

The Portrayal of Women: A Corpus Analysis of the Representation around the Word 'Wanita' in the Malaysian Hansard Corpus

Anis Nadiah Che Abdul Rahman

aniscar@ukm.edu.my

Centre for Research in Language and Linguistics
Universiti Kebangsaan Malaysia, Malaysia

ABSTRACT

The term "women's issues" is often used too broadly, which can obscure the specific challenges women face. Understanding how parliamentarians categorize women's issues is crucial for grasping their perspectives from a political dimension. This allows politicians to shape discussions through their speeches, as their chosen words significantly influence how women's issues are framed and contextualized. This paper examines the Malaysian Hansard Corpus, a collection of parliamentary debates, to analyze how politicians conceptualize and present women as a political issue over a period of 61 years. Using corpus linguistics, this study investigates the representation of women through noun collocates of 'wanita' (woman). WordSmith Tool 7.0 and a Part of Speech tagger extracted a list of collocates with semantic content to analyze the semantic preferences associated with the term 'wanita' over time. The findings indicate that politicians semantically associate 'wanita' with categories related to national identity, attributes, conflict, culture, religion, public affairs, and quantification. Notably, discussions surrounding women's issues have become more favorable throughout parliamentary sessions. This research contributes to corpus linguistics by demonstrating how linguistic patterns can reveal shifts in political discourse over time. By systematizing these findings, the study provides a framework for understanding women's issues in Malaysian parliamentary debates through their semantic representations, illuminating how language use reflects evolving attitudes and priorities regarding women's rights and representation in political contexts.

Keywords: Corpus linguistics; women; political discourse; Malaysian Hansard Corpus; semantic preference

INTRODUCTION

Women have always played dynamic roles in society, greatly contributing to national stability and growth. Their responsibilities have developed from largely domestic roles to active participation in several sectors, including politics and the workforce. Despite this improvement, women remain underrepresented in political sectors, with only 25.5% of national Members of Parliament (MPs) globally being female (United Nations, 2021). Research demonstrates that higher female presence in politics is associated with improved health outcomes for women and children (Dhatt et al., 2017). Social movements campaigning for increased access to education and legislative opportunities are major drivers of these transformations (United Nations, 2021). This ongoing struggle highlights the need for continued advocacy and support for women's representation in all the areas of society.

Gender inequality continues to be a crucial issue, particularly in the context of global challenges such as the COVID-19 epidemic. This pandemic has exposed the systemic inequities that undermine women's health, economic security, and social protection (Badri, 2020). In Malaysia, women typically experience misconceptions and discrimination, combined with conventional roles that promote family obligations (Kok, 2020). Although there has been significant improvement, demonstrated by an increase in female candidates during the 14th general election, only 11% were elected, emphasizing the ongoing socio-cultural impediments (Pey, 2018). These challenges hinder women's ability to construct identities and discourses, sometimes determined by more important political individuals (Baker et al., 2008). In addition to that, the study by Jamaluddin (2019) explores gender issues associated with the terms '*wanita*' (woman) and '*perempuan*' (woman) in Malaysian parliamentary debates, revealing trends in how women's issues are discussed.

The term "women's issues" is commonly abused, bearing negative implications that can hamper legitimate dialogue. Politicians play a key impact in moulding public perceptions and policy decisions through their speeches and debates, framing talks about women in ways that can empower or marginalize them. Effective public policies originate from identifying social challenges and modifying existing policies during parliamentary discussions. Parliaments offer useful sources for assessing political speech, as they are dynamic institutions where elected members engage in communication and debate (Schwab et al., 2020). Research in linguistics has explored the portrayal of women in diverse discourses, indicating a largely negative representation. Studies have demonstrated that the term 'women' often arises in circumstances that perpetuate stereotypes and bias (Pearce, 2008; Ndambuki & Janks, 2010; Dabbagh, 2016; Jia et al., 2016).

This study aims to examine the construction of women's issues in Malaysian parliamentary discourse by exploring the use of the lexical '*wanita*' (woman) from Parliament 1 to Parliament 14. It will evaluate the semantic preferences related to women and categorise the issues discussed, thereby contributing to comprehending how women are represented in political discourse. The results will be compared to the existing literature to identify patterns in the representation of women in various contexts.

PREVIOUS STUDIES

PARLIAMENT OF MALAYSIA AND STUDIES ON HANSARD MALAYSIA

The Malaysian Parliament is vital for the country's governance, notably through its legislative discussions and processes. Over the previous 60 years, the Dewan Rakyat (House of Representatives) has conducted more than 3,500 discussions. As the lower house of Parliament, which also comprises the Senate (Dewan Negara) and the King, its principal job is to propose and debate bills. Both chambers must approve these bills before they get royal approval from the King. Members of the Dewan Rakyat, known as Members of Parliament (MPs), are chosen from federal constituencies, representing the interests of their constituents in legislative concerns.

The role of Members of Parliament (MPs) in addressing women's issues is particularly crucial, even with the presence of other advocacy groups. Numerous MPs have actively offered and debated measures involving women's rights and related topics. Such initiatives are vital, as they can lead to real legislative actions that positively benefit women's lives.

According to Willis (2017), while politicians typically convert information into effective solutions, they tend to alter issues to meet their ideological positions, notably in conversations concerning climate change. This dynamic emphasizes the intricate interplay between political ideology and legislative processes in tackling significant societal concerns.

The Hansard serves a critical function in the legislative process by giving the official transcripts of parliamentary debates, thus promoting transparency and accountability in government. This website allows researchers and individuals to examine political speech and learn MPs' opinions on numerous issues. Transcripts of debates give a plethora of information for data mining regarding the ideas and attitudes of politicians and the parties they represent on global and local topics. Additionally, these data can provide insights into the democratic processes occurring within legislative bodies, as observed by Abercrombie and Batista-Navarro (2017).

As online electronic archives become increasingly accessible, quantitative methods have evolved as a favored analysis method (Baker, Gabrielatos, & McEnery, 2013). These technologies enable researchers to mine enormous databases over extended periods. Van Dijk (2004) underlines that records of parliamentary sessions are useful resources across multiple disciplines, including Digital Humanities and Social Sciences, due to their distinctive content, structure, and language. Previous investigations of Hansard-related corpora have offered useful insights into politicians' motivations and strategies. For instance, Rheault, Beelen, Cochrane, and Hirst (2016) analysed the emotional expressions of politicians in the British House of Commons Hansard, revealing light on the emotional components of political speech.

Extensive research into the Malaysian Hansard has been done within the Malaysian academic debate. Nonetheless, a substantial gap persists in examining matters concerning women's representation in legislative discourse. Prior research has investigated diverse subjects, such as economic representation (Nor Fariza et al., 2019), water management (Norsimah et al., 2019), and semantic shifts in legislative language (Sabrina et al., 2020). While Jamaluddin (2019) examined gender issues related to women in the Malaysian Hansard, the particular facet of women's representation in legislative texts was not comprehensively investigated. The disparity is significant, as the inclusion of women in legislative discussions is essential for comprehending the representation and priority of women's issues in public policy. Additional research could provide significant insights into the visibility and framing of women's issues in legislative discussions and resolutions.

REPRESENTATION OF WOMEN

In recent years, extensive research has been conducted on the representation of women, exploring various aspects from diverse perspectives. While numerous studies have focused on media representation, relatively little attention has been paid to how women's issues are represented in parliamentary contexts, particularly in relation to the number of women Members of Parliament (MPs). Today, the representation of women in the political sphere has evolved significantly. Studies examining women's representation across different discourses often reveal patterns of negative portrayal.

For example, Ponterotto (2014) analyzed the representation of female athletes in the media using a corpus-assisted approach, focusing on sports sections of major British and Italian newspapers, specifically *The Times* and *La Repubblica*. The study centered on tennis player Maria Sharapova, describing her as “the sexiest woman in the world” in thematic analysis. At the same time, metaphorical language depicted her as experiencing “bloodcurdling howls.” This research

concluded that such representations are driven by gender stereotypes and reflect inherent biases against women.

Similarly, Ndambuki and Janks (2010) conducted a Critical Discourse Analysis of women's issues in the Makueni District of Kenya. Their study highlighted that women often lack the platform to voice their opinions on important issues, portraying them as a marginalized group in need of male guidance. The findings emphasized that women are frequently represented within a patriarchal framework, characterized as impoverished and outdated.

Dabbagh (2016) further examined women's statuses in Morocco by analyzing the discourse surrounding their representation on two prominent Arab television networks. This analysis revealed that women are often associated with domestic and traditional roles, while alternative narratives portray them as vital contributors to modernity and economic growth.

In another study, Jia, Lansdall-Welfare, Sudahar, Carter, and Cristianini (2016) explored gender representation in online English media by analysing both words and images. Their research found that women appear more frequently in images than in text across various topics, and they are often the grammatical subjects of verbs (Pearce, 2008).

The studies highlight how binary portrayals of women in Malaysian society significantly limit their identities and roles, ultimately hindering their ability to navigate societal expectations and assert their rights. For instance, research on parliamentary discourse reveals that terms like '*wanita*' and *perempuan* often frame women's issues solely in contrast to men's, reinforcing traditional gender roles (Jamaludin, 2019). Similarly, an examination of religious commitment among rural older women shows that cultural expectations constrain their identities and participation in societal matters (Wan Ibrahim, 2023). Additionally, literature reflecting on ethnicity and gender in Malaysia underscores the need to move beyond simplistic categorizations to acknowledge the complexities of women's experiences (Raihanah, 2023). Together, these insights demonstrate that such binary frameworks not only restrict women's self-expression but also diminish their opportunities across various spheres.

Based on previous research, it is evident that earlier studies largely overlooked general linguistic phenomena and did not thoroughly examine specific domains, such as political discourse. Additionally, while they provide valuable insights into language patterns and associations, they often do not address the implications of these findings for social issues, particularly regarding women's representation and rights. A significant gap exists in the literature regarding the representation of women specifically within parliamentary contexts. While substantial research has addressed media representation and societal portrayals, there is a lack of studies examining how women's issues are framed and discussed in political discourse. This gap indicates a need for research that explores the linguistic representation of women in political speech. The primary objectives of this study are to:

1. Analyze the collocational trends of collocates of the lexical item '*wanita*' in the Malaysian Hansard Corpus to understand how parliamentarians frame women's issues within political discourse overtime
2. Examine the topics associated with '*wanita*' to uncover the broader socio-political themes linked to discussions about women.

By focusing on these objectives, this study aims to contribute to a deeper understanding of how political language shapes perceptions of women and their representation in the Malaysian parliamentary context.

CORPUS LINGUISTICS AND SEMANTIC PREFERENCE

Compared to traditional quantitative and qualitative methods, the corpus approach offers a more thorough, empirical, and systematic framework for linguistic analysis. This study demonstrates the potential of corpus linguistics in examining collocations and semantic preferences related to the term '*wanita*' (woman). By utilizing a corpus approach, the research improves the reliability of results through the analysis of extensive and authentic language data, enabling researchers to identify patterns and differences in language more clearly than would be attainable with smaller data. The study emphasizes collocates, illustrating how legislators associate women's issues with broader socio-political topics, thereby underscoring the importance of the corpus approach in analyzing intricate phenomena like political discourse. This strategy not only reveals the linguistic framing of women's issues but also portrays the changing attitudes within political contexts.

Collocation and semantic preference are the two concepts employed interchangeably in this study. Collocation is the initial term used by Firth (1958) to describe the words that frequently co-occur with specific words. According to Cotteril (2001), collocation can expose a semantic shape of a word or phrase. Semantic preference refers to the tendency of particular words or phrases to be associated with specific meanings or contexts. It demonstrates how certain terms are more likely to appear together based on their meanings, leading to a favored usage in particular scenarios (Sinclair, 1991). Focusing on semantic preference is crucial because it reveals the underlying attitudes and values reflected in political discourse, offering insights into how specific contexts shape the representation of women's issues.

The literature on semantic preference examines the tendency of certain words to co-occur with meanings or contexts. Significant research by scholars such as Partington (1998) and Sinclair (1991) elucidates the patterns of these relationships, demonstrating their reflection of broader linguistic tendencies. Research has investigated semantic preference in diverse circumstances, encompassing collocations and cross-linguistic studies, like Begagić's examination of semantic prosody (2020) and Wei and Li's research of semantic preference in English and Chinese (2020). Liu (2020) examined the semantic prosody and preference of the verb "seem," whereas Selmistraitis (2020) analysed synonymous adjectives within the COCA corpus. This corpus of work highlights the significance of semantic preference in theoretical linguistics and practical language instruction. Nonetheless, a notable gap exists in research that particularly examines semantic choice in the political domain and its consequences for women. This research seeks to address that deficiency by examining the manifestation of semantic preferences in political speech concerning women.

METHODOLOGY

DATA COLLECTION

This study applies quantitative and qualitative methodology by combining the methodology associated with culturomics, corpus linguistics and semantic preference and semantic prosody. By employing quantitative and qualitative modes of enquiry, the researcher attempt to illuminate the trend of '*wanita*' over time. According to Baker (2006), a conventional corpus-driven analysis was inadequate to explain or construe the causes of specific linguistics patterns as it does not take into account political, historical and cultural context of the data.

Therefore, this study employs a corpus-driven analysis of the representation of the word ‘wanita’ (woman) in the Malaysian Hansard Corpus, a parliamentary corpus based on Malaysian parliamentary debates over a period of 61 years from 1959 to 2020.

This particular study utilised the Malaysian Hansard Corpus by Imran, Anis Nadiyah and Azhar (2021). The data brought together the Malaysian parliamentary records, or Malaysian Hansard, for the period 1959-2020 (Parliament 1 – Parliament 14). The raw corpus amounted to 165 million words in 3,684 texts (taken from one proceeding day), representing the output from nearly 2329 speakers throughout 60 years of parliamentary proceedings in Malaysia. The distribution of parliament according to date (year) can be depicted in Table 1.

TABLE 1. The distribution of Parliament according to year

Parliament	Year
1	1959 – 1964
2	1964 - 1969
3	1971 – 1973
4	1974 - 1978
5	1978 - 1981
6	1982 - 1986
7	1986 -1990
8	1990 - 1994
9	1995 -1999
10	1999 - 2003
11	2004 - 2007
12	2008 – 2012
13	2013 – 2018
14	2019 – 2020

Figure 1 illustrates the size of the corpus based on the Malaysian Hansard Corpus, which consists of 14 sub-corpora. During the time of data collection, Parliament 14 was still ongoing. As shown in Figure 1, Parliament 12 contains the highest number of words (tokens), whereas Parliament 1 has the lowest token count.

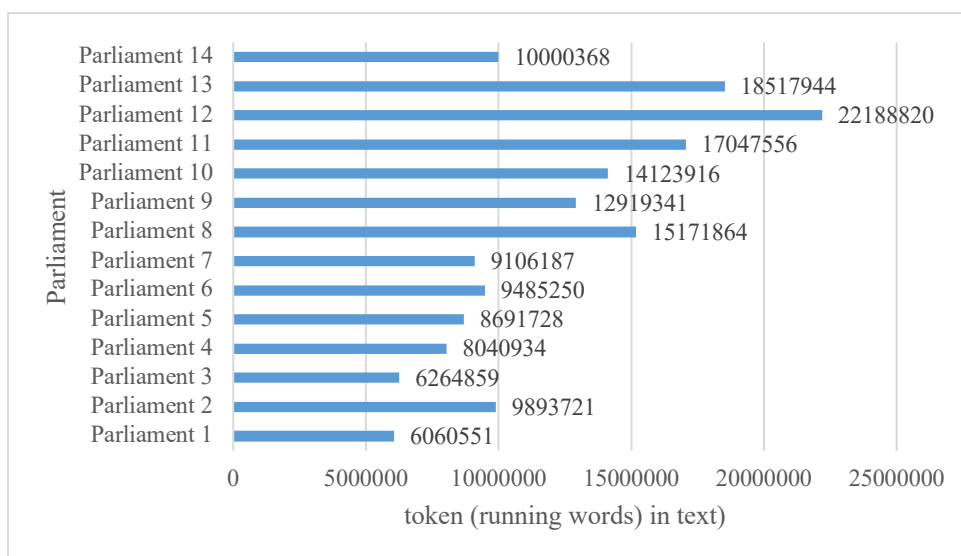


FIGURE 1. The size (token of running words) of sub corpora (parliament) in Malaysian Hansard Corpus (Imran et al. 2021)

PROCEDURAL ANALYSIS

This section describes the data analysis procedure for this study. It begins with pre-processing using WordSmith Tools 7.0 to extract basic statistics from the corpus, including its size and word list. Based on these statistics, a culturomics trend was developed using the search term ‘wanita’ which was not case sensitive. Next, the data processing involved categorizing topics related to ‘wanita.’ This required semi-automated annotation, as these topics could not be directly tagged from the raw corpus. The annotation process included selecting noun collocates and organizing them in Microsoft Excel. Additionally, a concordance analysis was performed using WordSmith Tools 7.0 to verify the contextual relevance of the categorization. Finally, a close examination of keywords in context (KWIC) was conducted to explore the semantic preference of the term ‘wanita’ in the Malaysian Hansard Corpus. The overall process is illustrated in Figure 2.

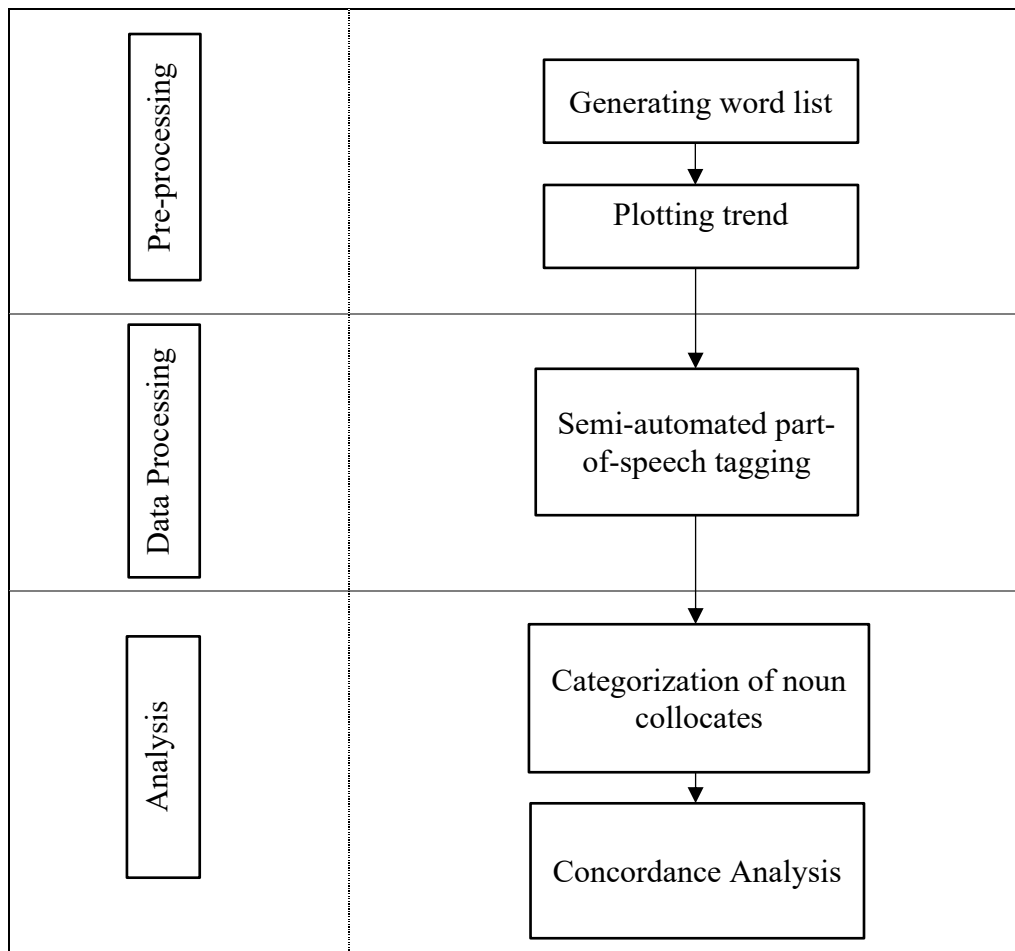


FIGURE 2. Procedural analysis

DATA ANALYSIS

CULTUROMICS

In the attempt to chart the trend of the lexis ‘wanita’ in Malaysian parliamentary proceedings across time, culturomics’ N-gram was generated by utilising the framework from McEnery and Hardie (2012). Culturomics is the quantitative examination of cultural trends through extensive text data, enabling researchers to investigate the evolution of language and cultural expressions throughout time (Michel et al., 2011). An N-gram refers to a sequence of 'n' items from a text, which can be used to track the frequency of specific phrases or terms, revealing patterns in language usage (Sculley & McCallum, 2003). To do this, a set of wordlist from each parliament was generated to see the number of mentions (frequency) of the node word ‘wanita’. The frequency threshold set for this study employs the threshold by Brezina, McEnery and Wattam (2015), where relative frequency is used rather than absolute frequency threshold because the corpus contains sub-corpora of different times and was generally and unevenly distributed. To get relative frequency, the frequency of the node word in each parliament was normalised. According to Warren and Pearse (2008), normalisation is a process that compares corpora with diverse sizes fairly. The normalisation can be achieved through the use of relative frequency. The formula can be seen in formula 1 as follows:

$$\text{Relative frequency} = \frac{\text{frequency of node word in corpus } X}{\text{total token of corpus } X} (1,000,000)$$

FORMULA 1. Normalisation process of frequency

COLLOCATIONAL ANALYSIS

The setting set for this particular study follows Collocational Parameter Notation (CPN) by Baker, McEnery and Brezina (2017). The distribution of CPN can be seen in Table 2.

TABLE 2. Collocational Parameter Notation (CPN) adapted from Baker, McEnery and Brezina (2017)

Notation categories	Statistic ID	Statistic Name	Statistic cut-off value	L and R span	Minimum collocate freq. (C)	Minimum collocation freq. (NC)	Filter
Example	2a	MI	3	L5 - R5	5	1	function words removed
In-text notation: 2a-MI(3), L5-R5, C5-NC1; function words removed (example)							

The association measure utilised in this study is MI-score (Mutual Information) which is the standard collocation measure. MI-score ascertains a relatively unique combination of words. Following the removal of function words/ stop words, the collocates were semi automatically tagged in terms of Part of Speech (POS). Semi-automated tagging is done to identify POS in Microsoft Excel. This is to provide a framework for future research on POS tagging in Malay.

CATEGORISATION OF NOUN COLLOCATES OF *WANITA*

This study investigates the representations of the term '*wanita*' and examines the diverse topics related to it discussed in the Malaysian Parliament. By analyzing the changes over time and the disparities between different usages of the term, the research highlights varied representations of '*wanita*' beyond the straightforward meanings of '*wanita*'. The focus is on how Members of Parliament use the lexical '*wanita*' during the specified parliamentary sessions.

To categorize the noun collocates of '*wanita*', this study employs the methodology developed by Baker, Gabrielatos, and McEnery (2013), which emphasizes contextual meanings rather than solely dictionary definitions. The analysis identifies several thematic categories: conflict, religion, culture, ethnic/national identity, character attributes, and groups/organizations. This contextual approach allows for a more subtle understanding of how the term is used in political discourse.

The thematic categorization of collocates offers important insights into the typical contexts in which '*wanita*' appears. For example, the term '*kaum*' (race) may not traditionally align with the category of ethnic or national identity; however, it can also denote a group of people. Analyzing these collocates reveals patterns that indicate semantic preferences related to '*wanita*' (Partington, 2004). To further refine the analysis, specific collocates are selected based on their frequency and mutual information (MI) score, facilitating a clearer understanding of how '*wanita*' is contextualized within parliamentary discourse. This approach not only highlights prevalent themes but also uncovers the societal influences that shape discussions about women in this context. Ultimately, this collocational analysis reveals the semantic preferences surrounding the term '*wanita*', while also illuminating broader societal attitudes and narratives. It demonstrates how language reflects and influences perceptions of women's roles and issues within the political sphere.

RESULTS

The data suggest two main findings: (1) The trends of '*wanita*' in Malaysian Parliament over time, and (2) the categorisation of '*wanita*' where women issues in Malaysian parliament can be categorised into eleven main categories namely national identity, attributes, conflict, culture, religion, group, public affairs and quantifier.

THE TRENDS OF 'WANITA' IN MALAYSIAN PARLIAMENT OVER TIME

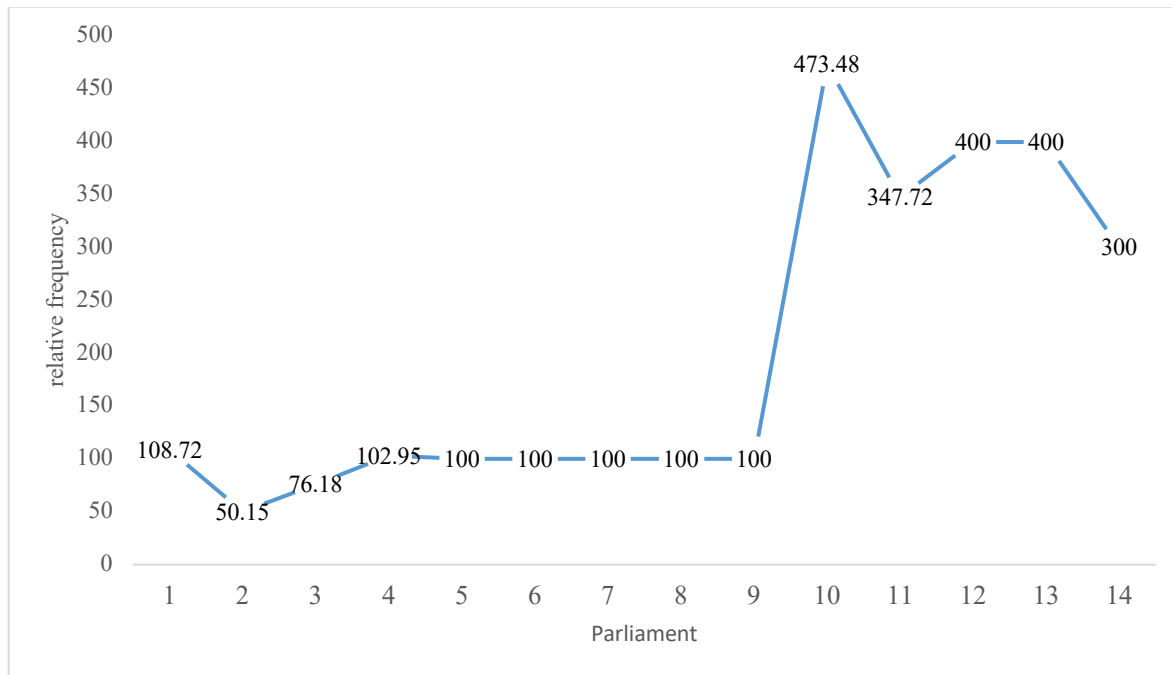


FIGURE 3. The n-gram of 'wanita' by parliament, over time in The Malaysian Hansard Corpus

Figure 3 illustrates the distribution of mentions of 'wanita' by parliament from Parliament 1 (1959) to Parliament 14 (2020). This upward trend generally indicates that the discussions on women in Malaysian Parliament have increased over the course of time. The discussions on women is gradual on Parliament 5 to Parliament 9. The mentions of 'wanita' are the lowest in Parliament 1 (year 1959 – 1964) and peak the most in Parliament 10 (year 1999-2003). The number of mentions, despite being normalised, indicates the significance of women's issues in that particular parliament. Take example from Parliament 10 where 'wanita' is mentioned the most throughout all parliaments, presumably due to the establishment of Malaysia's Ministry of Women, Family and Community Development on January 2001. It was first known as The Ministry of Women Affairs and the name was changed to The Ministry of Women and Family Development in February, of the same year. The establishment of the ministry was an execution based on the 1995 Fourth World Conference on Women in Beijing, which was a watershed moment in the global movement for gender equality. Ever since, the number of mentions of women has increased tremendously. An analysis was also done to compare the mention of 'wanita' with the exclusion of the phrases <kementerian pembangunan 'wanita' dan keluarga> (ministry of women and family) in Parliament 10, for example, indicates that out of 732 concordances on 'wanita' and *pembangunan* (development), 26 instances are related to Pelan Tindakan Pembangunan 'wanita' (National Women's Policy and Women's Development Action Plan), 154 instances are related to the ministry. The remaining concordances on *pembangunan* (development) are on the development of women in terms of national planning/ agenda. The portrayal of 'wanita' + *pembangunan* can be exemplified in Table 3, where the MPs are planning towards the development of women and single mothers in the state of Sarawak in Malaysia.

TABLE 3. The Context of “pembangunan” (development) in Parliament 10

Excerpt	Translation	Source
<i>perbincangan mengenai perancangan pembangunan khusus untuk ‘wanita’ luar bandar dan ibu tunggal di Sarawak</i>	discussion on development planning specifically for rural women and single mothers in Sarawak	P10
<i>berhubung dengan pembangunan usahawan ‘wanita’ ini, satu perkara yang saya lihat perlu dilaksanakan...</i>	in relation to the development of these women entrepreneurs, I can already see one thing that needs to be implemented	

NOUN COLLOCATES CATEGORISATION PARLIAMENT 1 TO PARLIAMENT 14

This section will further elaborate on the categories presented in Table 4. Generally, there are eight major categories of semantic preferences of ‘wanita’ in Malaysian parliamentary debates. The result can be elaborated further in the subsequent subtopics.

TABLE 4. Overall Categories of Noun Collocates of Wanita

Categories and sub-categories	Example of noun collocate
Conflict	<i>jenayah</i> (crime), <i>pembunuhan</i> (murder), <i>penderaan</i> (abuse), <i>diskriminasi</i> (discrimination), <i>keganasan</i> (violence), <i>pelacuran</i> (prostitution)
people	<i>kemiskinan</i> (poverty) taliban, <i>pendatang</i> (immigrant), <i>minoriti</i> (minority),
social problem	<i>mangsa</i> (victim), <i>penagih</i> (addicts)
religion	islam, <i>syariah</i> (sharia), <i>agama</i> (religion) <i>aurat</i> (aurat), <i>akhlak</i> (moral)
Culture	
view/ attitude/ emotion	<i>suara</i> (voice), <i>emosi</i> (emotion)
perspective	<i>pandangan</i> (view), <i>sudut</i> (angle), <i>perspektif</i> (perspective), <i>pakaian</i> (clothes), <i>tudung</i> (scarf/ hijab), <i>perkahwinan</i> (marriage)
social practices	<i>budaya</i> (culture), <i>falsafah</i> (philosophy), <i>kehidupan</i> (life)
culture	<i>martabat</i> (dignity), <i>maruah</i> (dignity)
value	<i>keldai/keldai dadah</i> (donkey/ drug dealer)
metaphor/ figurative	
Ethnic/ National Entity	
governance	<i>pengundi</i> (voters), <i>calon</i> (vandidates), <i>perwakilan</i> (representatives), <i>ketua</i> (head), <i>pemimpin</i> (leader),
governance/ legislation	<i>kementerian</i> (ministry), <i>kabinet</i> (cabinet), <i>parlimen</i> (parliament), <i>pindaan</i> (amendment), <i>akta</i> (act)
area/ location	Malaysia, Jepun <i>desa</i> (village), <i>negara</i> (country), <i>padang</i> (field), <i>ladang</i> (farm), <i>stesen</i> (station), <i>pasar</i> (market), <i>kilang</i> (factory), <i>bazar</i> (bazaar), <i>rumah</i> (house), <i>bilik</i> (room)
population	<i>manusia</i> (human), <i>orang</i> (people), <i>diri</i> (self)
Characterising/ differentiating attributes	
age/ sex	<i>bayi</i> (baby), <i>budak</i> (kid), <i>remaja</i> (teenager), <i>belia</i> (youth)
social recognition	Tuan (sir), puan (madam), Dato, Datuk, Datin,
occupational/ role	<i>guru</i> (teacher), <i>polis</i> (police), <i>tentera</i> (army), <i>doktor</i> (doctor)
ethnicity/ race/ nation	Melayu (Malay), Cina (Chinese), India (Indian)

family/ relationship	<i>suami</i> (husband), <i>isteri</i> (wife), <i>bapa</i> (father), <i>ibu</i> (mother), <i>balu</i> (widow)
Position/ status	<i>keadaan</i> (condition), <i>kedudukan</i> (position - location), <i>urusetia</i> (secretariat), <i>anggota</i> (member), <i>pasukan</i> (team)
Group/ organisation	<i>tokoh</i> (figure), <i>ikon</i> (icon)
icon	<i>bahagian</i> (division), <i>peringkat</i> (level), <i>sektor</i> (sector),
division	<i>institusi</i> (institution), <i>lembaga</i> (board), <i>organisasi</i> (organisation), <i>biro</i> (bureau), <i>kesatuan</i> (union)
Numeral indicator	
quantifier	<i>peratus</i> (percent), <i>peratusan</i> (percentage), <i>statistik</i> (statistic)
time/ day	<i>hari</i> (day), <i>malam</i> (night), <i>tahun</i> (year), <i>waktu</i> (time)
Public Affairs	<i>zaman</i> (time/ period), <i>alaf</i> (millennium)
development	<i>kemajuan</i> (progress), <i>pembangunan</i> (development)
health	<i>klinik</i> (clinic), <i>hospital</i> (hospital), <i>kesihatan</i> (health)
education	<i>kanser</i> (cancer), <i>aids</i> (Acquired immunodeficiency)
economy	<i>pendidikan</i> (education)
sports	<i>usahawan</i> (entrepreneur), <i>keusahawanan</i> (entrepreneurship),
problem solving/	<i>perniagaan</i> (business)
restoration	<i>sukan</i> (sports), <i>hoki</i> (hockey), <i>gusti</i> (wrestling)
medium/ source	<i>perlindungan</i> (protection), <i>pemulihan</i> (recovery),
management	<i>kaunseling</i> (counselling), <i>sokongan</i> (support)
potential	<i>radio</i> (radio), <i>video</i> (video), <i>iklan</i> (ads)
career-related welfare	<i>butiran</i> (details), <i>contoh</i> (example), <i>data</i> (data), <i>maklumat</i> (information), <i>perkara</i> (item), <i>pelan</i> (plan)
political sphere	<i>kelayakan</i> (qualification), <i>kemahiran</i> (skill), <i>kepimpinan</i> (leadership), <i>potensi</i> (potential)
activity	<i>caruman</i> (contribution), <i>gaji</i> (salary), <i>cuti</i> (leave), <i>pencen</i> (pension), <i>khidmat</i> (service) <i>gerakan</i> (movement), <i>kempen</i> (campaign), <i>undi</i> (vote), <i>parti</i> (party), <i>politik</i> (politics), UMNO (United Malay National Organisation)
participation/ contribution	<i>latihan</i> (training), <i>latehan</i> (old spelling for <i>latihan</i> – training), <i>aktiviti</i> (activity), <i>seminar</i> (seminar), <i>penyertaan</i> (participation), <i>penglibatan</i> (involvement), <i>kerjasama</i> (cooperation)

CONFLICT

A closer inspection of all noun collocates that fall under this category indicates the discussions on conflict. The manifestation of conflict in parliamentary discourse's context varies from discussions that involve the revelation of social problems, exploitation and violence. This category also involves the discussion on the effects of conflicts. In addition to that, it also involves the procedure for resolving difficult or complex problems that occur in society not only locally but globally. The category 'conflict' largely has or indicates negative meanings or connotations like 'crisis,' 'murder' and 'violence.'

Women are also manifested as the receiver of the conflict through the discussion on 'exploitation,' 'humiliation,' and 'victim.' This tendency indicates profound societal pressures where women frequently face the consequences of conflict-related challenges. Historical and cultural influences shape the idea of women being vulnerable in conflict scenarios, frequently resulting in their prominent participation in dialogues regarding exploitation and victimisation. Rigid gender roles designate women mainly as carers, intensifying their susceptibility in times of

conflict (Murthy & Lakshminarayana, 2006). Furthermore, societal conventions may marginalise women's voices, portraying them as passive victims instead of active contributors to peacebuilding (Harris, 2021; Smith, 2022). Thus, focused efforts are crucial for safeguarding women and enhancing their empowerment via education and economic prospects (Harris, 2021). Comprehending these factors is essential for promoting a more just methodology in conflict resolution and rehabilitation (Smith, 2022).

The following excerpts exemplify the example from the noun collocate of ‘*diskriminasi*’ (discrimination).

TABLE 5. Conflict category – ‘*diskriminasi*’ (discrimination)

Excerpt	English Translation	Source
<i>Apakah program kerajaan untuk memartabatkan ‘wanita’ dalam negara kita. Jadi begitu juga berkaitan dengan diskriminasi wanita. Ada ‘wanita’ yang mereka dilarang memakai pakaian menutup aurat semasa kerja kerana mematuhi arahan syarikat.</i>	What is the government's program to uplift women in our country. It involves discrimination against women . There are women whom are prohibited from wearing clothes covering ‘aurah’ during work to comply to company’s instructions.	P14
<i>Dia daripada perspektif pakaian ‘wanita’ Islam, contohnya kalau ‘wanita’ Islam itu tidak boleh, tidak dibenarkan menutup aurat, maka itu diskriminasi. Dia seharusnya dibenarkan menutup aurat.</i>	He sees the perspective of Muslim women ’s clothing, for example if the Muslim woman cannot, is prohibited to cover her ‘aurah,’ and that is discrimination . She needs to cover her ‘aurah’.	

In addition, another example in the conflict category is ‘*pelacuran*’ (prostitution), as illustrated in Table 6. This table highlights the importance of addressing the vulnerability of young women, who are at high risk for temporary homelessness and involvement with illicit prostitution syndicates. The excerpts raise critical questions about the ongoing prevalence of this trade. Both citizens and non-citizens, including women, children, and men, are often targets of exploitation, commodified in ways that compromise their dignity and autonomy. Addressing these issues requires targeted measures to protect vulnerable groups and a comprehensive understanding of the socioeconomic factors that contribute to this problem.

TABLE 6. Conflict category – ‘*pelacuran*’ (prostitution)

Excerpt	English Translation	Source
<i>Mohon nyatakan berapa peratusan kah golongan ‘wanita’ muda-Yang Berhormat Menteri sebut golongan muda tetapi saya nak minta spesifik ‘wanita’ muda kerana mereka high-risk group untuk terdedah kepada sindiket-sindiket pelacuran haram dan gelandangan bertaraf sementara...</i>	Please state the percentage of young women -The Honorable Minister mentions young women but I want to ask specifically for young women because they are a high-risk group to be exposed to illegal prostitution syndicates and temporary homelessness...	P14
<i>Mengapa walaupun serbuan di atas serbuan dilaksanakan tetapi masih ada yang mahukan trade pelacuran ini masih berjalan. ‘wanita’ dan kanak-kanak, lelaki dan perempuan warganegara dan bukan warganegara turut menjadi mangsa eksploitasi sebagai komoditi menjual maruah dan tubuh badan</i>	Even though various of raids were carried out but there are people who want this prostitution trade. Why? Women and children, men and women of citizens and non-citizens are also victims of exploitation as a commodity to sell dignity and body	P13

RELIGION

The collocates under this category are concerned about religious matters related to Islam. As Malaysia’s official religion is Islam, the collocates that fall under this category typically are related to Islam and the Islamic way of life, including moral values, Islamic law and the appropriate way of dressing up. All noun collocates under this category imply religious belief related to Islamic practices like *Islam*, *Sharia* (law of Islam), *religion*, and *auras* (any parts of the body that are prohibited from being exposed to other people). The example can be seen in Table 7.

TABLE 7. Religion category –Allah

Excerpt	English Translation	Source
<i>Kami kaum ‘wanita’ di-negeri ini menguchap terima kaseh kapada Kerajaan dan menguchap shukor ka-hadrat Allah atas kejayaan2 Kerajaan dalam mema-jukan negeri ini</i>	Us women in this state would like thank the Government and thank Allah (God) for the successes of the Government in developing this state.	P14

CULTURE

All nouns that fall under this category can be seen as a set of behaviours, beliefs, perspectives, practices and values generally shared by Malaysians but encompass women. This includes the voice and emotion of women, as well as the appearance of women in terms of outerwear. Interestingly, women are also related to metaphors, including *keldai dadah* (literal translation, drug donkey), which means drug dealer. Women are also related with a hand (*tangan*) as a metaphor where women are associated as the receiver of help. Table 8 shows the use of culture category about women’s issues in Malaysian parliamentary discourse to indicate the dignity of Malaysian women.

TABLE 8. Culture category – *maruah* (dignity)

Excerpt	English Translation	Source
<i>perkara ini dihalang sebab ‘wanita’ ada maruah. Jadi, kita bukan menjadi alat dagangan. Jadi, seharusnya ada satu undang-undang yang menyekat supaya ‘wanita’ di Malaysia adalah ‘wanita’ yang bermaruah. Kerajaan bertanggungjawab untuk memartabatkan wanita, bukan menjadi alat dagangan bagi pihak peniaga.</i>	this is forbidden because women have dignity . So, we are not a trading tool. So, there should be a law that prohibits women in Malaysia from being dignified women. The government is responsible for dignifying women, not being a tool of trade on behalf of traders.	P14

ETHNIC/ NATIONAL ENTITY

This category comprises several sub-categories and is not exclusively related to ethnicity but the diversity of women as humans or entities tied to national entities. As opposed to race, which is more frequently seen as inherent in our genetics and consequently passed down from generation to generation, an individual often decides on ethnicity. The idea of ethnicity has the advantage of allowing people to adopt many identities because it encompasses everything from language to ethnicity, culture, and religion. This category consists of governance, legislation’s governance, area or location, and population. It includes words related to elections like *calon* (candidate) and *pengundi* (voters), which elucidate women’s participations in politics. Women are also involved

in governance through the association with the words like *kementerian* (ministry) and *kabinet* (cabinet). Issues on women are significantly discussed towards betterment and change through the use of words like *pindaan* (amendment), *rang* (bill) and *dasar/ polisi* (policy). Other than that, legislative associations for women include acts, policies, and amendments to women’s issues. In addition to that, women under this category are identified under the population, including *manusia* (human being), *orang* (person) and *diri* (self). Table 9 indicates the example from P14 .

This categorization reflects a comprehensive understanding of women's roles within the socio-political landscape. By linking women to governance and legislative processes, the discourse emphasizes their active participation and the recognition of their rights. Using terms related to politics and legislative change suggests a growing acknowledgement of the need for gender-sensitive policies and the importance of women's voices in shaping national agendas. Including broader identity terms also illustrates the multifaceted nature of women's experiences, allowing for a more inclusive approach to understanding their societal contributions and challenges.

TABLE 9. Ethnic/ national entity category – *polisi* (policies)

Excerpt	Translation	Source
<i>Kajian ini mempunyai empat output dan sekarang mereka sedang mengkaji keberkesanan semua polisi yang telah dilaksanakan untuk golongan wanita. Apabila kajian ini siap nanti kami tentu akan berkongsi dengan semua Ahli Yang Berhormat.</i>	The study has four outputs and now they are reviewing the effectiveness of all the policies that have been implemented for women . When this study is completed, we will surely share it with all the Honourable Members.	P14

CHARACTERISING/ DIFFERENTIATING ATTRIBUTES

Generally, this category is related to the attributes of women, including women as a gender, which is also associated with womanliness, like ‘*gadis*’ (girl) and ‘*lelaki*’ (male). This category is also closely related to ethnic and cultural categories. Other than that, it is apparent that women are associated with occupations or roles like teacher, police, and doctor. It could be seen that women's occupations have revolutionised over time, from being teachers, police, and army to being in parliament as an MP, even as an employer and activist. Other than being linked to ethnicities like Malay, Chinese, Indian and Penan, women also play social roles as wives and mothers and are related to husbands and fathers. Women are also associated with ‘*balu*’ (widow) under this category. In addition, women’s association under this category is also linked to a woman’s position and status like ‘*jawatan*’ (post), status, and ‘*kuota*’ (quota). The example of a woman can be portrayed in Table 10.

TABLE 10. Characterising/ differentiating attributes category – *isteri* (wife)

Excerpt	English Translation	Source
<i>isteri yang malang tadi terpaksa turun dari rumah kongsi hidup-nya dengan keadaan baju yang ada lekat di-badan-nya sahaja. Maka di-sini-lah kami kaum ‘wanita’ memerlukan satu Undang2 yang membenarkan harta yang di-chari atau kumpulan sa-masa hidup bersama2 itu di-bahagikan sama rata di-hari pencheraian itu</i>	the unfortunate wife had to come down from her shared house with the condition of the shirt sticking to her body only. So this is where us as women need a law that allows the property that is taken, or the group while living together to be divided equally on the day of the divorce.	P2

<i>Yang Berhormat Menteri, saya bangkit isu waktu kerja yang agak flexible untuk 'wanita' balik lebih awal sebagai menghormati hak mereka sebagai isteri.</i>	Honorable Minister, I raised the issue of relatively flexible working hours for women to return home from work earlier as a respect for their rights as wives .	P14
---	---	-----

Based on this example, women are associated with the word 'isteri' (wife) as a group of people. The discussion that can be seen from P2 to P14 is interchangeable, focusing on the right to protect women. This categorization highlights the diverse roles women play in society, showing both traditional and evolving identities. The inclusion of women in various occupations indicates progress toward gender equality. By connecting women to ethnic identities and social roles, the discourse reflects the complexity of their experiences. Additionally, mentioning status and quotas emphasizes ongoing discussions about women's representation and the need for continued advocacy for equal opportunities.

GROUP/ ORGANISATION

This category involves women as a group of people who participate in the society. This category also corresponds with the previously mentioned category. Nevertheless, the collocates clustered under this category have their own direction and objectives, which are more towards active organisational participation. The example can be seen in the collocates like 'profesional' (professional), 'korporat' (corporate) and 'anggota' (member of group). These examples are not attributed to the culture but more to professionalism. Other than that, women have also become icons or respected figures. This category highlights the recognition of women as active participants in professional and corporate environments. The focus on terms like 'profesional', 'korporat', and 'anggota' emphasizes their roles in organizations rather than cultural attributes. By portraying women as icons and respected figures, the discourse acknowledges their growing influence and leadership in society, indicating a positive trend toward gender equality and empowerment in various fields.

TABLE 11. Characterising/ differentiating attributes category – anggota (officer)

Excerpt	English Translation	Source
<i>Ini ia-lah untok persaim-bangan yang selalu sahaja di-sebut2 atau di-kehendaki oleh gulongan kaum 'wanita' di-dalam negeri kita ini dan saya fikir sangat-lah patut kita menam-bah anggota polis 'wanita' di-negeri ini sesuai dengan kes2</i>	This is for the purpose that is always mentioned or desired by the women in our state and I think it is very appropriate that we add more women police officers in this state in accordance with the case.	P2

NUMERAL INDICATOR

Numeral indicators involve quantifiers such as statistics related to women. It also involves the date and time like 'hari' (day), 'malam' (night) and 'zaman' (period). An example of the numeral indicator can be depicted in Table 12. Based on the example, the numeral indicator related to women also involves discussions to solve women's issues.

Overall, this category focuses on temporal references, highlighting the various timeframes in which women engage in work and social responsibilities. This emphasis reflects the demands placed on women as they balance professional roles with familial duties. Recognizing women's contributions at different times of the day and across various stages of life is crucial, as it calls for

the development of supportive structures that accommodate these multifaceted roles and empower women in society.

TABLE 12. Numeral indicator category – *malam/ petang* (night/ evening)

Excerpt	English Translation	Source
<i>kelas yang ada kita buka dalam pelajaran dewasa ini ia-lah kelas yang bersesuaian dengan masa orang2 di-dalam kampong itu. Jadi kalau-lah Penyelia ‘wanita’ ini berjalan waktu malam, waktu petang hari so-orang diri, tentu-lah memberikan bahaya sa-mata2 kepada mereka berjalan di-kampong2.</i>	The class that we have open in today's lesson is the class that is suitable for the time of the people in the village. So if this woman Supervisor walked at night , in the evening alone, it would certainly give them a mere danger to walk in the village.	P2

PUBLIC AFFAIRS

This category is the emerging category compared to the study’s framework. This category comprises four sub-categories: development, health, education, and economy. These are the main public domains of the issues debated in the parliament regarding women. This category also comprises women's association with sport, management, resources, problem-solving, and recovery. It also associates women with potential and welfare, activities and contributions of women. Table 12 exemplifies the example of women’s association under the public affairs category.

TABLE 12. Public affairs category – *kesihatan* (health)

Excerpt	English Translation	Source
<i>Bagi kaum ‘wanita’ pula, sa-lain daripada latehan2 yang di-sebutkan tadi, mereka juga akan di-beri latehan: (a) Urusan Rumah Tangga. (b) Kesihatan diri dan kesihatan keluarga</i>	For women , in addition to the trainings mentioned earlier, they will also be given other trainings: (a) Household Affairs. (b) Personal health and family health	P2
<i>bagi menangani kesedaran ‘wanita’ untuk hal kesihatan mental ini yang berisiko antaranya ‘wanita’ selepas bersalin, ‘wanita’ yang pernah menghadapi trauma, ‘wanita’ yang mengurus keluarga secara sendirian dan yang mengalami buli siber.</i>	to address women's awareness of this mental health issue that is at risk including women after childbirth, women who have experienced trauma, women who manage their families alone and who experience cyberbullying.	P14

Based on the extract, it can be seen that women's association with health, for example, is not solely related to physical and mental health. However, discussion in earlier parliament associates women’s health with training towards the goodness of their family.

This emerging category highlights the critical focus areas in parliamentary discussions about women, reflecting a broader recognition of their roles in essential public domains. By encompassing development, health, education, and economy, the discourse emphasizes the multifaceted contributions of women to society. Including associations with sport and management indicates a growing acknowledgement of women's capabilities beyond traditional roles. This comprehensive approach underscores the importance of addressing women's issues holistically,

recognizing their potential for positive impact in various sectors and advocating for policies that support their well-being and empowerment.

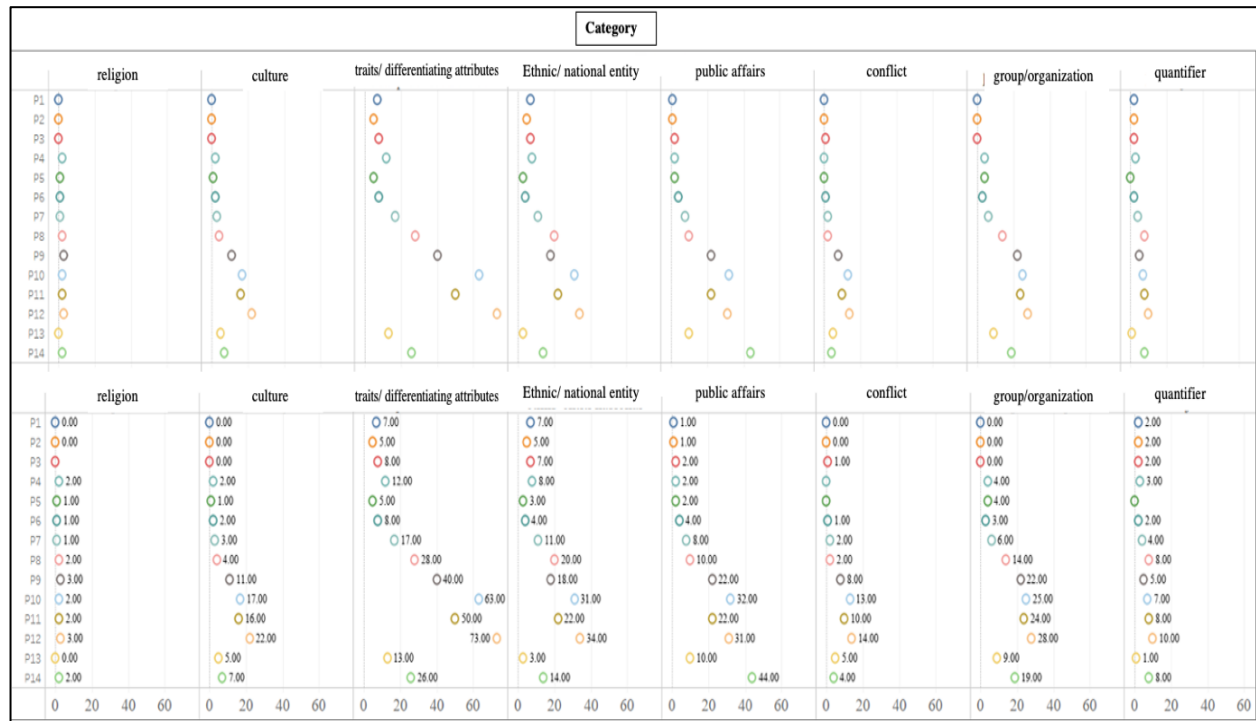


FIGURE 5. The Distribution of Categories related to Women in Malaysian Parliament from 1959-2020

Figure 5 depicts the distribution of the categorization of noun collocates of 'wanita' from Parliament 1 to Parliament 14 inside the Malaysian Hansard Corpus. This figure is illuminating in multiple aspects. Initially, it emphasizes the comparative prevalence of topic categories about women in the Malaysian Parliament from 1959 to 2020.

The selected categories comprise conflict, religion, culture, ethnic/national identity, character traits, and groups/organizations. Discourse concerning "conflict" may have significantly escalated during social upheaval, whilst "religion" might have gained prominence during talks on regulations impacting religious practices. The categories of "culture" and "character attributes" exhibit a significant decline in frequency during Parliament 13 and Parliament 14. This drop may signify a change in emphasis towards more urgent matters or a reaction to changing cultural objectives. This indicates that conversations regarding women's responsibilities are increasingly contextualized within modern concerns rather than conventional attributes.

The frequency of each category continually varies, signifying that the political landscape and objectives of Malaysian MPs adapt to shifting social dynamics and global trends. This variation exemplifies the representation of women's problems in parliamentary discourse and indicates broader cultural transformations about women's roles and societal expectations over the decades.

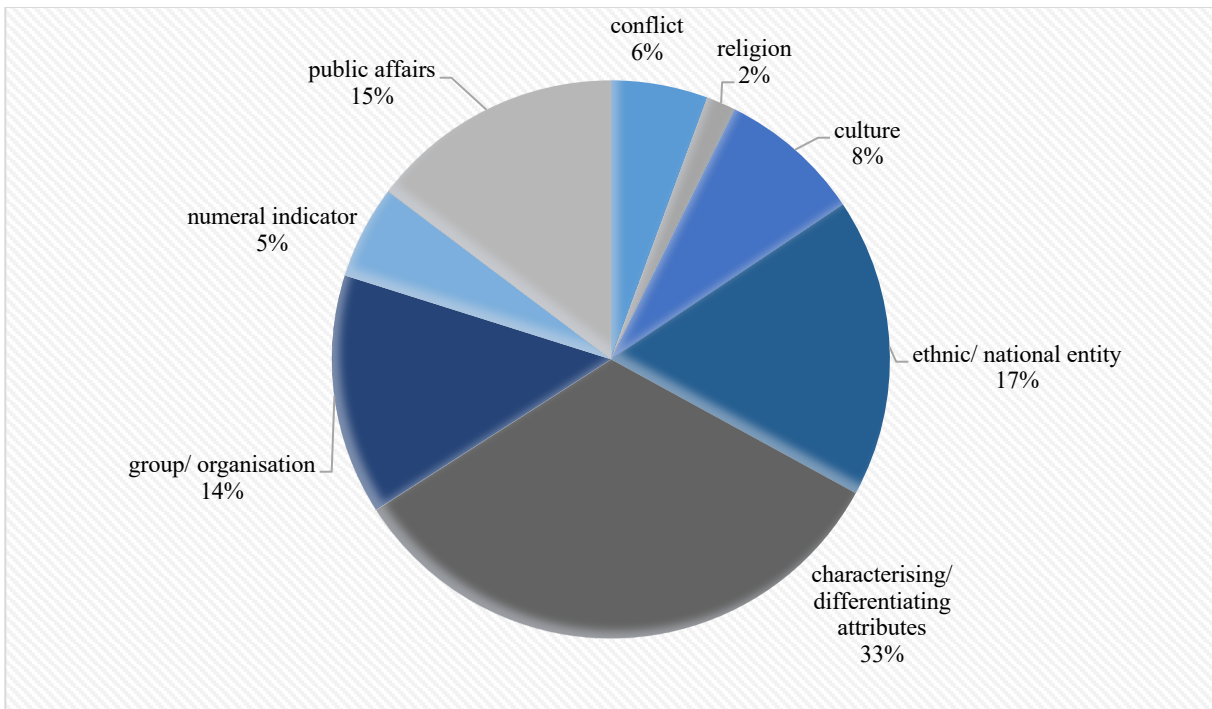


FIGURE 6. Frequency distribution of collocate category in terms of token

The collective frequency of each category from Parliament 1 to Parliament 14 is illustrated in Figure 6. The category "characterizing/differentiating attributes" has the highest occurrence in the pie chart, suggesting a substantial emphasis on the characteristics and qualities associated with women in parliamentary discourse. A more detailed examination of the chart reveals this. Conversely, the discourse regarding women as a "group/organization" is less prevalent than when they are framed as an "ethnic/national entity."

This discrepancy implies that, despite acknowledging women's collective identity, the discourse prioritises individual characteristics and attributes. The lexical richness of this area is emphasized by the prominence of the "characterizing/differentiating attributes" category, which suggests that the discussions surrounding these attributes are substantial and multi-dimensional. These discussions may cover a wide range of topics, including the representation of women in various contexts, personal qualities, and societal responsibilities.

The results depicted in Figure 6 underscore the intricacies of the discourse surrounding women in the Malaysian Parliament. The high frequency of characterising attributes indicates that these discussions are prevalent and reflect underlying societal attitudes towards women. The importance of comprehending how such representations influence perceptions and policies regarding women is underscored by the emphasis on traits and characteristics. To guarantee that women's concerns are adequately addressed in ongoing political discourse, it will be imperative to investigate the implications of these categories on legislative agendas and societal perceptions as this study progresses.

DISCUSSION AND CONCLUSIONS

The methodology utilised in this study to analyse the depiction of women in political discourse exhibits considerable value and flexibility. This approach to various node words inside the Malaysian Hansard Corpus facilitates the classification of collocates pertinent to women's problems. From 1959 to 2020, this investigation provides significant insights into the engagement and interpretation of women's problems by Malaysian Members of Parliament (MPs). The results validate the theoretical underpinnings of the corpus-driven methodology, especially on women's issues in Malaysia's parliament. The tendency in legislative debates has intensified, mirroring changes in political issues related to both global and local contexts. The diverse issues discussed in various parliamentary sessions indicate that the political agenda is fluid and attuned to prevailing circumstances. This dynamic is also reflected in the UK, where a top-down methodology allows parliamentary committees to establish agendas, guaranteeing thorough consideration of diverse subjects (Warren & Pearse, 2008; Farrell, Suiter, & Harris, 2019; Fein, Goethals, & Kugler, 2007). Improving democratic processes in agenda development can address the concerns of members desiring increased input in discussions.

While the study clarifies reoccurring preferences regarding '*wanita*' in the Malaysian Parliament, conclusive insights into their perception by other MPs or citizens remain unattainable. The perception of attitudes is contingent upon context and shaped by multiple elements, confounding the classification of these stimuli (Fein, Goethals, & Kugler, 2007). Nonetheless, the findings reveal that women are progressively portrayed in various categories during parliamentary sessions, indicating a change in opinions among Malaysian MPs. Significantly, discussions regarding women have been persistently addressed during six decades of parliamentary sessions. This signifies that MPs are examining and addressing issues concerning women, indicating a shift towards more constructive and sophisticated discourse. This pattern indicates an increasing recognition of women's issues inside the Malaysian Parliament, potentially reflecting a change in cultural attitudes. Future studies should investigate whether this transition signifies authentic societal advancement or a political discourse byproduct.

Future studies could examine the prosody of Malaysian MPs while addressing women's issues in Parliament, potentially uncovering the underlying attitudes and intents that influence legislation priorities. Moreover, additional investigation is required to understand how politicians portray diverse situations to sway legislation, especially in reaction to pressing matters such as the COVID-19 pandemic and climate change. The results of this study have substantial significance for future practice. Researchers might enhance their comprehension of women's political difficulties using analogous approaches. This methodology enhances academic dialogue and educates policymakers on the progression of talks about women's issues throughout 61 years of legislative sessions. This research highlights the necessity of critically evaluating political narratives and their influence on social perceptions. It stresses the importance of continuous dialogue and action to ensure that legislative agendas sufficiently address and prioritise women's issues.

REFERENCES

- Abdullah, K., Noor, N. M., & Wok, S. (2008). The perceptions of women's roles and progress: A study of Malay women. *Social Indicators Research*, 89, 439–455. <https://doi.org/10.1007/s11205-008-9242-7>
- Badri, B. (2020). The impact of COVID-19 on women. *UN Chronicle*. <https://www.un.org/en/un-chronicle/impact-covid-19-women>
- Baker, H., Brezina, V., & McEnery, T. (2017). Ireland in British parliamentary debates 1803–2005. pp. 83–107.
- Baker, P., Gabrielatos, C., Khosravinik, M., Krzyzanowski, M., McEnery, T., & Wodak, R. A. (2008). Useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse and Society*, 19, 273–306. <https://doi.org/10.1177/0957926508088962>
- Baker, P., Gabrielatos, C., & McEnery, T. (2013). Sketching Muslims: A corpus driven analysis of representations around the word “Muslim” in the British press 1998-2009. *Applied Linguistics*, 34, 255–278. <https://doi.org/10.1093/applin/ams048>
- Begagić, M. (2020). “Utterly content in each other’s company”: Semantic prosody and semantic preference. *Journal of Language and Linguistic Studies*, 16(1), 1-15.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.
- Brezina, V., McEnery, T., & Wattam, S. (2015). Collocations in context. *International Journal of Corpus Linguistics*, 20, 139–173. <https://doi.org/10.1075/ijcl.20.2.01bre>
- Cotteril, J. (2001). Domestic discord, rocky relationships: Semantic prosodies in representations of marital violence in the O.J. Simpson trial. *Discourse & Society*, 12, 291–312.
- Conceição, P. (2019). *United Nations Development Programme Human Development Report 2019: Beyond income, beyond averages, beyond today: Inequalities in human development in the 21st century*; ISBN 9789211264395.
- Dabbagh, A. (2016). Gender representation under critical image analysis: The case of Iranian ELT textbooks evaluation of learning objectives using Bloom’s revised taxonomy: Comparing vision and four corners view project.
- Dhatt, R., Theobald, S., Buzuzi, S., Ros, B., Vong, S., Muraya, K., Molyneux, S., Hawkins, K., González-Beiras, C., & Ronsin, K. (2017). The role of women’s leadership and gender equity in leadership and health system strengthening. *Global Health, Epidemiology and Genomics*, 2. <https://doi.org/10.1017/gheg.2016.22>
- Farrell, D. M., Suiter, J., & Harris, C. (2019). ‘Systematizing’ constitutional deliberation: The 2016–18 citizens’ assembly in Ireland. *Irish Political Studies*, 34, 113–123.
- Fein, S., Goethals, G. R., & Kugler, M. B. (2007). Social influence on political judgments: The case of presidential debates.
- Hajar Abdul Rahim. (2005). Impak konotasi budaya terhadap leksis: Satu kajian semantik berasaskan korpus, ke atas perkataan perempuan dan wanita. *Jurnal Bahasa*, 5, 83–111.
- Haas, W. J. R. Firth. (1958). *Papers in Linguistics, 1934–1951*. London, Etc.: Oxford University Press, 1957. 35s. *Bulletin of the School of Oriental and African Studies*, 21, 668–671. <https://doi.org/10.1017/S0041977X00060559>
- Hishamudin Isam, N. M. A. Setia (Loyal) from semantics prosody perspective: A corpus-based analysis. *GEMA Online® Journal of Language Studies*, 12, 359–374.

- Hunston, S., & Francis, G. (n.d.). *Pattern grammar: A corpus-driven approach to the lexical grammar of English*.
- Imran Ho Abdullah, Anis Nadiah Che Abdul Rahman & Azhar Jaludin. (2021). The development of the Malaysian Hansard corpus: A corpus of parliamentary debates 1959-2020. *Jurnal Linguistik*, 25.
- Jamaluddin Aziz (2019). Exploring gender issues associated with wanita/woman and perempuan/woman in Malaysian Parliamentary Debates: A culturomic approach. *GEMA Online® Journal of Language Studies*, 19(4), 278-303.
- Jia, S., Lansdall-Welfare, T., Sudhahar, S., Carter, C., & Cristianini, N. (2016). Women are seen more than heard in online newspapers. *PLoS ONE*, 11. <https://doi.org/10.1371/journal.pone.0148434>
- Joshi, A., Neely, B., Emrich, C., Griffiths, D., & George, G. (2015). Gender research in AMJ: An overview of five decades of empirical research and calls to action. *Academy of Management Journal*, 58, 1459–1475.
- Kok, A. (2020). IPSOS study reveals Malaysians’ top issues faced by women and misperceptions of women empowerment.
- Liu, W. (2020). A corpus-based study of semantic prosody and semantic preference of “seem.” *Open Journal of Social Sciences*, 8(1), 43-56.
- Louw, B., & Chateau, C. (2000). *Semantic prosody for the 21st century: Are prosodies smoothed in academic contexts? A contextual prosodic theoretical perspective*.
- Michel, J.-B., Shadish, W. R., & Teytelman, L. (2011). Quantitative analysis of culture using millions of digitized books. *Science*, 331(6014), 176–182. <https://doi.org/10.1126/science.1199644>
- Milesi, P., & Alberici, A. I. (2019). “Frontrunners”: An investigation of the discursive construction of “women politicians” intersectional identity. *Europe’s Journal of Psychology*, 15(3), 459–478. <https://doi.org/10.5964/ejop.v15i3.1557>
- Morley, J., & Partington, A. (2009). A few frequently asked questions about semantic — or evaluative — prosody. *International Journal of Corpus Linguistics*, 14, 139–158. <https://doi.org/10.1075/ijcl.14.2.01mor>
- Ndambuki, J., & Janks, H. (2010). Political discourses, women’s voices: Mismatches in representation. 4, 73–92.
- Norsimah Mat Awal Azhar Jaludin, Anis Nadiah Che Abdul Rahman and Imran Ho-Abdullah, I. (2019). ‘Is Selangor in Deep Water?’: A Corpus-driven Account of Air/water in the Malaysian Hansard Corpus (MHC). *GEMA Online® Journal of Language Studies*, 19(2), 99–120. Doi: <https://doi.org/10.17576/gema-2019-1902-07>
- Nor Fariza Mohd Nor, Anis Nadiah Che Abdul Rahman, Azhar Jaluddin, Imran Ho Abdullah, and Sabrina (2019). A Corpus Driven Analysis of Representations Around the Word ‘ekonomi’ in Malaysian Hansard Corpus. *GEMA Online® Journal of Language Studies*, 19(4), pp.66–95. Doi: <https://doi.org/10.17576/gema-2019-1904-04>
- Partington, A. (1998). *Patterns and meanings: Using corpora for English language research and teaching*. J. Benjamins Pub. ISBN 9027222703.
- Partington, A. (2004). “Utterly content in each other’s company” semantic prosody and semantic preference.
- Pearce, M. (2008). Investigating the collocational behaviour of man and woman in the BNC using Sketch Engine. *Corpora*, 3, 1–29. <https://doi.org/10.3366/e174950320800004x>

- Ponterotto, D. (2014). Trivializing the female body: A cross-cultural analysis of the representation of women in sports journalism.
- Pey Jung Yeong. (2018). How women matter - gender representation in Malaysia's 14th general election. *The Round Table*, 107, 771–786.
- Rheault, L., Beelen, K., Cochrane, C., & Hirst, G. (2016). Measuring emotion in parliamentary debates with automated textual analysis. *PLoS ONE*, 11. <https://doi.org/10.1371/journal.pone.0168843>
- Sabrina Tiun, Nor Fariza Mohd Nor, Azhar Jalaludin and Anis Nadiah Che Abdul Rahman. (2019). Word Embedding for Small and Domain-specific Malay Corpus. *Lecture Notes in Electrical Engineering*, 603), 435–443. Doi: https://doi.org/10.1007/978-981-15-0058-9_42
- Schwab, K., Crotti, R., Geiger, T., & Ratcheva, V. (2020). *World Economic Forum. Global gender gap report 2020 insight report*; World Economic Forum, 2019. ISBN 9782940631032.
- Sculley, D., & McCallum, A. (2003). *Web-scale classification and new data mining algorithms*. In Proceedings of the 25th International Conference on Machine Learning (ICML).
- Selmistraitis, L. (2020). Semantic preference, prosody and distribution of synonymous adjectives in COCA. *GEMA Online® Journal of Language Studies*, 20(1), 67-80.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
- Sinclair, J. (1996). The empty lexicon. *International Journal of Corpus Linguistics*, 1, 99–119. <https://doi.org/10.1075/ijcl.1.1.07sin>
- Sinclair, J. (2003). *Reading-concordances-an-introduction*. Pearson Longman.
- Stubbs, M. (1995). Collocations and semantic profiles. *Functions of Language*, 2, 23–55. <https://doi.org/10.1075/foL.2.1.03stu>
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work* (6th ed.). John Benjamins Publishing Company.
- UN Women. (2021). *Women in politics: 2021*. Digital library: Publications. Retrieved from <https://www.unwomen.org/en/digital-library/publications/2021/01/women-in-politics-2021>
- van Dijk, T. A. (2004). On the analysis of parliamentary debates on immigration.
- Wan-Ibrahim, W. A., & I. Zainab. (2014). Factors Related to Religious Commitment of Rural Older Woman in Malaysia. *Middle East Journal of Scientific Research*, 20(12), 1746–1749. <https://doi.org/10.5829/idosi.mejsr.2014.20.12.21079>
- Warren, E. M., & Pearse, H. (2008). Designing deliberative democracy: The British Columbia citizen's assembly. In *Cambridge University Press*.
- Wei, N., & Li, X. (2020). Exploring semantic preference and semantic prosody across English and Chinese: Their roles for cross-linguistic equivalence. *International Journal of Linguistics*, 12(1), 72-85.
- Willis, R. (2017). Taming the climate? Corpus analysis of politicians' speech on climate change. *Environmental Politics*, 26, 212231. <https://doi.org/10.1080/09644016.2016.1274504>

ABOUT THE AUTHOR

Anis Nadiah Che Abdul Rahman is a senior lecturer at the Center for Research in Language and Linguistics. Her areas of interests include corpus linguistics and semantics, focusing on political and social media discourse.