Exploring Gender Issues Associated With \textit{wanita}/woman and \textit{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 \textit{wanita}/woman and \textit{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 \textit{wanita}/woman and \textit{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 \textit{wanita}/woman and \textit{perempuan}/woman collocate rather consistently with the word \textit{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 \textit{wanita}/woman and \textit{perempuan}/woman have been discussed, that is, from women functioning in private sphere (collocates like \textit{anak}/child and \textit{mengandung}/pregnant to a more public one (collocates like \textit{guru}/teacher and \textit{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; \textit{wanita}/woman; \textit{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 Maiddin (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 Dendance (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, 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.

| 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 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 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 *wanita/woman* and *perempuan/woman* in MHC and the highest frequencies for *perempuan/woman*. |
| 5          | The lowest frequencies of *wanita/woman* and *perempuan/woman* in MHC |
| 10         | The highest frequencies of *wanita/woman* in MHC |
| 13         | The last discussion of *wanita/woman* and *perempuan/woman* 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*)

| Freq | MI Scores | Collocate          |
|------|-----------|--------------------|
| 100  | 9.7141    | *Kaum/group*       |
| 27   | 9.67348   | *Laki/man*         |
| 12   | 9.6222    | *Pemudi/young lady* |
| 100  | 9.62055   | *Wanita/woman*     |
| 10   | 9.20304   | *Lelaki/man*       |
| 11   | 9.13409   | *Lemah/weak*       |
| 7    | 7.99659   | *Maju/advance*     |
| 11   | 7.86661   | *Askar/soldier*    |
| 18   | 7.80246   | *Merasa/feel*      |
| 23   | 7.50355   | *Gaji/salary*      |
| 26   | 7.48181   | *Polis/police*     |
| 14   | 7.21898   | *Selalu/often*     |
| 16   | 7.12943   | *Tangan/hand*      |
| 13   | 6.90889   | *Latehan/training* |
| 20   | 6.83859   | *Berikan/to give*  |

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*)

| Freq | MI Scores | Collocate             |
|------|-----------|-----------------------|
| 104  | 11.99299  | *Wanita/woman*        |
| 33   | 11.3265   | *Gadis/young girl*    |
| 19   | 10.39127  | *Lelaki/man*          |
| 21   | 9.88301   | *Pakaian/clothing*    |
| 11   | 8.97968   | *Perlindungan/protection* |
| 62   | 7.90469   | *Kaum/group*          |
| 22   | 7.38398   | *Anggota/body*        |
| 43   | 7.38367   | *Islam/Islam*         |
| 42   | 6.83062   | *Pekerja/worker*      |
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.

| Freq | MI Scores | Collocate |
|------|-----------|-----------|
| 32   | 6.82488   | Seorang/a person |
| 15   | 6.45794   | Polis/police |
| 10   | 6.42872   | Asing/ alien |
| 13   | 6.13176   | Terutama/main |

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*.

| Freq | MI Scores | Collocate |
|------|-----------|-----------|
| 120  | 11.03604  | Selangau/a town in Sarawak |
| 20   | 10.89165  | Tuala/towel |
| 36   | 10.13761  | Hamil/pregnant |
| 18   | 9.84656   | Incubator |
| 10   | 9.52908   | Pencinta/lover |
| 51   | 9.47965   | Mengandung/pregnant |
| 1298 | 9.18145   | Keluarga/family |
| 229  | 9.17325   | Lelaki/man |

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

| Freq | MI Scores | Collocate                      |
|------|-----------|--------------------------------|
| 39   | 11.49836  | *Lelaki*/man                   |
| 68   | 11.33789  | *Laki*/man                     |
| 21   | 9.45327   | *Budak*/a child                |
| 18   | 9.14472   | *Tangga*/steps                 |
| 80   | 9.08107   | *Murid*/student                |
| 11   | 9.04291   | hostel                         |
| 22   | 9.02705   | *Lanjutan*/extension           |
| 45   | 8.89394   | *Kaum*/group                   |
| 15   | 8.88692   | *Maktub*/institute             |
| 139  | 8.86855   | *Anak*/child                   |
| 110  | 8.82302   | *Guru*/teacher                 |
| 32   | 8.64038   | *Perempuan*/woman              |
| 15   | 8.57236   | *Terdiri*/constitute           |
| 38   | 8.55976   | *Gaji*/salary                  |
| 18   | 8.53215   | *Penuntut*/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*)**

| Freq | MI Scores | Collocate                      |
|------|-----------|--------------------------------|
| 90   | 13.13537  | *Lelaki*/man                   |
| 38   | 11.54082  | *Perempuan*/woman              |
| 21   | 10.84713  | *Laki*/man                     |
| 13   | 8.25983   | *Pesakit*/patient              |
| 10   | 7.70925   | *Menengah*/secondary           |
| 29   | 7.47806   | *Murid*/student                |
| 15   | 7.47224   | *Penuntut*/student             |
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 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.
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 shows the binary pattern of *wanita-woman-lelaki-man*. 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/man* 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*.
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.

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

| Parliament | Freq (Wanita/woman) | Freq (Perempuan/woman) | 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  |

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
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:

| Parliament | Frequency | MI Scores | Collocate          |
|------------|-----------|-----------|--------------------|
| 1          | 100       | 9.62055   | Wanita/woman       |
| 5          | 104       | 11.99299  | Wanita/woman       |
| 10         | 1436      | 8.82114   | Wanita/woman       |
| 13         | 1298      | 9.18145   | Keluarga/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.

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-nyaa belum saya berharap kapada Kementerian yang berkenaia supaya mengadakan Pegawai2 Wanita dalam soal ini, kerana kemajuan di-luar bandar itu ada-lah sa-bahagian besar-nya di-pegang oleh tanggongan 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 bermacham2 pandangan kapada 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.

The context of the issue is identified as the following:

9 DESEMBER 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 dita as oleh kerana ianya tidak sejajar dengan padaan wanita-wanita Islam, dan jika ia, sama ada Kerajaan saya yakin dan percaya apa yang diharapkan oleh wanita-wanita Islam khasnya dan umat Islam di negara innya yang mana hukuman hanyalah dikenakan ke atas wanita-wanita Islam manakala pasangan yang terdiri dari lelaki kenaan 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 dikehendak supaya ditambah dause tertentu, khususnya kepada wanita-wanita Islam selain daripada kebenaran Menteri atau Ketua 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 dibenarkan ran kecil nanti mengekori peraturan ini khas bagi wanita-wanita Islam. Tuan Yang di-Pertua, Adakah Yang

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.

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 ah dengan kekhlasan untuk meningkatkan kedudukan wanita. Sekiranya kerajaan menganggap wanita adalah sama rata dengan

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.

The context of the issue is identified as the following:

21.11.2016

[...] saya ingin bertanya kepada kementerian tentang tindakan yang dibuat supaya tidak belaku 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:

| 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 quality of the training and education of daughters.

tanggal 5021 6 FEBRUARY 1961 5022 ia-itu di-mana anak2 perempuan kita dapat di-beri latehan dan pendidekan ini ia-ah bagi mem-beri peluang kapada anak2 perempuan kita daripada sekolah kebangsaan yang di-kampung tidak berapa mengambil berat tentang hal kemajuan anak2 perempuan kita di-Indonesia. Beliau pergi hanya untok mem-beri pendidekan yang sa-luas2-nya kapada anak2 perempuan kita di-dalam negeri ini. Saya merasa ini supaya membanyak-kan sekolah bagi pendidekan anak perempuan kita di-Tanah Melayu ini. Enchex92 -nya datang-lah register maka terpaka-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 kapada anak2 perempuan kita ini maka dapat-lah kita menjadikan anak2 itu sa-bagai pelateh di-kampong2 sa-bagaimana dasar Kerajaan pada hari ini akan mengadakan pendidekan yang sa-baik2-nya bagi pendudok2 di-luar bandar baik bagi anak2 laki2 mahu pun anak2 perempuan kita.
[...] Oleh sebab itu saya rasa, kalau pihak Kerajaan lebeh2 lagi pihak 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 kapada anak kita untuk di-jadikan pemimpin atau pelateh bagi wanita 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 is 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.

The contextual evidence is found as the following:

3 DESEMBER 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.

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.

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/10.1002/978118611463.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-idUSKCN1IM11U

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** |
| 100 | 9.7141 | kaum | 12 | 12.703 | pandu | 12 | 12.671 | juri | 10 | 12.0781 | bersuami | 104 | 11.792 | wanita |
| 27 | 9.67348 | laki | 14 | 11.214 | pengakap | 61 | 12.601 | gadis | 292 | 11.7137 | wanita | 33 | 11.326 | gadis |
| 12 | 9.6222 | pemudi | 189 | 11.064 | kaum | 104 | 11.666 | wanita | 84 | 11.4633 | gadis | 19 | 10.391 | lelaki |
| 100 | 9.62055 | wanita | 13 | 11.038 | gadis | 21 | 10.672 | perlindungan | 29 | 11.0736 | berkahwin | 10 | 9.883 | pakai |
| 10 | 9.20304 | lelaki | 25 | 10.814 | taman | 10 | 9.9451 | lelaki | 41 | 10.2513 | lelaki | 11 | 8.0976 | pakaian |
| 11 | 9.13409 | lemah | 16 | 9.7942 | laki | 76 | 9.1626 | kaum | 31 | 9.25698 | perlindungan | 62 | 7.904 | wanita |
| 7.99659 | maju | 11 | 9.4917 | akhlak | 20 | 7.5246 | islami | 12 | 9.03909 | bersalin | 22 | 7.3839 | anggota |
| 7.86661 | asar | 20 | 9.2353 | wanita | 10 | 6.7856 | latihan | 10 | 8.90327 | cantik | 43 | 7.3836 | wanita |
| 18 | 7.80246 | merasa | 13 | 8.7410 | tunjut | 6 | 7.320 | gaji | 10 | 8.92289 | memakai | 42 | 6.8396 | pekerja |
| 23 | 7.50355 | gaji | 11 | 8.2282 | perempuan | 6 | 10.5646 | seorang | 15 | 8.16257 | pakai | 32 | 6.8248 | seorang |
| 26 | 7.48181 | polis | 39 | 7.9963 | latihan | 27 | 6.4653 | pekerja | 11 | 8.12436 | akhlak | 15 | 6.4747 | polis |
| 14 | 7.21898 | selalu | 10 | 7.7082 | pertubuhan | 20 | 8.09674 | pemulihan | 10 | 6.4287 | pakaian | 12 | 6.1317 | berkahwin |
| 16 | 7.12943 | tangan | 28 | 7.6080 | gaji | 88 | 7.98785 | kaum | 13 | 6.1317 | asing |
| 13 | 6.90889 | perempuan | 30 | 7.5569 | polis | 71 | 7.51328 | wanita | 14 | 6.5237 | berkahwin |
| 20 | 6.83859 | ibu | 17 | 7.5523 | berkahwin | 35 | 7.41724 | alic | 16 | 6.17385 | perempuan |

| 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** |
| 236 | 11.03 | wanita | 14 | 13.64 | bersuami | 12 | 10.64 | penghidupan | 28 | 9.026 | mendiskriminasi | 14 | 10.318 | wanita |
| 466 | 10.97 | gadis | 31 | 11.87 | gaji | 312 | 10.21 | wanita | 25 | 9.253 | mendiskriminasi | 14 | 10.247 | wanita |
| 1688 | 10.71 | gacok | 105 | 9.87 | gadi | 1038 | 9.484 | wanita | 263 | 10.16 | mendiskriminasi | 14 | 10.052 | wanita |
| 13 | 10.60 | laki | 11 | 11.63 | laki | 18 | 9.65 | wanita | 10 | 9.713 | mendiskriminasi | 14 | 10.05 | wanita |
| 57 | 10.17 | bersal | 238 | 10.89 | bersal | 48 | 8.465 | wanita | 10 | 9.453 | mendiskriminasi | 14 | 9.536 | wanita |
| 30 | 9.856 | memem | 38 | 10.08 | wanita | 12 | 0.08 | wanita | 12 | 0.08 | mendiskriminasi | 14 | 9.536 | wanita |
| 20 | 9.805 | berka | 82 | 0.12 | wanita | 182 | 0.12 | wanita | 97 | 0.12 | mendiskriminasi | 14 | 9.536 | wanita |
| 89 | 9.668 | berka | 36 | 9.780 | wanita | 182 | 0.12 | wanita | 97 | 0.12 | mendiskriminasi | 14 | 9.536 | wanita |
| 236 | 8.894 | berka | 47 | 8.686 | wanita | 182 | 0.12 | wanita | 97 | 0.12 | mendiskriminasi | 14 | 9.536 | wanita |
| 17 | 8.446 | status | 13 | 8.94 | wanita | 182 | 0.12 | wanita | 97 | 0.12 | mendiskriminasi | 14 | 9.536 | wanita |
| 16 | 8.195 | penuh | 11 | 9.416 | wanita | 182 | 0.12 | wanita | 97 | 0.12 | mendiskriminasi | 14 | 9.536 | wanita |
| 11 | 7.725 | wanita | 11 | 9.052 | wanita | 182 | 0.12 | wanita | 97 | 0.12 | mendiskriminasi | 14 | 9.536 | wanita |
| 114 | 7.658 | wanita | 11 | 9.052 | wanita | 182 | 0.12 | wanita | 97 | 0.12 | mendiskriminasi | 14 | 9.536 | wanita |
| 23 | 7.657 | wanita | 11 | 9.052 | wanita | 182 | 0.12 | wanita | 97 | 0.12 | mendiskriminasi | 14 | 9.536 | wanita |
| 38 | 7.433 | wanita | 11 | 9.052 | wanita | 182 | 0.12 | wanita | 97 | 0.12 | mendiskriminasi | 14 | 9.536 | wanita |
| 11 | 6.36 | wanita | 11 | 9.052 | wanita | 182 | 0.12 | wanita | 97 | 0.12 | mendiskriminasi | 14 | 9.536 | wanita |

---

*Note:* Collocate means the word that is frequently collocated with the word in question.

---

**eISSN:** 2550-2131  
**ISSN:** 1675-8021
| Freq | Wanita Parliament 11 | MI Scores | Collocate | Freq | Wanita Parliament 12 | MI Scores | Collocate | Freq | Wanita Parliament 13 | MI Scores | Collocate |
|------|----------------------|-----------|-----------|------|----------------------|-----------|-----------|------|----------------------|-----------|-----------|
| 213  | 10.61986             | jelebu    | 10        | 12.45268 | bedonisene          | 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 | semi              | 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        | 10.71372 | tatassila         | 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 | ceduw |
### TABLE OF COLLOCATION OF PEREMPUAN/WOMAN

| Perempuan Parliament 1 | Perempuan Parliament 2 | Perempuan Parliament 3 | Perempuan Parliament 4 | Perempuan Parliament 5 |
|------------------------|------------------------|------------------------|------------------------|------------------------|
| Fre | MI Scores | Collocate | Fre | MI Scores | Collocate | Fre | MI Scores | Collocate | Fre | MI Scores | Collocate |
| 39 | 11.4983 | 13.2206 | 51 | 12.8843 | lelaki | 12.7263 | lelaki |
| 68 | 11.3378 | 13.0333 | 51 | 12.8843 | lelaki | 12.7263 | lelaki |
| 21 | 9.4537 | 12.8767 | 10 | 5 | shah | 10.5245 | shah |
| 18 | 9.1447 | 12.6526 | 13 | 8.4635 | kanak | 9.92793 | kanak |
| 80 | 9.0817 | 12.5139 | 17 | 7.91882 | seorang | 9.07903 | seorang |
| 11 | 9.0429 | 10.8855 | 15 | 7.34878 | murid | 8.80973 | murid |
| 45 | 8.8939 | 10.6868 | 14 | 7.04739 | raja | 8.01146 | raja |
| 15 | 8.8862 | 10.4080 | 22 | 6.75859 | pekerja | 7.75778 | pekerja |
| 139 | 8.8655 | 9.90608 | 12 | 6.42536 | bilangan | 7.49407 | bilangan |
| 110 | 8.8230 | 9.91295 | 22 | 6.41264 | sekolah | 7.14942 | sekolah |
| 32 | 8.6403 | 9.47266 | 74 | 6.28006 | orang | 6.80289 | orang |
| 15 | 8.5723 | 9.46169 | 12 | 6.12467 | guru | 6.5577 | guru |
| 38 | 8.5597 | 8.88081 | 14 | 6.09411 | anak | 6.46409 | anak |
| 18 | 8.53215 | 8.88358 | 17 | 6.43228 | pelajar | 6.96965 | pelajar |

### TABLE OF COLLOCATION OF LELEKAN/LELEKAN

| Perempuan Parliament 6 | Perempuan Parliament 7 | Perempuan Parliament 8 | Perempuan Parliament 9 | Perempuan Parliament 10 |
|------------------------|------------------------|------------------------|------------------------|------------------------|
| Fre | MI Scores | Collocate | Fre | MI Scores | Collocate | Fre | MI Scores | Collocate | Fre | MI Scores | Collocate |
| 19 | 14.7531 | 12.7851 | 17 | 13.7927 | bersuami | 12.7263 | lelaki |
| 68 | 11.8214 | 12.4283 | 17 | 12.2053 | lelaki | 12.3043 | lelaki |
| 32 | 10.5355 | 10.9275 | 11 | 12.1033 | lelaki | 11.0602 | lelaki |
| 12 | 10.2869 | 11.8593 | 9 | 11.6235 | lelaki | 10.1641 | lelaki |
| 15 | 8.7890 | 8.71639 | 16 | 8.80391 | lelaki | 9.92793 | lelaki |
| 14 | 8.50131 | 8.5326 | 28 | 9.2852 | lelaki | 8.65881 | lelaki |
| 24 | 8.22719 | 8.22665 | 11 | 9.20852 | lelaki | 8.65881 | lelaki |
| 31 | 7.76101 | 8.04712 | 36 | 8.42177 | lelaki | 8.7302 | lelaki |
| 35 | 7.28131 | 8.03941 | 13 | 8.71675 | lelaki | 7.3926 | lelaki |
| 16 | 7.06721 | 7.02301 | 14 | 8.6874 | lelaki | 7.93461 | lelaki |
| 11 | 7.06495 | 7.00187 | 30 | 7.9054 | lelaki | 7.3152 | lelaki |
| 11 | 6.04306 | 6.82803 | 85 | 7.35927 | lelaki | 7.25275 | lelaki |
| 13 | 6.02223 | 6.54688 | 73 | 7.25429 | lelaki | 7.25275 | lelaki |
| 14 | 6.31152 | 6.91153 | 90 | 7.1859 | lelaki | 6.80289 | lelaki |

**Notes:**
- Fre: Frequency
- MI: Mutual Information
- Collocate: Collocation
- Scores: Scores
### 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.

| Perempuan Parliament 11 | Freq | MI Scores | Collocate |
|------------------------|------|-----------|-----------|
| lelaki                 | 394  | 11.98522  |           |
| perempuan              | 170  | 11.46335  |           |
| jantina                | 11   | 10.09979  |           |
| berkahwin              | 26   | 9.80564   |           |
| gila                   | 11   | 9.15258   |           |
| kahwin                 | 13   | 8.71888   |           |
| gender                 | 10   | 8.38309   |           |
| berumur                | 13   | 8.32149   |           |
| simpanan               | 12   | 8.03917   |           |
| cantik                 | 13   | 7.8853    |           |
| warganegara            | 13   | 7.67465   |           |
| suami                  | 14   | 7.67407   |           |
| kem                    | 10   | 7.50994   |           |
| isteri                 | 20   | 7.40714   |           |

| Perempuan Parliament 12 | Freq | MI Scores | Collocate |
|------------------------|------|-----------|-----------|
| lelaki                 | 394  | 11.98522  |           |
| perempuan              | 170  | 11.46335  |           |
| jantina                | 11   | 10.09979  |           |
| berkahwin              | 26   | 9.80564   |           |
| gila                   | 11   | 9.15258   |           |
| kahwin                 | 13   | 8.71888   |           |
| gender                 | 10   | 8.38309   |           |
| berumur                | 13   | 8.32149   |           |
| simpanan               | 12   | 8.03917   |           |
| cantik                 | 13   | 7.8853    |           |
| warganegara            | 13   | 7.67465   |           |
| suami                  | 14   | 7.67407   |           |
| kem                    | 10   | 7.50994   |           |
| isteri                 | 20   | 7.40714   |           |

| Perempuan Parliament 13 | Freq | MI Scores | Collocate |
|------------------------|------|-----------|-----------|
| zainab                 | 216  | 12.50901  | lelaki    |
| laki                   | 18   | 12.42791  |           |
| perempuan              | 64   | 11.48653  |           |
| mengandung             | 11   | 10.68672  |           |
| kahwin                 | 18   | 10.02749  | berkahwin |
| remaja                 | 22   | 9.75873   |           |
| adik                   | 10   | 9.71393   |           |
| budak                  | 22   | 9.23144   |           |
| kahwin                 | 10   | 9.02896   |           |
| berumur                | 13   | 8.66312   |           |
| warganegara            | 27   | 8.65117   |           |
| umur                   | 15   | 8.09486   |           |
| kem                    | 78   | 7.80439   |           |
| seorang                | 75   | 7.481     |           |