa-

The contextual evidence is as follows:

6 FEBRUARY 1961

[...] Dengan sa-masak2 pendidekan atau pelajaran yang di-berikan kepada anak2 perempuan kita ini maka dapat-lah kita menjadikan anak2 itu sa-bagai pelateh di-kampung2 sa-bagaimana dasar Kerajaan pada hari ini akan mengadakan pendidekan yang sa-baik2-nya bagi penduduk2 di-luar bandar baik bagi anak2 laki2 mahu pun anak2 perempuan kita.

[...] Oleh sebab itu saya rasa, kalau pehak Kerajaan lebeh2 lagi pehak Kementerian Pelajaran supaya mengambil perhatian, kalau dapat saya minta supaya gedong yang endah dan molek itu saperti yang ada di-dalam bandar Kuala Lumpur ini di-mana kita memberi latehan kepada anak2 kita untok di-jadikan pemimpin atau pelateh bagi wanita2 di-luar bandar supaya di-pindahkan dasar-nya.

[...] with sufficient and quality education given to our daughters, they can become trainers in rural areas. This is in line with the present government's policy to provide quality education to rural folks regardless of whether they are boys or girls.

[...] Therefore, I feel that the government especially the Ministry of Education to pay attention to this issue by constructing state of the art educational facilities similar to the ones that can be found in Kuala Lumpur. This facility will help training youngsters to train women in rural areas.

The speaker hopes that the government pays more attention to the quality of the education for *perempuan-anak*/woman-daughter. In this context, the word *perempuan*/woman is used as a gender marker. This point is also raised by Hajar Abdul Rahim (2005), where she points out that the word *perempuan*/woman in this case is neutral, but can be used pejoratively. However, in this case, it is used in a neutral context.

ISSUE 2: HUKUMAN JENAYAH SYARIAH

The second issue identified in P5 with collocate *perempuan-lelaki*/woman-man is concerning the punishment under the Syariah (Islamic) law.

sendirinya orang itu ada hubungan seks di antara lelaki dengan perempuan itu. Dato\92 Hisbamuiddin bin Haji hormat dari Semerah itu yang mahu pelajar-pelajar lelaki diasingkan daripada pelajar-pelajar perempuan di sekolah-sekolah boleh menjadi guru mengajar agama itu, kalau lelaki disebut ustaz, kalau perempuan disebut ustazah. Kalau sudah tadi dan yang menyeluk sehingga dalam rahim orang perempuan ini tadi adalah doktor lelaki dan dengan sebab n oleh Yang Berhormat kes mengenai khalwat antara perempuan Islam dengan lelaki yang bukan Islam, ini telah lelakinya bukan Islam; contoh-nya kita kata yang perempuan Islam Melayu dan yang lelaki bukan Islam\97

The contextual evidence is found as the following:

3 DISEMBER 1980

Secara ringkas bolehlah dinyatakan bahawa segala jenis perhubungan seks di antara perempuan dan lelaki yang bukan suami isteri dari khalwat hingga persetubuhan haram adalah tertakluk di bawah bidangkuasa Mahkamah Syariah. Saya ingin bertanya kepada Yang Berhormat Menteri, apakah cara-cara yang akan dibuat oleh pihak Kerajaan secara tertentu mengenai masalah khalwat ini, jika yang berkhalwat itu yang perempuannya Islam, yang lelakinya bukan Islam; contoh-nya kita kata yang perempuan Islam Melayu dan yang lelaki bukan Islam orang India, bagaimana hukumannya?

In short, it can be said that all forms of sexual relationship between a man and woman who are not married, from close promiximity to out of wedlock sexual intercourse, are under the jurisdiction of the Syariah Court. I would like to ask the Minister, what are the actions taken by the government with regards to the case of close proximity that involves Muslim women and non-muslim men. For example, between a Malay Muslim woman and a non-Muslim Indian man. What is the punishment?

The speaker in this context raises the issue of inter-racial/religious relationship involving Muslim *perempuan*/woman and non-muslim man. In this context, it can be seen that the word *perempuan*/woman is used as the issue discussed is a negative one. The speaker believes that the

relevant agency is not carrying out its duty well; hence, the word *perempuan*/woman is used negatively.

ISSUE 3: LACK OF PROGRESSIVE POLICY

The third issue identified in P13 with collocate *perempuan-lelaki*/woman-man is concerning the lack of progressive policy on women.

an, menjatuhkan seseorang dan janganlah satu puak lelaki dan perempuan juga menjatuhkan kalangan wanita-wanita mereka' dan
ak timbul kerana lelaki diurus oleh pihak pegawai lelaki dan perempuan juga diasingkan oleh pengawal pihak perempuan. In
uga tidak ada pilihan, kalau ada kedua-dua doktor lelaki dan perempuan, kita seboleh- bolehnya akan menunaikan perminta
uju ke negara maju, mesti ada keseimbangan antara lelaki dan perempuan. Kita tidak mahu bila orang perempuan ke
asar yang meliputi perbezaan jantina, kanak-kanak lelaki dan perempuan, kumpulan umur, pendidikan, pendapatan, bandar da
ntang kes keganasan rumah tangga. Statistik 2015, lelaki dan perempuan, mangsa lelaki ialah 23 kes, mangsa ada juga

The textual evidence of the issue is as the following:

16.3.2017

Akan tetapi malangnya di Malaysia kita lihat pembudayaan wanita masih lagi ketinggalan dengan tidak adanya dasar-dasar baru yang lebih progresif yang boleh meletakkan wanita ada kesaksamaan mereka dengan rakyat atau dengan orang lelaki di negara ini.

Jadi orang lelaki juga itu bertanggungjawab untuk memperdayakan wanita, untuk memartabatkan nasib wanita. Jadi kalau discrimination to women, maknanya juga adalah discrimination to the nations.

It is rather unfortunate that in Malaysia the empowerment of women is very much behind due to the lack of new policies that are progressive and able to promote equality, either among the people or the men in this country.

Therefore, men are also responsible in empowering women, to bring dignity to their lives. This simply means discrimination against women equals discrimination against the nation)

The speaker in this context expresses her frustration at the speed in which women empowerment is promoted in Malaysia. The lack of more progressive policies for women development creates a more negative tone to her discourse. For that reason, the speaker uses the word *perempuan*/woman as opposed to *wanita*/woman to express her disappointment. *Perempuan*/woman collocates with *lelaki*/man as the speaker believes that gender equality needs support from men.

CONCLUSION

It can be concluded that both nodes *wanita*/woman and *perempuan*/woman point to not only how issues related to women in MHC are discussed but also the cultural connotations of both nodes. In addition, culturomic as an approach allows for cultural trend to be proven empirically with the generation of n-grams. As a corpus-driven study, one interesting pattern that emerges is the gender binary system or dichotomy that structures how women issues are discussed, that is, in relation to men. *Lelaki*/man collocates consistently with both nodes especially in the earlier parliament discourses, hinting at how society views men as the standard to aspire to. Besides that, there is also a shift in trend in terms of how *wanita*/woman and *perempuan*/woman have been discussed, that is, from women functioning in private sphere (collocates like *anak*/child and *mengandung*/pregnant) to a more public one (collocates like *guru*/teacher and *gaji*/salary). However, there is no apparent shift in terms of women being defined as a biological category

(sex) or a cultural category (gender); this points at the dearth of consciousness about gender as the new technocratic or operational term in the nation's development program.

The present study has shown via n-grams the cultural trend of the nodes *wanita*/woman and *perempuan*/woman in MHC. The frequency of both lexicals has also shown a decrease in usage; nonetheless, the n-grams show that the lexical *perempuan*/woman is getting less frequently used compared to the lexical *wanita*/woman'. This is due to the association the node *perempuan*/woman has with informality and *wanita*/woman with a more formal discourse setting like the parliamentary discourse. This finding seems to support the study done by Hajar Abdul Rahim (2005) on newspaper corpus, arguing that the positive semantic prosody of the lexical *wanita*/woman sets it quite apart from its *perempuan*/woman counterpart. One common similarity between this study and that of Hajar's is the formal context in which these nodes appear, that is, parliamentary and newspaper discourses. Unlike the newspaper discourse, nonetheless, the different tones such as skepticism, criticism and 'playfulness' used by parliamentarians in parliamentary discourses make it possible for the word *perempuan*/woman to continue to be found in MHC. This study has therefore contributed to the current interest in researching on Hansard data. With this research, more knowledge about MHC is revealed, especially with regards to women issues. For future research, it is important for a more critical scrutiny of the key words used to be undertaken so that MHC can be utilized to unpack the ideology in Hansard data.

ACKNOWLEDGMENT

This research is supported by research grant coded KRA-2018-05, Universiti Kebangsaan Malaysia.

REFERENCES

- Anonymous. (2012). *Women and Politics: Malaysia*. Putrajaya: Ministry of Women, Family and Community Development.
- Anonymous. (2018). Malaysia improves ranking in global gender gap report. Retrieved September 18, 2019 from <https://www.nst.com.my/news/nation/2018/11/431681/malaysia-improves-ranking-global-gender-gap-report>
- Anonymous. (2019). Perils of gender and geography hamper global development. Retrieved September 18, 2019 from <https://www.channelnewsasia.com/news/world/gates-foundation-goalkeepers-report-sustainable-development-goal-11912694>
- Azianura H. S. et al. (2017). Tracking trends: the corpus analysis of lexical item *jenayah* in the Malaysian Hansard. *Prosiding Kolokium Penerokaan Dataraya Kulturomik*. 40-45.
- Bouferrouk, A. & Dendane, Z. (2018). Muslim women issues in the press: a corpus linguistic analysis of the burka's representation in the Daily Telegraph. *European Journal of Language Studies*. 5(1). <http://www.idpublications.org/ejls-vol-5-no-1-2018/>
- D. Uppsala University. (2011). *Normalizing Results*. Introduction to Academic Writing English A1/HS1?T1, Fall 2011.
- DeFancisco, V. (1997). Gender, power and practice: or, putting your money (and your research) where your mouth is. In Ruth Wodak (Ed.), *Gender and Discourse*. Sage Publications Ltd.: London

- Glasinki, F. et al. (2016). "He Said She Said" – a Male/Female Corpus of Polish. *Proceedings of the Tenth International conference on language resources and evaluation (LREC'16)*. Retrieved October 18, 2019 from <https://www.aclweb.org/anthology/L16-1648.pdf>
- Hajar Abdul Rahim. (2005). Impak konotasi budaya terhadap leksis: satu kajian semantik berasaskan korpus ke atas perkataan perempuan dan wanita. *Jurnal Budaya*. 5(1), 83-111.
- Hajar Abdul Rahim. (2014). Corpora in language research in Malaysia. *Kajian Malaysia*. 13(Supp. 1), 1-16.
- Hirshman, C. (2016). Gender, the status of women, and family structure in Malaysia. *Malaysian Journal of Economic Studies*. 53(1), 33-50.
- Imran Ho Abdullah et al. (2017). The Malaysian Hansard Corpus. Unpublished doctoral dissertation, Culturomics Workshop UKM. 27 April 2017.
- Ilie, C. (2015). *Parliamentary Discourse. The International Encyclopedia of Language and Social Interaction*. 1st Edn. John Wiley & Sons Inc. Retrieved from <https://doi.org/10.1002/9781118611463.wbielsi201>
- Kesumawati A. Bakar. (2014). ATTITUDE and identity categorizations: a corpus-based study of gender representation. *Procedia: Social and Behavioral Sciences*. 112, 747-756
- Kim Hua Tan, et al. (2017). Trend penggunaan Bahasa samar dalam parlimen Malaysia. *GEMA Online® Journal of Language Studies*. 17(4), 84-100.
- Lih Yi, B. (2018). Women sidelined in new Malaysian government, despite campaign promises. Retrieved September 18, 2019 from <https://www.reuters.com/article/us-malaysia-women-politics/women-sidelined-in-new-malaysian-government-despite-campaign-promises-idUSKCN1I11U>
- Loffler, M. (2019). Neutral masculinity: An analysis of parliamentary debates on Austria's neutrality law. *Men and Masculinities*. 22(3), 444-464.
- Michel, Jean-Baptiste, et al. (2011). Quantitative analysis of culture using millions of digitized books. *Science*. 331(6014), 176-182. doi:10.1126/science.119964
- Minderjeet Kaur. (2017). Did EC insult Parliament, saying Hansard can't be believed, ask MPs. Retrieved September 15, 2019 from *FreeMalaysiaToday.com*
- Noorli Khamis & Imran-Ho Abdullah. (2018). What do different word lists reveal about the lexical features of a specialized language? *3L: The Southeast Asian Journal of English Language Studies*. 24(3), 26-42
- Norsimah Mat Awal et al. (2019). "Is Selangor in Deep Water ?": A Corpus-driven Account of Air / water in the Malaysian Hansard Corpus (MHC). *Journal of Language Studies*. 19 (2), 12-99.
- Rheault, L. et al. (2016). Measuring emotion in parliamentary debates with automated textual analysis. *PLoS ONE*. 1-8
- Sahul Hamid Mohamed Maiddin. (2017). Sejarah penubuhan dan perkembangan agensi propaganda: pengalaman Malaysia sejak darurat sehingga konfrontasi Malaysia-Indonesia (1948-1965). *Jebat: Malaysian Journal of History, Politics & Strategic Studies*. 44(2), 137-158.
- W. Kate. (2019). 7 surprising and outrageous stats about gender inequality. Retrieved September 27, 2019 from <https://www.weforum.org/agenda/2019/03/surprising-stats-about-gender-inequality/>

APPENDIX A

TABLES OF COLLOCATION OF WANITA/WOMAN'

Wanita Parliament 1			Wanita Parliament 2			Wanita Parliament 3			Wanita Parliament 4			Wanita Parliament 5		
Freq	MI Scores	Collocate	Freq	MI Scores	Collocate	Freq	MI Scores	Collocate	Freq	MI Scores	Collocate	Freq	MI Scores	Collocate
100	9.7141	kaum	14	12.70318	pandu	12	12.69126	juri	10	12.07817	bersuami	104	11.99299	wanita
27	9.67348	laki	14	11.21489	pengakap	61	12.60165	gadisa	292	11.71376	wanita	33	11.3265	gadisa
12	9.6222	pemudi	189	11.0648	kaum	104	11.66626	wanita	84	11.46334	gadisa	19	10.39127	lelaki
100	9.62055	wanita	13	11.03827	gadisa	21	10.67265	perlindungan	29	11.07366	berkahwin	21	9.88301	pakaian
10	9.20304	lelaki	25	10.81479	taman	10	9.94514	lelaki	41	10.2513	lelaki	11	8.09768	perlindungan
11	9.13409	lemah	16	9.79429	laki	76	9.16261	kaum	31	9.25698	perlindungan	62	7.90469	kaum
13	7.99659	maju	11	9.49176	akhlak	20	7.52462	islam	12	9.03909	bersalin	22	7.38398	anggota
11	7.86661	askar	20	9.23539	wanita	10	6.78568	latihan	10	8.93027	cantik	43	7.38367	islam
18	7.80246	merasaa	13	8.74105	tuntutan	16	6.73205	gaji	10	8.92289	memakai	42	6.83062	pekerja
23	7.50355	gaji	11	8.22827	perempuan	10	6.5646	seorang	15	8.16257	pakaian	32	6.82488	seorang
26	7.48181	polis	39	7.99635	latehan	27	6.46536	pekerja	11	8.12436	akhlak	15	6.45794	polis
14	7.21898	selalu	10	7.7082	pertubuhan				20	8.09674	pemulihan	10	6.42872	asing
16	7.12943	tangan	28	7.60808	gaji				88	7.98075	kaum	13	6.13176	terutama
13	6.90889	latehan	30	7.5569	polis				71	7.51328	polis			
20	6.83859	berikan	17	7.55237	ibu				35	7.41274	bekerja			

Wanita Parliament 6			Wanita Parliament 7			Wanita Parliament 8			Wanita Parliament 9			Wanita Parliament 10		
Freq	MI Scores	Collocate	Freq	MI Scores	Collocate	Freq	MI Scores	Collocate	Freq	MI Scores	Collocate	Freq	MI Scores	Collocate
236	11.03696	wanita	14	13.64228	bersuami	12	11.44064	penghibur	28	11.07926	mendiskriminasi	14	10.31471	bersuami
19	10.97466	gadisa	31	12.61919	gusti	312	10.32399	wanita	25	10.32061	beijing	11	10.24242	lelaki
13	10.71668	telefonis	105	11.87674	gadisa	109	10.21038	lelaki	263	10.16324	lelaki	28	10.05691	ppb
57	10.60241	lelaki	10	11.30885	laki	18	9.44896	hawa	10	9.79087	menghina	43	10.03092	mendiskriminasi
30	10.17249	bersalin	238	11.13131	wanita	48	9.15142	gadisa	10	9.56346	hoki	13	9.82447	hamil
20	9.85602	memaikai	86	10.89312	lelaki	11	8.46573	memaikai	11	9.45304	aurat	90	9.55699	gadisa
18	9.80589	berkahwin	12	10.0836	suntikan	182	8.44997	kaum	658	9.40708	wanita	358	9.14623	lelaki
10	9.68609	dekad	13	9.78047	berkahwin	14	8.38659	pelacuran	49	9.13287	wanita	187	9.08796	rajang
11	9.48717	status	12	9.68653	bersalin	13	8.11489	keganasan	18	9.11378	mengandung	19	9.06939	aurat
16	8.44663	pakaian	11	9.46653	urusetia	25	7.90014	isteri	53	9.07869	gadisa	89	8.89563	berkahwin
11	8.19557	pemudi	11	9.41661	kahwin	11	7.89232	suami	42	9.04389	berkahwin	16	8.84109	mengandung
114	7.72542	kaum	10	9.05251	diskriminasi	28	7.71366	penglibatan	41	8.94565	hawa	1436	8.82114	wanita
24	7.65813	ehwalbekerja	66	8.85362	perlindungan	68	7.63936	pertubuhan	17	8.92528	bersalin	16	8.78299	dirogol
38	7.65743	cuti	16	8.57565	rogol	12	7.57411	status	405	8.78567	kaum	763	8.7219	keluarga
11	7.43336	cuti	26	8.52413	ehwal	61	7.54637	bekerja	22	8.62772	sedunia	43	8.66277	hawa

Wanita Parliament 11			Wanita Parliament 12			Wanita Parliament 13		
Freq	MI Scores	Collocate	Freq	MI Scores	Collocate	Freq	MI Scores	Collocate
213	10.61986	jelebu	10	12.45268	hedonisme	120	11.03604	selangau
38	10.343	mendiskriminasikan	10	10.92217	asimilasi	20	10.89165	tuala
1723	9.77193	keluarga	29	10.89187	resam	36	10.13761	hamil
12	9.59565	ceiling	461	10.86044	seni	18	9.84656	inkubator
362	9.30907	lelaki	16	10.84125	ketimuran	10	9.52908	pencinta
43	9.12819	pembuat	25	10.81525	merakyatkan	51	9.47965	mengandung
230	8.76294	datin	95	10.80984	istana	10	9.45108	pampers
17	8.75022	dirogol	10	10.71572	tatasusila	14	9.42955	placement
29	8.73564	gadis	14	10.10803	cultural	55	9.26362	pembuat
20	8.52526	kesaksamaan	10	9.89469	kejutan	14	9.23024	dirogol
1689	8.52141	masyarakat	32	9.88283	culture	1298	9.18145	keluarga
758	8.51212	wanita	10	9.86772	sosio	229	9.17325	lelaki
54	8.30614	berkahwin	594	9.82151	budaya	59	9.1512	datin
165	8.2089	chew	11	9.78283	lepak	80	8.98133	diskriminasi
1867	8.04985	pembangunan	16	9.74609	teater	10	8.86612	cedaw

APPENDIX B

TABLE OF COLLOCATION OF PEREMPUAN/WOMAN'

Perempuan Parliament 1			Perempuan Parliament 2			Perempuan Parliament 3			Perempuan Parliament 4			Perempuan Parliament 5		
Fre	MI	Collocat	Fre	MI	Collocat	Fre	MI	Collocat	Fre	MI	Collocat	Fre	MI	Collocat
q	Scores	e	q	Scores	Collocat	q	Scores	e	q	Scores	e	q	Scores	e
39	11.4983	lelaki	24	13.2206	bersuami	51	12.8843	lelaki	12	13.0803	merogol	90	13.1353	lelaki
68	11.3378	laki	167	13.0333	laki	36	11.3131	perempu	119	12.7121	lelaki	38	11.5408	perempu
21	9.45327	budak	32	12.8767	perkumpul	10	11.0669	shah	68	11.4585	perempu	21	10.8471	laki
18	9.14472	tangga	19	12.7096	an	17	8.59442	penyakit	28	10.5245	laki	13	8.25983	pesakit
80	9.08107	murid	44	12.6526	ls	13	8.48635	kanak	21	9.92793	pergurua	10	7.70925	menenga
11	9.04291	hostel	23	12.5139	gadis	17	7.91882	seorang	22	9.07903	maktab	29	7.47806	murid
22	9.02705	lanjutan	10	10.8855	berkahwin	15	7.34878	murid	19	8.80973	budak	15	7.47224	penuntut
45	8.89394	kaum	66	10.6686	perempua	14	7.04739	raja	41	8.01146	doktor	34	7.41252	seorang
15	8.88692	maktab	12	10.4080	an	22	6.75859	pekerja	13	7.57758	pesakit	18	6.96965	hospital
139	8.86855	anak	17	9.98068	angkat	12	6.42536	bilangan	33	7.49407	murid	14	6.54566	pelajar
110	8.82302	guru	11	9.91295	perlindung	22	6.41264	sekolah	10	7.14942	penyakit	19	6.21322	guru
32	8.64038	perempu	12	9.47266	an	74	6.28006	orang	28	6.80289	seorang	18	6.12982	dia
15	8.57236	terdiri	38	9.46169	akhlak	12	6.12467	guru	17	6.5587	islam			
38	8.55976	gaji	10	8.86801	kanak	14	6.09411	anak	13	6.46409	penuntut			
18	8.53215	penuntut	11	8.68358	pemuleha				17	6.43228	melayu			

Perempuan Parliament 6			Perempuan Parliament 7			Perempuan Parliament 8			Perempuan Parliament 9			Perempuan Parliament 10		
Fre	MI	Collocat	Fre	MI	Collocat									
q	Scores	e	q	MI Scores	e									
19	14.7531	bersuami	112	12.7851	lelaki	17	13.7927	bersuami	21	12.7263	laki	274	11.836	lelaki
68	12.6905	lelaki	72	12.4283	perempu	152	12.2053	lelaki	204	12.3043	lelaki	110	11.265	perempu
32	11.8214	perempu	11	10.9275	an	11	12.1033	wahai	64	11.0602	perempu	17	9.6432	kahwin
12	10.5355	berumur	10	7	kahwin	94	11.6235	perempu	14	10.1641	membun	14	9.1690	tingkatan
15	10.2869	laki	12	9.99539	gadis	27	11.0353	an	16	10.1591	berkahwi	19	8.9815	cantik
10	8.79809	tua	42	8.79129	tua	16	10.9342	nabi	19	9.80391	berumur	10	8.8173	berkahwi
14	8.50131	pesakit	10	8.71639	kanak	28	11.0353	musa	39	9.26853	budak	20	8.4903	n budak
24	8.22719	kanak	14	8.53236	umur	11	11.0353	budak	10	8.65881	suami	11	8.4600	berumur
31	7.76101	seorang	27	8.22605	muda	36	9.20852	cantik	13	8.42717	remaja	16	8.3532	kamu
35	7.28131	anak	48	8.04712	murid	11	9.19789	umur	77	8.37203	kanak	28	7.7302	murid
16	7.06721	murid	23	7.83941	seorang	14	8.71675	kamu	12	7.93461	umur	16	7.7227	isteri
11	7.06495	doktor	35	7.02301	pelajar	30	8.68749	isa	14	7.9054	tua	36	7.3926	kanak
11	6.04306	pelajar	22	7.00187	anak	85	7.43142	kanak	11	7.25725	india	20	7.3153	doktor
13	6.02223	dia	13	6.82803	kaum	34	7.35937	anak	73	7.24529	seorang	12	7.3055	tua
			14	6.54688	melayu	55	7.04195	pelajar	90	7.1859	anak	12	7.2831	umur
				6.31152	bilangan		6.91153	seorang					2	

Perempuan Parliament 11			Perempuan Parliament 12			Perempuan Parliament 13		
Freq	MI Scores	Collocate	Freq	MI Scores	Collocate	Freq	MI Scores	Collocate
394	11.98522	lelaki	394	11.98522	lelaki	18	14.64514	zainab
170	11.46335	perempuan	170	11.46335	perempuan	216	12.50901	lelaki
11	10.09979	jantina	11	10.09979	jantina	18	12.42791	laki
26	9.80564	berkahwin	26	9.80564	berkahwin	64	11.48653	perempuan
11	9.15258	gila	11	9.15258	gila	11	10.68672	mengandung
13	8.71888	kahwin	13	8.71888	kahwin	18	10.02749	berkahwin
10	8.38309	gender	10	8.38309	gender	22	9.75873	remaja
13	8.32149	berumur	13	8.32149	berumur	10	9.71393	adik
23	8.21223	budak	23	8.21223	budak	22	9.23144	budak
12	8.03917	simpanan	12	8.03917	simpanan	10	9.02896	kahwin
13	7.8853	cantik	13	7.8853	cantik	13	8.66312	berumur
13	7.67465	warganegara	13	7.67465	warganegara	27	8.65117	umur
14	7.67407	suami	14	7.67407	suami	15	8.09486	isteri
10	7.50994	kem	10	7.50994	kem	78	7.80439	kanak
20	7.40714	isteri	20	7.40714	isteri	75	7.481	seorang

ABOUT THE AUTHOR

Jamaluddin Aziz (Ph.D.) is an Associate Professor in Critical Media Studies at Centre for Research in Media and Communication, FSSK, UKM. His research centres on the media and culture, gender and digital humanities. He was seconded to the Ministry of Women, Family and Community Development from 2015-2016 as a gender expert.