

## “Is Selangor in Deep Water?”: A Corpus-driven Account of *Air*/water in the Malaysian Hansard Corpus (MHC)

Norsimah Mat Awal

[norsimah@ukm.edu.my](mailto:norsimah@ukm.edu.my)

Center for Literacy and Socio Cultural Transformation,  
Faculty of Social Sciences and Humanities,  
Universiti Kebangsaan Malaysia

Azhar Jaludin

[azharj@ukm.edu.my](mailto:azharj@ukm.edu.my)

Center for Literacy and Socio Cultural Transformation,  
Faculty of Social Sciences and Humanities,  
Universiti Kebangsaan Malaysia

Anis Nadiah Che Abdul Rahman

[p87706@siswa.ukm.edu.my](mailto:p87706@siswa.ukm.edu.my)

Center for Literacy and Socio Cultural Transformation,  
Faculty of Social Sciences and Humanities,  
Universiti Kebangsaan Malaysia

Imran Ho Abdullah

[imranho@ukm.edu.my](mailto:imranho@ukm.edu.my)

Center for Literacy and Socio Cultural Transformation,  
Faculty of Social Sciences and Humanities,  
Universiti Kebangsaan Malaysia

### ABSTRACT

*Air* or its English equivalent ‘water’ is very important in our everyday life so much so that when the tap runs dry, it even made it as one of the topics debated by politicians. This paper looks at the issues that surround *air*/water in Malaysian Parliamentary debates by specifically focusing its relation to the state of Selangor. The *air*/water related issues were examined based on the collocates of *air* and *Selangor* in the Malaysian Hansard Corpus (MHC) from Parliament 1 (P1) to Parliament 13 (P13). The findings show that *air* is consistently present as one of the collocates of *Selangor* from Parliament 4 (P4) to Parliament 13 (P13). However, *air* started to show an upward trend starting in Parliament 7 (P7) and continued to Parliament 13 (P13). The recurring issues during those periods are the never-ending water-related problems and the steps taken by the government to overcome the problems. In P7 and P8, the focus is on the source of water as it collocates with *pembersihan logi air* (water treatment plant) and *kawasan tadahan air* (water catchment area). In Parliament 10, Parliament 11, Parliament 12 and Parliament 13 the recurring issue with *air* and *Selangor* is *penyaluran air mentah* (the transfer of raw water) from the neighbouring state Pahang to Selangor. Another issue observed is *penstrukturan air* (restructuring) of water supply and services which was first observed in Parliament 12 and continues to Parliament 13. Thus, by focusing on the collocates of *air* this corpus-driven account has managed to show the trend of the parliamentary debates in relation to *air* and *Selangor*. Therefore, parliamentary debates where various issues of national interest are often raised offer opportunities for more critical analysis of issues that are important to the public.

**Keywords:** *air* (water); Selangor; Malaysian Hansard Corpus; parliamentary debates & collocates

## INTRODUCTION

Water is one of the most precious resources, livelihoods and prosperity to human and other life forms. Without water, all forms of life will not survive and would cease to exist. It is an essential element in maintaining life and needed for the survival of all organisms. It is also a critical ingredient in all social and economic endeavours, such as agriculture, industry, energy, transport. Water plays an equally powerful role through its wide impact on the factors such as food production, hygiene, sanitation and health and the environment (Chamhuri & Ferdoushi, 2014; Wen, 2004; Malek et al., 2013). Hence, it can be said that should any country be hit by water crisis, the country would also be facing crisis in humanity, economy and health. Malaysia has had its share of water crisis in 2014, where Selangor, Federal Territory of Kuala Lumpur (FTKL) and Putrajaya were badly effected and water rationing had to be implemented (Syed Muhammad Shahabudin, 2015; Rasyidah Md Khalid, 2018). Since water plays a very important and powerful role in the economic growth and sustainable development of a country, it would be interesting to examine issues pertaining to water are discussed and debated by Malaysian law makers in the parliament.

## PARLIAMENT

Parliament is the legislative body of a government. Power (2012) described parliament as:

Parliaments are the indispensable institutions of representative democracies around the world. Whatever their country-specific rules, their role remains the same: to represent the people and ensure that public policy is informed by the citizens on whose lives they impact...Parliament is therefore the central institution of democracy and constitutes an expression of the very sovereignty of each nation. Parliament is a political institution. It is a place for political, and often confrontational, debate.

(Power, 2012, pp. 1-2)

In Malaysia, Parliament is the highest legislative body of the country and it is the place where laws are being enacted. *Dewan Rakyat* or the House of the Representatives is one of the three components in the Parliament of Malaysia, the premier legislative body in Malaysia after the *Dewan Negara* (Senate) and His Majesty The King. The House of Representatives has 222 members that are elected from the General Election of Malaysia that is held in every five years. Each Member of Parliament (MP) represents a Parliamentary Constituency (The Official Portal of Parliament of Malaysia, 2018). The role of the House of Representative stands as the body to pass, emend and repeal the laws of the country. The parliamentary debates of House of Representative and Senates are all recorded in an official record called Hansard.

## HANSARD

Hansard is an official parliament report that contains accurate and verbatim debates during the *Dewan Rakyat* (House of Representatives) and the *Dewan Negara* (Senate) sittings. The publication of this report gives the public an opportunity to follow and understand all the debates involving members of both houses. The public will also be able to understand and made aware of the processes and proceedings that occur in parliament without having to attend the sessions. The term 'Hansard' is actually taken after the name of Thomas C. Hansard, the publisher of the debates in British Parliament in the early 19th century (Vice & Farrell, 2017). The Malaysian Parliament Hansard contains the Yang di-Pertuan Agong's speech, the attendance of the Members of Parliament (MPs), the questions and answers of the parliament sittings, bills and acts tabled, MPs' votes and others. It is important to note that

the Members of Parliament are free to voice out any views during the session without fear of being subjected to legal action as any statements made in Parliament cannot be prosecuted in any court.

Malaysian hansard in its current state (in isolated PDF forms) cannot be fully utilized for in-depth research. One of the ways to harness the data in the hansard is to turn the PDF form of the Hansard to a digital corpus form. The construction of hansard corpus has already taken place in other parts of the world, such as researchers in Britain, Canada and the European countries. They have been building corpus of Hansard Parliament of their respective countries such as the British Hansard corpus (or collection of texts) which contains nearly every speech given in the British Parliament from 1803-2005 (<https://www.hansardcorpus.org/>); The CLARIN ERIC infrastructure, an internet service organization, has listed and provided access to 18 European parliamentary corpora, covering almost all of the languages spoken in European countries (CLARIN, 2019). CLARIN ERIC also highlighted the importance of Parliamentary corpora:

Parliamentary corpora are a very important multidisciplinary language resource that can be approached from many research perspectives, including not only political science, but also sociology, history, psychology, and applicative approaches to linguistics, for instance, critical discourse analysis. (CLARIN, 2019)

Another hansard corpus is the Canadian Hansard Corpus and it is well known for its parallel corpus. This corpus contains debates from the Canadian Parliament in both English and French (the country's official languages), and has been utilized for translation research and to build bilingual dictionary (Canadian Hansard Corpus, 1996). Australia, too, has its own hansard corpus called the Australian Parliament Question Time Corpus, consisting of the Question Time records of the Australian Federal Parliament House of Representatives from 1998 to July 2014 (Australian Parliament Question Time Corpus, 2015).

In Malaysia, the Malaysian Hansard Corpus (henceforth MHC) has been developed by a team of researchers from Universiti Kebangsaan Malaysia (UKM) to make available the hansard documents in electronic form which can be accessed by the public, upon request. The texts and information contained in MHC can be utilized as a set of data in cross-discipline researchs such as corpus linguistics, sociolinguistics, discourse analysis, political science and other disciplines, including the study of culturomics that can be carried out in the quantitative analysis of digital texts.

#### PREVIOUS STUDIES ON HANSARD

Various studies have used data taken from parliamentary debates. Baker (2006) examined keywords in the House of Commons in 2002 and 2003 debates on fox hunting in the UK. Using concordance analyses, the study identified different discourses that speakers utilized in order to persuade others of their standpoints. The study further explored ways in which keyness can be used to find salient language differences. The study postulated that a keyword list is useful for identifying lexical differences between texts. Bachmann (2011) also utilized the British Hansard to examine the language used in civil partnership debates in both Houses of the UK Parliament. The study employed a corpus-driven approach. By grouping keywords thematically and analysing them in context, scrutinising collocations and concordance lines, he was able "to see how (recurrent) uses of language construct gay and lesbian relationships" (Bachmann, 2011, p. 77).

Employing a corpus-based approach, Bayley et al. (2004) investigate lexical choices used by English, German and Italian MPs to communicate fear and reaction to fear in parliamentary debates. Linguistically, Italian MPs are said to use low-intensity words to

evoke danger and fear. Bayley and San Vicente (2004) also employed a corpus approach to investigate how UK and Spanish MPs talk about 'work'. The researchers employed concordance and collocation tools to study how collocation patterns characterise 'work' and to see whether there are any similarities or differences between English and Spanish. They concluded that concordancing software allows for identifying certain linguistic details that cannot be observed by simple perusal.

Valvason (2017) employed a corpus-based study to explore Hungarian politicians' attitudes towards the EU. This study utilized the corpus of Hungarian parliamentary debates (the HUNPOL corpus). Semantic preference and semantic prosody were spotted in the corpus through collocation analysis and concordance lines. The study showed that, in the years preceding Hungary's joining the European Union, Hungarian politicians referred to the EU with mixed voices, alternating positive stances by claiming the need to join the EU to develop further and faster, and negative standpoints by signaling warning that EU-membership could cause a loss of national strength.

A corpus-based study by Bijeikienė and Utkā (2006) focused on gender-specific features in parliamentary discourse from a combined sociolinguistic and corpus-linguistic methodological perspectives. The source data for the spoken language corpus consists of stenographs of the Lithuanian Parliament. It is shown that gender-related language variation in political communications does not follow a dichotomous pattern.

In Malaysia, several studies have been conducted utilizing the Malaysian Hansard. Ngeow et al. (2010) examined the use of sexist language and gender stereotypes in the parliamentary debates in the Malaysian parliament. The main data elicitation source of this study is the hansard in the Parliament official website from 2004 to 2008. This study did not adopt the corpus approach in its methodology, instead it employed qualitative content analysis to identify instances of gender stereotyping and sexist language use in Malaysian parliament. This study identified the emergence of three categories of frequent episodes involving gender stereotypes and/or sexist language, namely, beautiful women are sexual objects for lust; female divorcees are sexually promiscuous and menstruation is demonised.

Tan et al. (2017) studied the trend of vague language, in the form of adjunctive and disjunctive phrases, used by politicians in the Malaysian Parliament. The study utilized the Malaysian Parliament Hansard corpus specifically, from 1970 to 1979 and comparing it to the corpus from 2000 to 2009. Utilizing corpus tools viz collocation, concordance and keywords analysis the study concluded that, the trend of usage from the 70's to 80's showed a high preference for the use of adjunctives over disjunctives based on frequency per million. The findings showed that in the year 2000s the function of adjunctive phrase such as *dan sebagainya* has been expended to include both issues and solutions while the disjunctive phrase such as *atau lebih* was used extensively to quote legal acts.

Azianura et al. (2017) also utilized the Malaysian Hansard Corpus to look at how issues pertaining to crime were discussed in Malaysian Parliament. They found that there was a dramatic increase in the mention of the lexical *jenayah* or crime in Malaysian Parliament from the year 1995 to 2016. She concluded that this increase correlated with the increase of crime statistics in Malaysia from 1980 to 2004.

Past studies utilizing hansard corpus have delved into various issues brought up in parliamentary debates such as fox hunting or civil partnership in the UK parliamentary debates. There are also comparative studies on expressions of and reaction to fear by English, German and Italian MPs and another that looked at how British and Spanish MPs characterized 'work'. On the local front, there are various studies that utilized the Malaysia hansard data such as Ngeow et al. (2010) which looked at the use of sexist language and

gender stereotypes. Tan et al. (2017) and Azianura et al. are two studies to date that utilized the MHC.

There are obviously various issues that can be explored using the corpus on parliamentary debates. Kwabena (2018) through UK and Ghanaian parliamentary data has shown that a corpus-driven method can help in determining one's research focus in the study of parliamentary debates. Therefore, this study aims to examine the issues that are frequently associated with one of the states in Malaysia, specifically, Selangor. Selangor was chosen as the focus of this study because it has the highest Gross Domestic Product (GDP) in Malaysia (Department of Statistics Malaysia, 2017). Hence, it would be interesting to look into the issues related to Selangor discussed in parliament. The lexical 'Selangor' was also chosen because of its consistently high frequency in the MHC (refer to Figure 2) compared to other lexicals which represent states in Malaysia such as Johor, Kuala Lumpur and Pulau Pinang and frequency of co-occurrence of linguistics variable is the most commonly used statistic in corpus linguistics (Gries, 2010).

## METHODOLOGY

The current research adopts a corpus-driven approach utilizing both quantitative and qualitative methods. According to McEnery (2013:6) corpus-driven approach is "an inductive process where corpora are investigated from the bottom up and patterns found therein are used to explain linguistic regularities and exceptions of the language variety/genre exemplified by those corpora". Hence, by utilizing the MHC we will explore the patterns of usage of the lexical 'Selangor'.

## THE CORPUS

The corpora used in the present study is the Malaysian Hansard Corpus (MHC). MHC was developed from 2016 to 2018 from the reports of Malaysian Parliamentary Debates in the House of Representative (*Dewan Rakyat*). To date, the corpus contains texts from 3,511 files of reports from 13 parliamentary proceedings (Parliaments 1 – 13). The first parliamentary proceeding was convened from September 1959 to January 1964 while the last entry of the corpus was taken from the 13th parliamentary proceeding which was convened from June 2013 to April 2018 (the 13th Parliament was dissolved on 7 April 2018). This diachronic corpus of parliamentary discourse contains over 157 million words.

## THE COMPOSITION OF THE MHC

The existing MHC is divided into smaller subcorpora which are sorted according to the Parliaments (*Parlimen*), Sessions (*Penggal*) and Meetings (*Mesyuarat*). Each meeting has a number of reports because it is held for several days. Each day of the meeting is reported in one file. Therefore, a Meeting in a Session could comprise of at least 2 report files, for example in Parliament 1, Session 1, Meeting 2 could have as many as 95 files as is the case in Parliament 8, Session 1, Meeting 2. One report denotes one sitting day.

To date, there are a total number of 3,511 Hansard reports of the House of Representatives collected in MHC with the total of over 157 million words (token).

TABLE 1. The Divisions of House of Representative according to Parliaments and Sessions: Last Measured in April 2018 (Imran et al., 2018)

Parliament Session	1	2	3	4	5	6	7	8	9	10	11	12	13	Total
1	12	50	43	63	108	109	103	109	70	4	68	80	54	
2	46	60	50	72	59	66	65	96	84	77	77	70	57	
3	44	56	45	61	64	60	64	4	86	77	91	83	93	
4	38	31		71		21	24	67	72	61	85	63	53	
5	40	44							46	67		70	57	
6												20		
<b>Total</b>	182	241	138	267	231	256	255	276	385	286	321	366	334	3,511

Based on Table 1, the total number of 3,511 reports are compiled from Parliament 1 to Parliament 13. Parliament 1 has the least number of reports with 182 reports while Parliament 9 has the most number of reports (385). Customarily, each Parliament has 5 sessions except for Parliament 13 which has 6 sessions. The number of reports denotes the number of days of the parliamentary sittings.

#### THE SIZE OF MHC

Figure 1 shows the token (running words) in texts of each parliamentary sessions.

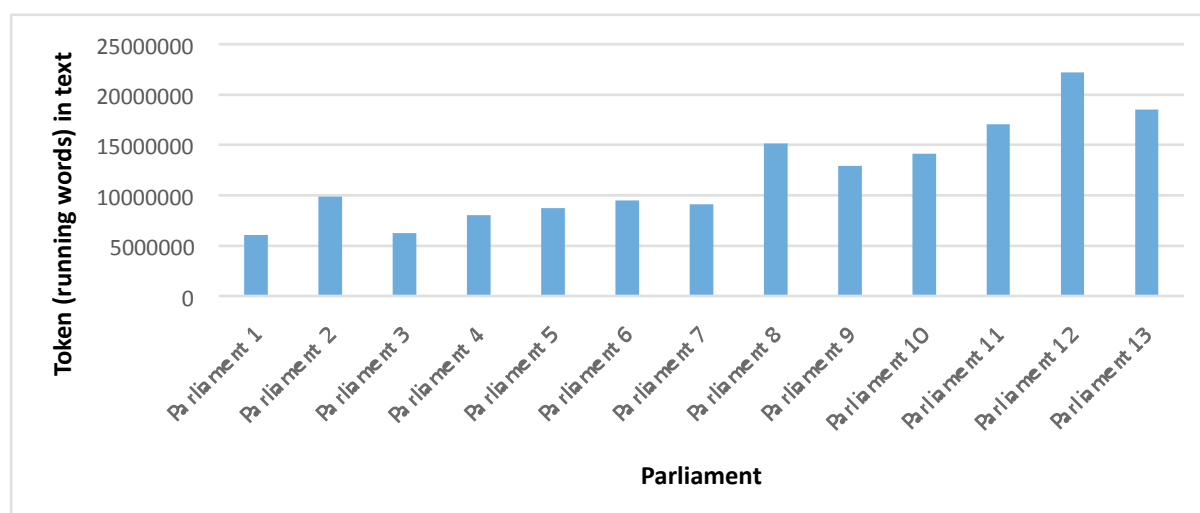


FIGURE 1: Division of Basic Statistical Information according to Parliaments (Generated using WordSmith 5.0) (Imran et al., 2018)

Based on figures illustrated in Figure 1, there is an upward trend in the number of token over the years. Parliament 3 has the least number of word token while Parliament 12 has the most number of word token. The increasing number of word token may indicate that the number of debates have also increased in the parliament.

#### THE CORPUS TOOLS

The corpus tool utilized to process and analyze the target corpus for the current research is WordSmith Tools (henceforth, WST), version 5 (Scott, 2008). There are three main tools within WST, namely; Concord, KeyWord and Wordlist and 12 utilities. A brief explanation of each of the main tools are as follows.

#### CONCORD

The primary purpose of a concordance is to see many lexical items in context (Key Word in Context, KWIC format). Scott (2008, p. 106) states that “you get a much better idea of the use of a word by seeing lots of examples of it, and it’s by seeing or hearing new words in context lots of times that you come to grasp the meaning of most words in your native language”. The concordance allows a researcher to analyse a word, part of a word or lexical chunk in its linguistic context, and thereby notice recurring lexical features or idiosyncratic usages of a lexical item or chunk (Guillén Nieto et al., 2008).

#### KEYWORDS

The purpose of the KeyWords is to compare two word lists: a reference corpus and a study corpus. Scott (2008, p. 156) states that “key-words provide a useful way to characterise a text or a genre”. The comparison results in a list of keywords where the frequencies are significantly different between the reference corpus and the study corpus. For example, if a word has a frequency of 6% in the reference corpus, and 5% in the study corpus it will not be a keyword, even though it may be the most frequently occurring word (Scott, 2008). The KeyWord tool is useful for observing lexical similarities or differences between two texts and lastly, noticing the writer’s stylistic preferences (Guillén Nieto et al., 2008).

#### WORDLIST

Scott (2008) describes the purpose of the WordList tool as (1) to analyse the vocabulary used, (2) to identify common word clusters, (3) to compare the frequency of a word in different text files or across genres, (4) to compare the frequencies of cognate words or translation equivalents between different languages and, (5) to get a concordance of one or more words in a list. The current research, however, uses only two out of the three WST main tools, namely Concord and Wordlist tools.

#### THE PROCEDURE

The following 4 steps were followed in order to meet the aim of the research:

**Step 1:** : Producing the frequency count of the lexical ‘Selangor’ in each parliamentary sitting i.e. from the first parliament through to the thirteenth parliament. The frequency count (raw frequencies) for ‘Selangor’ in each parliament was determined using WST and the frequency count was then normalized. Normalisation is a process to standardise word count which is done to accurately compare corpora (or sub-corpora) of different sizes (Evison, 2010). The line graph is then plotted and the plot indicates the trend of occurrence for the word ‘Selangor’ spanning from the first parliament to the thirteenth parliament

**Step 2:** Generating the collocates of ‘Selangor’ in each Parliamentary sitting. According to McEnery (2013), 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 ‘. Even though there are a number of statistical tests that can be used to measure collocational strength, for example the MI (mutual information), z, t, log-likelihood, log-log, and MI3 scores (Xiao and Mcenery 2006), the current research opted to use MI (mutual information) as a measure of collocation strength. A higher MI score means a stronger association and thus a more coherent relationship between words (Salazar 2014). MI score works well with content words rather than function words. MI score of 3 or

higher indicates that the two lexical items frequently co-occur (Xiao & Mcenery, 2006). However, this current research sets the minimum MI score to 6, and items must have a minimum co-occurrence frequency of 10 (within a 4–4 window span) to be considered as a collocate of a given node word. These settings allow for greater statistical significant differences of the collocations.

**Step 3:** Identifying the most consistent collocate/s of ‘Selangor’.

This step was done manually. The MI value and the normalized co-occurrence frequencies of the identified collocate/s were recorded and then plotted on a line graph.

**Step 4:** Identifying the context of use by concordance analysis (Keyword in context) of the collocate identified in step 3. WST is used here to determine the context of use of the collocate in relation to the lexical term ‘Selangor’.

## FINDINGS AND DISCUSSION

The research findings will be presented in three categories:

The frequency of ‘Selangor’ in the MHC.

The collocates of ‘Selangor’ in each parliamentary session in the MHC.

The most consistent collocates of ‘Selangor’ throughout the 13 Parliamentary sessions – ‘air’/water.

### ‘SELANGOR’ IN MALAYSIAN HANSARD CORPUS (MHC)

This study focuses on the profile or trend of one of the states, namely ‘Selangor’. As the most developed state in Malaysia (The Sultan of Selangor Opening Address in the 3<sup>rd</sup> Session of the 13<sup>th</sup> State Assembly, *Hansard Selangor*, 2015) ‘Selangor’ or issues pertaining or related to Selangor would often be brought up in the Malaysian Parliamentary debates. Utilizing WST, ‘Selangor’ was set as the node to determine its frequency in 13 parliamentary sessions to plot the trend of its occurrences in those parliamentary debates. Figure 2 indicates the trend of ‘Selangor’ in MHC.

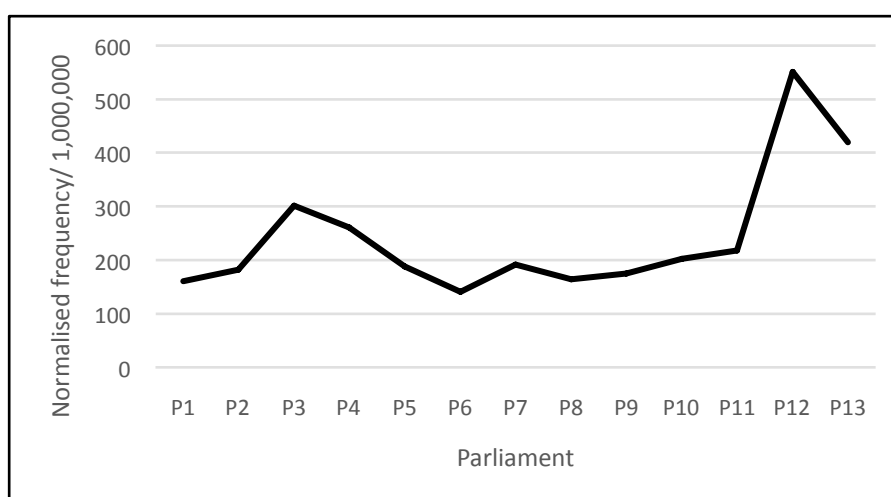


FIGURE 2. The Trend of “Selangor” in MHC

Figure 2 shows that ‘Selangor’, indeed, is widely mentioned in Parliament 1 to Parliament 13 particularly in P3, P7, P11 and 12. Hence, our next aim is to look at the issues that surround ‘Selangor’. In order to get an insight on the issues discussed or debated



pertaining to ‘Selangor’, the collocates of ‘Selangor’ from P1 to P13 were generated. The collocates of ‘Selangor’ was generated for all 13 parliamentary periods. The aim is to look at the pattern of the top 15 collocates of the node ‘Selangor’ with Mutual Information (MI) score and frequency set as  $> 6$  and  $> 10$  respectively (see Appendix A). The list is grouped according to the Parliament and sorted according to MI score. Based on the word list generated, most of the top collocates with ‘Selangor’ are nouns, specifically proper noun such as names of states *Perak, Pahang, Kedah* or names of persons such as *Chen, Lee, Daud* or *Ibrahim*. Another trend observed from the list is the consistent presence of the lexical *air* or water as one of the collocates of ‘Selangor’ from P4 to P13, although *air*/water did not occupy the upper rank of the collocates list. Hence, we will next look at the context of occurrence of *air*/water in MHC.

#### ‘AIR’/WATER AS A CONSISTENT COLLOCATE OF SELANGOR

As one of the consistent collocates for *Selangor*, a detailed information on *air*/water needs to be examined. Figure 3 shows the trend of the collocate *air*/water with the node *Selangor* based on its MI score for each Parliament. The two dotted lines signify the parameters of the MI score for the study. The study sets the collocate results of MI score at 6.0 and above.

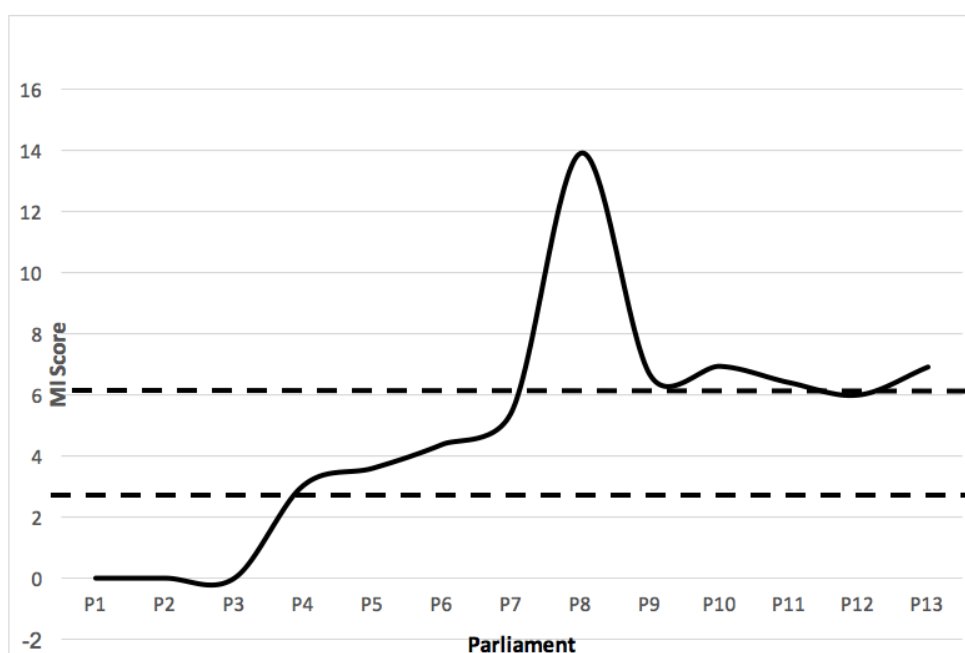


FIGURE 3: The MI Score of ‘air’/water as a collocate for ‘Selangor’ from P1-P13

Figure 3 shows that *air*/water is one of the consistent collocates of the node *Selangor* from P4 to P 13. This, however, does not mean that *air*/water did not appear in P1, P2 or P3 (1959-1973), on the contrary, *air*/water did appear but not as a collocate of *Selangor*. Additionally, *air*/water was spelled as *ayer* during those periods until the formation of the *Majlis Bahasa Indonesia – Malaysia (MBIM)* – Indonesia-Malaysia Language Council in 1972, which also launched the standardization of the romanised spelling system of the Malay language in Malaysia and Indonesia (Awang Sariyan, 2002). *air*/water was also referred to by its English equivalent ‘water’ during those three parliamentary periods.

In terms of frequency, *air*/water also shows an interesting trend as demonstrated in Figure 4.

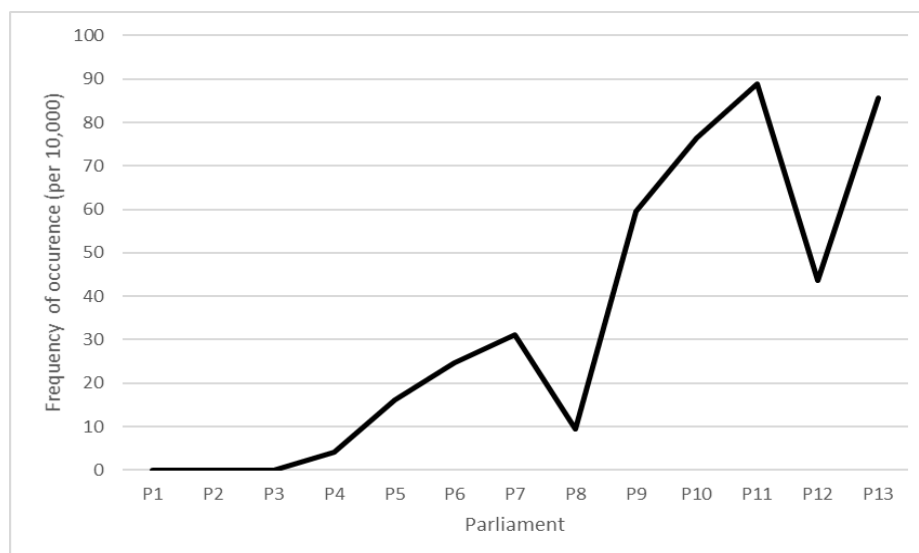


FIGURE 4. The Trend of *air/water* as a collocate of *Selangor*

Figure 4 shows the trend of *air/water* as the collocate of *Selangor*. The figures are generated from the frequencies of the collocate *air/water* with the node *Selangor*. To compare between the parliaments, each of the frequencies were normalized. *air/water* started to show an upward trend in P7 but went down in P8. However, from P8 *air/water* seems to regain momentum as it continuously became one of the issues frequently brought up in parliamentary debates and reached its peak in P11. *air/water*, however took a downturn in P12 but still continues to be one of the frequently discussed topics in P13 as shown in Figure 4. In light of this interesting trend, it is felt that the issues that surround *air/water* should be closely examined by looking at its contextual usage. This paper will focus on P7 to P13 because beginning with P7 the MI score for *air/water* is 6 which is the MI score set for this study (see Figure 3 for details). We will next look at the concordances of *air/water* and its contextual usage.

#### THE CASE OF AIR/WATER

The issue of water and Selangor has a long history. In 1998, the water crisis in Selangor reached its peak when residents in the Klang Valley were subjected 150 days of water rationing and it has been dubbed as ‘the country’s worst water shortage’ (The Star Online, 26 February 2014a). The dry spell was caused by drought-inducing El-Nino phenomenon with no rainfall in the catchment area between January and March 1998 (Rasyidah Md Khalid, 2018). As a consequence of the water shortage, 3.2 million residents in Klang Valley had to be put on water rationing exercise from April 20, 1998 to September 16 during the same year. In 2014 Selangor again experienced a severe water shortage. This time it is blamed on ‘over development and lack of planning’ (The Star, March 19, 2014b). With 2 major water crisis in Selangor, we will now turn to the Malaysian Hansard Corpus (MHC) to observe the trend of *air/water* as it is debated in parliament.

#### AIR/WATER IN PARLIAMENT 7 (1986-1990)

In Parliament 7, *air/water* collocates mostly with *pembersihan* which literally means cleaning. However, contextual usage from concordances show that *pembersihan* referred to a different kind of cleaning i.e. the treatment of water in water treatment plant as shown in examples below:

### Concordance 1:

The treatment of water

cadangan untuk menswastakan operasi loji pembersihan air Sungai Semenyih, Bentuk penswastakan yang sedang jumlah air yang dibekalkan daripada loji pembersihan air. Saya telah difahamkan bahawa syarat-syarat kontrak kontrak operasi dan penyelenggaraan loji pembersihan air kepada sektor swasta. Mengikut cara penswastakan ini,

2 December 1986

- a. *“Mengenai penswastakan bekalan air Selangor itu, saya telah difahamkan bahawa Kerajaan Selangor telah menimbang cadangan untuk menswastakan operasi loji pembersihan air Sungai Semenyih, Bentuk penswastakan yang sedang dipertimbangkan ialah secara kontrak operasi dan penyelenggaraan loji pembersihan air kepada sektor swasta. Mengikut cara penswastakan ini, Kerajaan Negeri Selangor akan membayar kontraktor bekalan jumlah air yang dibekalkan daripada loji pembersihan air.”*

“In the privatization of the water supply in Selangor, I was informed that the Selangor Government has considered the proposal to privatize the water treatment plant in Sungai Semenyih. The privatization being considered is privatizing the operation and maintenance of water treatment to the private sector. According to the terms of the privatization, the State Government of Selangor will pay the contractor for the amount of water supplied from the water treatment plant.”

The issue at hand is the proposal by the Selangor state government to privatize the water treatment plant in Sungai Semenyih. This plan is signaled by the word *cadangan* or proposal. This debate took place on 2 December 1986 which is before the big water crisis in 1998.

#### AIR/WATER IN PARLIAMENT 8 (1990-1994)

In Parliament 8, *kawasan* or area is the top collocate for *air*/water, specifically water catchment area:

### Concordance 2:

Water catchment area

hutan. Itupun 30% daripada hutan itu untuk 'catchment area', *kawasan* tadahan air. Yang Berhormat menteri tidak berusaha untuk memperluaskan hutan di Negeri , ini yang di tepi yang dekat dengan Sungai Buluh dan juga di *kawasan* Tadahan Air. Cuma Yang Berhormat Menteri Besar Selangor kata sebahagiannya disimpan melalui rancangan pengairan Tanjong Karang yang memperolehi punca bekalan air dari *kawasan* Tadahan Sungai bernaam dan Sungai Tinggi. Bekalan air diambil dari

4 August 1993

- b. *“Saya lebih satu dekad di Dewan Undangan Negeri sentiasa membangkitkan hal ini, mengenai perhutanan di Negeri Selangor yang mempunyai hanya 30% humi negeri itu adalah hutan. Itupun 30% daripada hutan itu untuk 'catchment area', kawasan tadahan air.”*

“I have raised this issue in the State Assembly for more than a decade and, about forestry in the state of Selangor which covers only 30% of the state. And that 30% is designated as ‘catchment area’.”

25 October 1991

- c. *“Tuan Yang di-Pertua, saya hendak jelaskan di sini, di Sungai Buluh yang dekat dengan Forest Reserved Institute, kawasan itu memang hutan simpanan yang kekal. Forest Reserve Institute kita tidak boleh ditebang, ini yang di tepi yang dekat dengan Sungai Buluh dan juga di kawasan Tadahan Air.”*

“The Honourable Speaker, let me clarify here, in Sungai Buluh which is close to the Forest Reserve Institute, a permanent forest reserved area. Cutting trees in the Forest Reserve Institute is not allowed. This area is adjacent to Sungai Buluh and it is also in the water catchment area.”

From both examples (b) and (c), the debates were about areas designated as water catchment area. In (b) the concern was about limited area in Selangor (about 30%) that can serve as as water catchment area. In (c) it is about a water catchment area which is adjacent to the Forest Reserved Institute.

#### AIR/WATER IN PARLIAMENT 9 (1995-1999)

In Parliament 9, there were two collocates which frequently co-occur with *air*/water and they were *Enakmen* (Enactment) and *penswastaan* or privatization.

18 December 1999

- d. *“Saya mengucapkan banyak terima kasih kepada pihak kementerian dan juga kepada kerajaan seterusnya atas inisiatif ini untuk mengemaskinikan dan meminda undang-undang yang termaktub pada tahun 1974 untuk diselaraskan dengan pindaan pada undang-undang yang sama Enakmen Bekalan Air Selangor.”*

“I would like to thank the ministry and the government for this initiative to update and amend the law that was formed in 1974 so that it is in line with the amended law within the Selangor Water Supply Enactment.”

### Concordance 3:

#### Privatisation of Water Supply

berharga RM2 bilion ialah satu produk penswastaan skim bekalan air sungai Selangor fasa III yang telah diserahkan kepada TSWA Sendiri dan bekalan masih merupakan tanggungjawab Jabatan Bekalan Air Selangor seperti sebelum penswastaan. Akan tetapi krisis bekalan Selangor menimbang semula rancangan penswastaan bekalan air di Selangor. Bukan bermakna langkah penswastaan akan cadangan baru amatlah kami alu-alukan. Kedua, masalah penswastaan air Selangor kepada syarikat Puncak Niaga Bhd. - Sebagaimana Jabatan Bekalan Air negeri Selangor melalui konsep penswastaan air negeri. Sebilangan besar rakyat yang tinggal di kawasan Lembah

19 April 1999

- e. *“Projek empangan Sungai Selangor berharga RM2 bilion ialah satu produk penswastaan skim bekalan air sungai Selangor fasa III yang telah diserahkan kepada TSWA Sendirian Berhad iaitu syarikat milik Tan Sri Wan Azmi, Gamuda Berhad dan kumpulan Darul Ehsan Berhad yang ditubuhkan untuk mengatasi masalah kekurangan air yang dijangka melanda sekitar Lembah Klang tahun 2003.”*

“The Sungai Selangor dam project valued at RM2 billion is the product of the Phase III in the privatization of Sungai Selangor water supply that was given to TSWA Sendirian Berhad. It is a company owned by Tan Sri Wan Azmi, Gamuda Berhad and the Darul Ehsan Berhad group that was established with the purpose to overcome water shortage problem that is expected to hit the Klang Valley area in 2003.”

In Parliament 9, the top collocates for *air*/water were *enakmen* or enactment and *penswastaan* or privatization. Both examples, (d) and (e), were taken from meetings in 1999 – a year after the 1998 water crisis which has been labelled as the worst water shortage to have happened in the country. Hence, the Selangor Water Supply Enactment was amended and measures were taken to overcome water shortage by privatizing the water supply.

#### AIR/WATER IN PARLIAMENT 10 (1999-2003)

For Parliament 10, the two collocates closely associated with *air*/water were *dikorporatkan* or corporatization and *penyaluran* or to transfer (water).

#### Concordance 4:

##### Corporatization of JBS

satu perkara lagi, iaitu saya ingin tanya yang Jabatan Bekalan Air Selangor telah **dikorporatkan** pada 15 Mac baru-baru ini dan bagi membolehkan JBAS **dikorporatkan** menjadi Perbadanan Urus Air Selangor iaitu PUAS Berhad. Adakah dijangkakan dinikmati oleh pengguna air selepas PUAS dan Jabatan Bekalan Air Selangor **dikorporatkan**. Apakah pengguna akan menghadapi Pian Hon: Tuan Yang di-Pertua. soalan tambahan. Jabatan Bekalan Air Selangor akan **dikorporatkan** pada tahun depan dan saya minta Jabatan Bekalan Air Selangor. Walaupun Jabatan Bekalan Air Selangor **dikorporatkan** dengan nama baru PUAS dan semua

11 December 2001

f. “Untuk makluman Ahli Yang Berhormat, Dewan Undangan Negeri Selangor pada 1 November 2001 yang lalu telah meluluskan Enakmen Bekalan Air Syarikat Pengganti bagi membolehkan JBAS **dikorporatkan** menjadi Perbadanan Urus Air Selangor iaitu PUAS Berhad. Adakah dijangkakan pengkorporatan tersebut akan dapat dilaksanakan pada tahun hadapan.”

“Members of the House – for your information, the Selangor State Assembly had passed the Company Replacement for Water Supply Enactment on 1 November 2001. This is to enable JBAS to be corporatized to become Perbadanan Urus Air Selangor or PUAS Berhad. It is expected that the corporatization will take place next year.”

#### Concordance 5:

##### Water Transfer Project from Pahang to Selangor

air sambungan yang besar adalah projek **penyaluran** air mentah antara negeri iaitu daripada Pahang ke Berhormat, kos keseluruhan bagi projek **penyaluran** air mentah dari negeri Pahang ke negeri Selangor telah merancang di bawah Kepala 11200 - **Penyaluran** Air Antara Negeri, sebanyak RM200 juta bagi tujuan Sungai Selangor, Fasa Tiga dan juga Skim **Penyaluran** Air Mentah dari negeri Pahang ke negeri Selangor... yang sangat tinggi. Sebagai contoh, projek **penyaluran** air mentah dari Pahang ke Selangor yang sepatutnya . Yang kedua Tuan Yang di-Pertua, projek **penyaluran** air mentah dari negeri Pahang ke Selangor. sudah selesai. Mengenai soal projek **penyaluran** air dari Pahang ke Selangor, untuk makluman Ahli ialah sebanyak RM1.4 bilion dan bagi projek **penyaluran** air mentah dari Pahang ke Selangor ia merupakan satu siap pada Julai tahun hadapan serta projek **penyaluran** air Pahang-Selangor yang dijangka siap pada tahun

26 November 2000

g. “Antara projek bekalan air sambungan yang besar adalah projek **penyaluran** air mentah antara negeri iaitu daripada Pahang ke Selangor yang memerlukan peruntukan sebanyak RM2.6 bilion dalam Rancangan Malaysia Kelapan.”

“Among the big water supply connection is the project to transfer raw water between states which is from Pahang to Selangor that needs an allocation of RM2.6 billions in the Eighth Malaysian Plan.”

In Parliament 10 the debates on *air/water* centered on efforts taken to improve water supply services when the Selangor State Government took step to corporatize the Selangor Water Supply Department. Efforts were also taken to avoid another major water crisis in Selangor. To this end, the government has proposed to transfer raw water from neighbouring state Pahang to Selangor.

#### AIR/WATER IN PARLIAMENT 11 (2004-2007)

The top collocate for air in Parliament 11 is *penyaluran* or to transfer which can be observed in the examples below:

#### Concordance 6:

##### Water Transfer Project from Pahang to Selangor

dan rakan-rakan yang lain bahawa perjanjian **penyaluran air** mentah, Pahang ke negeri Selangor telah telah bangun dan menayakan dari segi projek **penyaluran air** mentah dari Pahang ke Selangor. DR.30.8.2006 101 ahli Yang Berhormat kerja-kerja fizikal, projek **penyaluran air** mentah dari Pahang ke Selangor akan dimulakan pada ini adalah dengan melaksanakan projek **penyaluran air** mentah dari negeri Pahang ke negeri Selangor dan beberapa isu mengenai pelaksanaan Projek **Penyaluran Air** Mentah dari Pahang ke Selangor. Untuk makluman ahli RM100 juta telah diperuntukkan. Projek **Penyaluran Air** Selangor-Pahang iaitu satu projek yang disenaraikan secara ringkasnya saya hendak tanya, mengenai bekalan air, **penyaluran** air daripada Pahang ke negeri-negeri lain tadi. Yang Berhormat Serdang, status projek **penyaluran air** negeri Pahang ke Selangor. Rakan saya Yang

23 November 2005

- h. “Saya hendak tanya apakah punca perkara ini berlaku dan seterusnya secara ringkasnya saya hendak tanya, mengenai bekalan air, *penyaluran air* daripada Pahang ke negeri-negeri lain terutamanya negeri Selangor.”

“I would like to ask what is the reason for this to happen and next, briefly I would like to ask about water supply, water transfer from Pahang to other states especially Selangor.”

27 November 2007

- i. “Saya ingin merujuk kepada butiran 050200 - Pasukan Projek *Penyaluran Air Mentah Pahang Ke Selangor*. Sebagai mana yang telah disebut oleh timbalan menteri dan rakan-rakan yang lain bahawa perjanjian *penyaluran air* mentah, Pahang ke negeri Selangor telah ditandatangani pada 22 November yang lalu.”

“I would like to refer to article 050200 – the project on Raw Water Distribution from Pahang to Selangor. The Deputy Minister and friends have mentioned that the agreement on the transfer of raw water from Pahang to Selangor has been signed on 22 November.”

The issue of transferring raw water from Pahang to Selangor continues to dominate debates on *air/water*. In 2005 the issues of water supply and water transfer from Pahang was the focus in debates on *air/water* and Selangor. The same issue was again mentioned in 2007 debates but this time it seemed to be reporting on its progress. This is signaled by the lexical *perjanjian* or ‘agreement’ and *ditandatangani* or ‘signed’ in (i). The Selangor State Government and the Pahang State Government signed a sale and purchase agreement of raw water on 22 November 2007 (Memorandum Persefahaman Antara Kerajaan Malaysia Dan Kerajaan Negeri Selangor Berhubung Penstrukturan Semula Industri Bekalan Air Selangor, 2014).

AIR/WATER IN PARLIAMENT 12 (2008-2012) & PARLIAMENT 13 (2013-2018)

In Parliament 12 and Parliament 13 there were two recurring issues brought up in relation to water, namely *penyaluran air* or the transfer of water and second is *penstrukturan air* or the restructuring of water services. Contextual information in the form of concordances are shown below:

**Concordance 7:**

Water Transfer Project from Pahang to Selangor in Parliament 12

projek yang satu lagi yang di sebelah **Penyaluran Air** Mentah Pahang-Selangor ataupun (PAMPS) itu yang tahu sejauh manakah terowong-terowong untuk **penyaluran air** daripada Pahang ke Selangor ini telah dibina? Di mana Berhormat Kepong secara bertulis. Bagi projek **penyaluran air** dari Pahang ke Selangor ini, ingin saya jelaskan kepada kita tahu bahawa satu projek mega projek besar **penyaluran air** dari Sungai Pahang ke Selangor. Kenapa ini perlu Air Langat 2 dan sistem agihan. Projek **Penyaluran Air** Mentah dari Pahang ke Selangor dijadualkan siap di Selangor ini, kita pada hari ini bincang projek **penyaluran air** kepada Pulau Ketam dan itulah dalam apa yang kita Yang di-Pertua, mengenai pelaksanaan Projek Rawatan Air Langat 2 dan Projek Penyaluran Air Mentah yang telah memohon penjelasan tentang jumlah kualiti air mentah yang dibekalkan oleh Penyaluran Projek Air Tuan Pengerusi, saya juga ingin menyentuh soal krisis air dan Projek **Penyaluran Air** Pahang - Selangor. Kerajaan kerugian kewangan kerajaan dalam kontrak terowong air bagi projek **penyaluran air** mentah Pahang ke Selangor kualiti air mentah yang dibekalkan oleh **Penyaluran Projek Air** Mentah Pahang-Selangor. Saya kira projek-projek yang tidak termasuk pada projek PPAMPS (**Penyaluran Projek Air** Mentah Pahang-Selangor). Mungkin dalam forum yang

30 November 2010

j. *“Akhirnya Tuan Pengerusi, saya juga ingin menyentuh soal krisis air dan Projek Penyaluran Air Pahang - Selangor.”*

“Lastly Mr Chairman, I would like to touch on the water crisis and the Pahang-Selangor Water Transfer Project.”

**Concordance 8:**

Water Transfer Project from Pahang to Selangor in Parliament 13

ini adakah Kerajaan Negeri Selangor membenarkan pembinaan air dan juga **penyaluran air** daripada Pahang ke Selangor dan kerana Rancangan Bekalan Air Negeri Selangor, Butiran 11200 - **Penyaluran Air** Antara Negeri. Mula-mula kita beri gula-gula air percuma, Sukalah saya maklumkan di Dewan ini bahawa projek **penyaluran air** mentah dari Pahang ke Selangor merangkumi empat komponen Tuan Yang di-Pertua, projek ini berkaitan dengan projek **penyaluran air** mentah dari Pahang ke Selangor. Sukalah saya maklumkan di Bekalan Air Negeri Selangor dan Butiran 11200 - **Penyaluran Air** Antara Negeri. Kesemua perkara ini ada hubungannya dengan Air. Saya pergi kepada Butiran 020100 - Pasukan Projek **Penyaluran Air** Mentah Pahang ke Selangor. Saya sebenarnya merakamkan merujuk Butiran B.29 - Butiran 020000 - Pasukan Projek **Penyaluran Air** Mentah Pahang ke Selangor yang berjumlah RM2.95 juta. Jumlah

11 March 2014

k. *“Komponen yang pertama ialah projek penyaluran air mentah Pahang - Selangor yang di dalamnya terdiri daripada komponen Empangan Kelau di Pahang, terowong penyaluran air mentah, dan Stesen Pam Semantan dan juga saluran paip dan kerja-kerja berkaitan.”*

The first component is the Pahang-Selangor raw water transfer project which includes other components such as the Kelau Dam in Pahang, tunnel to transfer raw water, and the Semantan Pump Station and pipeline and other related works.

The inter-state raw water transfer project from Pahang to Selangor continued to be one of the issues debated in Parliament 12 and Parliament 13. The issue was debated over two parliamentary periods and this could be attributed to the fact that in the 2008 General Election (P12) when Selangor fell to the opposition, hence the project that was agreed upon by the previous government needed to be re-examined. It could also be because of the recurring

incidents of water crisis as brought up by one of the MPs in (j). Water crisis caused by severe drought hit Selangor in 1998 and 2014 (Rasyidah Md Khalid, 2018) and therefore efforts taken to stem it need to be debated.

We will next look at *penstrukturan air* or the restructuring of water services.

### Concordance 9:

#### Restructuring of Water Services in Parliament 12

tahu bahawa konsesi ini akan juga beralih, *penstrukturan* semula *air* negeri Selangor akan juga berlaku dan pada masa ini kita tidak yang masih dalam proses perbincangan. *Penstrukturan* semula *air* Selangor ini adalah selaras dengan hasrat Akta Industri air secara berterusan. Tuan Yang di-Pertua, mengenai *penstrukturan* *air* Selangor, Ahli Yang Berhormat Hulu Langat, Klang, Bandar Tun isu berkaitan *penstrukturan* semula industri perkhidmatan bekalan *air* di negeri Selangor. Saya ingin menegaskan di Dewan yang mulia tanah sehingga *penstrukturan* semula industri perkhidmatan bekalan *air* di negeri Selangor dimuktamadkan mengikut pendekatan terma kerana telah hampir dua tahun kemelut *penstrukturan* semula bekalan *air* belum selesai. Adalah tidak adil untuk mengamalkan double giat dijalankan untuk memuktamadkan *penstrukturan* semula industri *air* yang mana saya ingin komitmen jitu pihak kementerian yang untuk merundingkan *penstrukturan* semula perkhidmatan bekalan *air* bagi negeri Selangor tiada banyak bezanya dengan apa yang telah ataupun concessionaire, tetapi bila *penstrukturan* semula bekalan *air* di negeri Selangor ini akan diambil kira dan buat masa ini ia tidak yang belum muktamad mengenai *penstrukturan* semula bekalan *air* ini. Terbaru Kerajaan Selangor berdegil untuk tidak menyetujui khas iaitu Jawatankuasa *Penstrukturan* Semula Industri Bekalan *Air* Selangor dan di dalam jawatankuasa ini juga diwakili oleh wakil

4 March 2009

- l. *“Oleh yang demikian, dasar dan pendekatan yang digunakan untuk merundingkan penstrukturan semula perkhidmatan bekalan air bagi negeri Selangor tiada banyak bezanya dengan apa yang telah dilaksanakan oleh Kerajaan Persekutuan bagi negeri-negeri lain”*

Therefore, the policy and approach used to negotiate the restructuring of water supply services for the state of Selangor is not much different from what has been implemented by the Federal Government in other states.

### Concordance 10:

#### Restructuring of Water Services in Parliament 13

soalan sahaja. Nombor satu isu mengenai penyelesaian *penstrukturan* *air* di negeri Selangor antara Kerajaan Pusat dengan Kerajaan Negeri dan yang kedua, pembinaan Loji Langat 2. Pertama, *penstrukturan* *air* Selangor, saya hendak dapatkan penjelasan sama ada perjanjian perkara berhubung *penstrukturan* semula industri perkhidmatan bekalan *air* di negeri Selangor dan meminta agar perjanjian berkenaan dapat jumlah kos keseluruhan bagi *penstrukturan* semula industri bekalan *air* Selangor, Wilayah Persekutuan dan Putrajaya, termasuk pembiayaan dilaksanakan sepenuhnya. Tuan Yang di-Pertua, masalah *penstrukturan* *air* di Selangor sehingga kini masih lagi belum dilaksanakan sepenuhnya. gus menyelesaikan *penstrukturan* semula industri perkhidmatan bekalan *air* negeri Selangor, Wilayah Persekutuan Kuala Lumpur dan Putrajaya.

9 July 2013

- m. *“Sehubungan dengan itu, Kerajaan Persekutuan telah pun memulakan perbincangan dengan Kerajaan Negeri Selangor bagi mengatasi masalah ini dan sekali gus menyelesaikan penstrukturan semula industri perkhidmatan bekalan air negeri Selangor, Wilayah Persekutuan Kuala Lumpur dan Putrajaya.”*

In line with that, the Federal Government has started talks with the Selangor State Government to overcome this problem and at the same time complete the restructuring exercise on water supply services in Selangor, Federal Territory of Kuala Lumpur and Putrajaya.



17 June 2015

- n. “Tuan Yang di-Pertua, ada masalah penstrukturan air di Selangor sehingga kini masih lagi belum dilaksanakan sepenuhnya. Saya telah menerima jawapan daripada Menteri Tenaga, Teknologi Hijau dan Air pada 27 Mei dan 9 Jun yang lepas berkenaan dengan dua isu di atas. Jawapan yang diterima jelas menunjukkan bahawa masalah air ini masih tidak nampak kesudahannya.”

Mr Speaker, there is problem in the water restructuring exercise in Selangor and to this date the restructuring exercise has not been executed completely. I have received answers from the Minister of Energy, Green Technology and Water on 27<sup>th</sup> May and 9<sup>th</sup> June pertaining to both issues. The answer received clearly showed that there is still no end to this water problem.”

The restructuring exercise on water supply and services in Selangor continues to make its way in the parliamentary debates. In (l), the MPs were informed that the policy and approaches used in restructuring water services is similar to what has been practised in other states despite the fact that Selangor is now controlled by the opposition. In Parliament 13 the water restructuring exercise is again given attention in the debates. However, the focus of the debate was on the problems encountered in the water restructuring exercises as shown by the lexical ‘problem’ in (m) and (n). The continuous problems faced in the restructuring exercise has prompted one of the MPs to remark in (n) that there seemed to be no end to the water problem.

The analysis of the collocates for *air/water* in the MHC has shown an interesting trend in the discussions and debates on *air/water* in Parliament. Starting with Parliament 7, the top collocate for *air/water* is *pembersihan* and its literal translation is cleaning. However, the cleaning that is referred to is the treatment of water, specifically water treatment plant. In Parliament 8, the top collocate for *air/water* is *kawasan* or its English translation is area. Specifically, the area referred to in Parliament 8 is water catchment area.

In Parliament 9 there were two top collocates for *air/water*, namely *enakmen* or enactment and *penswastaan* or privatization. *Enakmen* became one of the top collocates in Parliament 9 is probably driven by the country’s worst water crisis in 1998. As a result of the water crisis, residents in the Klang Valley (which includes part of Selangor) were forced to endure 150 days of water rationing exercise (Rasyidah Md Khalid, 2018). New enactment on water were discussed to perhaps be better prepared in case severe droughts were to occur in the future. Another recurring theme is *penswastaan* or privatization. Privatization of water services was discussed as one of the ways to provide better services to the public.

Next, in Parliament 10 the two top collocates for *air/water* were *dikorporatkan* or corporatized and *penyaluran* or to transfer. The House was informed that the Selangor Department of Water Supply has undergone corporatization exercise and is now known as PUAS. The other collocate, *penyaluran* or to transfer was also one of the top collocates in Parliament 11, Parliament 12 and Parliament 13. In all four parliamentary periods (P10, P11, P12 and P13) *penyaluran* was discussed in context of the transfer of raw water from Pahang to Selangor. In Parliament 10, *penyaluran* was introduced to inform members of the House of the suggestion and plan to transfer raw water from Pahang to Selangor. This was put forth after the 1998 severe drought that led to water crisis in the Klang Valley. The debate on the transfer of raw water from Pahang to Selangor continued in Parliament 11, Parliament 12 and Parliament 13. Another collocate that dominated the debates on water in Selangor is *penstrukturan* or restructuring. In both Parliament 12 and Parliament 13, *penstrukturan* was debated in terms of the restructuring of the water supply and services in Selangor. The trend of the debate on water in Selangor has not only shown that it is interesting but it has also shown that water is linked to various issues.

## CONCLUSION

This corpus-driven study has led us to ‘Selangor’. Through statistical analysis we were able to determine the issue or issues closely related to Selangor and *air* or water has been identified as one of the issues closely related to Selangor. By utilizing a corpus data, we were able to chart the trend of *air*/water throughout the thirteen parliaments. For example, the major water crisis experienced in 1998 which led to initiative and plan by the government to transfer raw water from Pahang to Selangor as discussed in P10 or the issue of transferring raw water from Pahang to Selangor which has spanned over four parliamentary periods.

Utilizing large data such as the Malaysian Hansard Corpus (MHC) has opened up a lot of possibilities for further research. Parliamentary corpora are multidisciplinary language resource (CLARIN 2019) and as such research from multi disciplines is made possible. This is shown by Azianura et al. (2017) who conducted a lexical analysis of the word *jenayah* and observed its dramatic increase in occurrence in the period 1995 to 2016 which correlates with the increase in crime statistics in Malaysia. As a multidisciplinary language resource, the MHC not only offers data for research in language and linguistics but also research in other fields such as political science, history and sociology.

Parliament represents the people and as such various topics of national interest are raised during the parliamentary debates. This paper has managed to present the trend one of the issues raised frequently raised i.e. ‘water’ and its relation with Selangor. There are, however, more issues of national interest that can be explored utilizing the Malaysian Hansard Corpus (MHC).

## ACKNOWLEDGMENT

This study was supported by Universiti Kebangsaan Malaysia research grant KRA-2018-005.

## REFERENCES

- Australian Parliament Question Time Corpus. (2015). Retrieved 1 March, 2019 from <https://people.eng.unimelb.edu.au/aturpin/qtCorpus/index.html>.
- Awang Sariyan. (2002). Ceritera tentang Kejayaan Perancangan Bahasa Melayu di Malaysia. Retrieved 1 March, 2019 from <http://pts.com.my/berita/perancangan-bahasa-melayu-di-malaysia/>
- Azianura Hani Shaari, Marlyna Maros, and Imran Ho Abdullah. (2017). Tracking Trends: The Corpus Analysis of Lexical Item *Jenayah* in the Malaysian Hansard. Prosiding Kolokium Penerokaan Dataraya Kulturomik Malaysia 2017, pp 40 – 45.
- Bachmann, I. (2011). ‘Civil partnership – “gay marriage in all but name”: A corpus-driven analysis of discourses of same-sex relationships in the UK parliament’. *Corpora*. 6(1), 77–105.
- Baker, P. (2006). “‘The question is, how cruel is it?’ Keywords, foxhunting and the House of Commons’, AHRC ICT methods network expert seminar on linguistics. Retrieved 1 March, 2019 from [www.methodsnetwork.ac.uk](http://www.methodsnetwork.ac.uk)
- Baron, R. A., Byrne, D. R. & Branscombe, N. R. (2006). *Mastering Social Psychology* (1st ed.). Pearson.
- Bayley, P. & San Vicente, F. (2004). ‘Ways of talking about work in parliamentary discourse in Britain and Spain’ in P. Bayley (Ed.), *Cross-Cultural Perspectives on Parliamentary Discourse* (pp. 237–69). Amsterdam: John Benjamins.

- Bayley, P., C. Bevitori & E. Zoni. (2004). 'Threat and fear in parliamentary debates in Britain, Germany and Italy' in P. Bayley (Ed.), *Cross-Cultural Perspectives on Parliamentary Discourse* (pp. 185–36). Amsterdam: John Benjamins.
- Bijeikienė, V. & Utkā, A. (2006). Gender-specific features in Lithuanian parliamentary discourse: An interdisciplinary sociolinguistic and corpus-based study. *SKY Journal of Linguistics*, 19, 63-99
- British Parliament. (2019). Retrieved 1 March, 2019 from <https://www.hansard-corpus.org/>
- Canada Hansard Corpus. (1996). Retrieved 1 March, 2019 from <https://spraakbanken.gu.se/lb/pedant/parabank/node6.html>.
- Chamhuri Siwar & Ferdoushi Ahmed. (2014). Water resources in Malaysia: Issues and challenges. *Pakistan Journal of Nutrition* 13(5), 281-286.
- CLARIN. (2019). Parliamentary corpora. Retrieved 1 March, 2019 from <https://www.clarin.eu/resource-families/parliamentary-corpora>
- Department of Statistics. (2017). Retrieved 1 March, 2019 from [https://www.dosm.gov.my/v1/index.php?r=column/cthemByCat&cat=102&bul\\_id=OE5UdnNRL1VZbjJjcU1STVdIdFIUdz09&menu\\_id=TE5CRUZCb1h4ZTZMODZlbmk2aWRRRQT09](https://www.dosm.gov.my/v1/index.php?r=column/cthemByCat&cat=102&bul_id=OE5UdnNRL1VZbjJjcU1STVdIdFIUdz09&menu_id=TE5CRUZCb1h4ZTZMODZlbmk2aWRRRQT09)
- Evison, J. (2010). What are the basics of analyzing a corpus? In A. O'Keefe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 122-135). London: Routledge.
- Gries, Th. S. (2010). "Useful statistics for corpus linguistics", in A mosaic of corpus linguistics: selected approaches. A. Sánchez and M. Almela (Eds.), *Peter Lang: Frankfurt am Main, Germany* (pp. 269–291).
- Guillén -Nieto, V., C Helo Vargas -Sierra, Pardiño -Juan, M., Martínez -Barco, P. & Suárez – Cueto, A. (2008). Exploring State-of-the-Art Software for Forensic Authorship Identification. *International Journal of English Studies (IJES)*. 8(1), 1-28.
- Hansard Selangor. (2015). Dewan Negeri Selangor yang Ketiga Belas Penggal ke Tiga Mesyuarat Pertama, pp 9. Retrieved 1 March, 2019 from <http://dewan.selangor.gov.my/assets/pdf/Hansard/2015/Sesi-1/01.HANSARD.300315.pdf>
- Imran Ho Abdullah, Anis Nadiah Che Abdul Rahman & Azhar Jaludin. (2018). Malaysian Hansard Corpus: The Parliamentary Documents in Malay/ Malaysian Setting. Unpublished.
- Kawabena, S-K. (2018). It's about people: identifying the focus of parliamentary debates through a corpus-driven approach. *Corpora*. 13(3), 393–430.
- Malek M. A., Nor M. A. M. & Leong Y P. (2013). Water security and its challenges for Malaysia. IOP Conf. Ser.: Earth Environ. Sci. 16 012123.
- Marlyna Maros, Afendi Hamat & Nor Fariza Mohd Noor. (2017). eProsiding Kolokium Penerokaan Dataraya Kulturomik Malaysia. Digital Humanities Research Group UKM. 2017. Retrieved 1 March, 2019 form <https://culturomicsukm.files.wordpress.com/2017/12/e-proceeding-kolokium-kulturomik-2017.pdf>.
- McEnery, T. (2013). *Corpus: Some key terms. The ESRC Centre for Corpus Approaches to Social Science (CASS)*. Lancaster University, UK.
- Ngeow, Y. M., Kuang, C. H. & David, M. K. (2010). Gender stereotypes in Malaysian Parliament: A content analysis. In: Proceedings of 2010 International Conference on Humanities, Historical and Social Sciences (CHHSS 2010), 26-28 February, 2010, Singapore. (Unpublished). Retrieved 1 March, 2019 form <http://eprints.um.edu.my/3181/>

- Power, G. (2012). *Global parliamentary report: The changing nature of parliamentary representation*. Geneva, Switzerland : Inter-Parliamentary Union. New York : United Nations Development Programme.
- Rasyidah Md Khalid. (2018). Review of the Water Supply Management and Reforms Needed to Ensure Water Security in Malaysia. *International Journal of Business and Society*. 19(3), 472-483.
- Salazar, D. (2014). *Lexical Bundles in Native and Non-Native Scientific Writing*. Amsterdam: John Benjamins.
- Scott, M. (2008). WordSmith Tools version 5, Liverpool: Lexical Analysis Software.
- Syed Muhammad Shahabudin. (2015). Avoiding Water Crisis in Malaysia : Lessons for the Future in. *Water Malaysia The Malaysian Water Association (MWA) Quarterly*. 28, 3-8.
- Tan, K. H., Abdullah @ Ho Yee Beng, I. H., Zulkifli, N. A. & Mohd Shukor, S. M. (2017). Trend penggunaan bahasa samar dalam persidangan parlimen Malaysia. *GEMA Online® Journal of Language Studies*. 17(4), 84-100.
- The Official Portal of Parliament of Malaysia. (2018). The Official Portal of Parliament of Malaysia. Retrieved 1 March, 2019 from <https://www.parlimen.gov.my/index.php?&lang=en> .
- The Star Online. (2014a). Worst water shortage occurred in 1998 with 150 days of rationing. Retrieved 1 March, 2019 from <https://www.thestar.com.my/news/nation/2014/02/26/worst-water-shortage-occurred-in-1998-with-150-days-of-rationing/#LQzLV3yPXWyLi66Q.99>
- The Star Online. (2014b). Current water crisis caused by over development and lack of planning. Retrieved 1 March, 2019 from <https://www.thestar.com.my/news/community/2014/03/19/weather-not-to-blame-current-water-crisis-caused-by-over-development-and-lack-of-planning/#22ovRgzlMC4kR6Bz.99>
- Valvason, E. (2017). Mixed Voices: A Corpus-based Study of Hungarian Parliamentary Debates about the European Union. The 9th International Corpus Linguistics Conference at the University of Birmingham. Retrieved 1 March, 2019 from <https://www.birmingham.ac.uk/Documents/college-artslaw/corpus/conference-archives/2017/general/paper169.pdf>
- Vice, J. & Farrell, S. (2017). *The history of Hansard*. London: House of Lords Hansard and the House of Lords Library.
- Weng C. N. (2004). A Critical Review of Malaysia's Accomplishment on Water Resources Management Under AGENDA 21. *Malaysian Journal of Environmental Management* 5, 55-78.
- Xiao, R. & McEnery, T. (2006). Collocation, Semantic Prosody, and Near Synonymy: A Cross-Linguistic Perspective. *Applied Linguistics*. 27(1), 103-129.

## Appendix A

### TOP 10 AND 11 COLLOCATE LIST FOR “SELANGOR” IN MHC

No	Word P1	Relation	f	No	Word P4	Relation	f
1	SELANGOR	12.62954521	996	1	NEGERI	6.009408951	481
2	PERAK	8.827691078	87	2	WING	10.84006596	167
3	COMMERCE	8.784480095	73	3	SUM	10.02421474	165
4	STATE	6.346147537	58	4	CHEN	9.20233345	165
5	LABOUR	7.171729565	44	5	MICHAEL	10.76914406	164
6	SEMBILAN	9.354406357	37	6	RAJA	7.314871311	133
7	PAHANG	8.403063774	32	7	ISHAK	8.568361282	131
8	TRENGGANU	6.694219112	29	8	NASRON	8.872180939	97
9	JOHORE	7.566857338	28	9	PERAK	7.625770569	96
10	SULTAN	8.256325722	26	10	WILAYAH	6.68702507	67
				11	AIR	3	6
<b>P2</b>				<b>P5</b>			
1	STATE	6.686161995	153	1	NEGERI	6.194514751	443
2	PERAK	8.321522713	148	2	PERSEKUTUAN	6.991550446	124
3	SEMBILAN	9.549764633	95	3	WILAYAH	6.927373886	114
4	AFFAIRS	7.499453068	93	4	UMNO	8.559337616	72
5	PAHANG	8.3122015	62	5	PERAK	7.542766571	63
6	NEGRI	8.785064697	51	6	SEMBILAN	8.576512337	36
7	LABOUR	6.12132597	40	7	PINANG	6.566469193	36
8	KEDAH	6.987283707	39	8	JOHOR	6.2122159	35
9	MELAKA	6.463280678	38	9	BARAT	7.597874641	33
10	TRENGGANU	6.207083225	36	10	LAUT	6.418653011	29
				11	AIR	3.5966	12
<b>P3</b>				<b>P6</b>			
1	WING	10.68507195	89	1	NEGERI	6.029141903	285
2	SUM	9.404016495	89	2	WILAYAH	6.817350388	78
3	MICHAEL	10.66073322	88	3	PERAK	8.337896347	73
4	CHEN	8.033283234	86	4	PERSEKUTUAN	6.423883438	73
5	RAJA	6.802962303	84	5	SAI	9.716993332	52
6	ISHAK	8.073349953	75	6	KIM	9.310194969	52
7	STATE	6.45844841	75	7	LEE	7.38002634	52
8	CHIK	7.553602695	72	8	JOHOR	7.707167625	43
9	PERAK	6.523229122	53	9	PULAU	6.037938595	37
10	SEMBILAN	7.93559742	40	10	PAHANG	8.67684269	36
				11	AIR	4.37	8
<b>P7</b>				<b>P10</b>			
1	SELANGOR	12.3943224	1794	1	NEGERI	7.270579815	187
2	NEGERI	6.173442841	394	2	SUKAN	7.790607929	257
3	PERAK	9.097774506	98	3	BELIA	7.799220085	255
4	DARUL	10.60398579	92	4	AIR	6.928060532	36
5	EHSAN	12.30126095	84	5	PERSEKUTUAN	7.430027962	91
6	WILAYAH	6.886706352	74	6	WILAYAH	8.556547165	93
7	PERSEKUTUAN	6.376464844	74	7	PERAK	9.345826149	81
8	PULAU	6.591117859	68	8	LUMPUR	6.76533556	56
9	SUBRAMANIAM	8.061209679	63	9	JOHOR	7.15233326	59
10	PINANG	7.310675621	55	10	SUNGAI	6.495744705	46
11	AIR	5.450642586	9				
<b>P8</b>				<b>P11</b>			
1	NEGERI	20.97286797	675	1	NEGERI	7.017210484	212
2	DAN	18.05793953	537	2	TARIHEP	11.90260887	321
3	YANG	20.12871361	376	3	DAUD	11.84880543	321
4	PALANIVEL	11.78103352	363	4	HULU	9.220503807	320

5	PERAK	18.40701485	171	5	AIR	6.392780781	59
6	WILAYAH	8.206079483	137	6	SUNGAI	6.705155373	61
7	PERSEKUTUAN	6.936355591	133	7	LUMPUR	7.3737607	93
8	PULAU	18.60754776	131	8	PAHANG	9.129398346	67
9	BESAR	14.27156162	120	9	PERSEKUTUAN	6.488616467	77
10	JOHOR	15.75632763	118	10	JOHOR	7.094004631	85
11	AIR	13.89605236	12				
<b>P9</b>				<b>P12</b>			
1	NEGERI	7.173819542	765	1	AHMAD	8.911207199	363
2	ADNAN	12.36036396	175	2	DZULKEFLY	10.82863045	363
3	JAMALUDDIN	8.540949821	174	3	NEGERI	6.520933151	276
4	ABDOL	10.65914059	164	4	PANCHANATHAN	10.82731819	191
5	MULOK	10.64920807	164	5	PINANG	7.293728828	128
6	PERSEKUTUAN	7.434614182	119	6	MIC	7.677032948	188
7	WILAYAH	8.614062309	114	7	KEDAH	6.179595947	110
8	PERAK	9.564443588	100	8	SIKIT	6.282687187	72
9	AIR	6.619233131	97	9	PAKATAN	6.078561783	89
10	SUNGAI	6.898245811	86	10	PERAK	6.421103954	84
				11	AIR	5.987871647	73
<b>P13</b>							
1	IRMOHIZAM	10.9547348	724				
2	SEKINCHAN	9.797624588	14				
3	MAKAM	9.337225914	11				
4	PKNS	9.167301178	11				
5	IJOK	9.070439339	16				
6	IBRAHIM	8.979006767	738				
7	SPLASH	8.703296661	28				
8	KUALA	8.523644447	3032				
9	PENYALURAN	8.242791176	17				
10	EXCO	7.870017529	37				
11	AIR	6.902613163	74				

## ABOUT THE AUTHORS

Norsimah Mat Awal (PhD) is an Assoc. Professor at the Center for Literacy and Socio Cultural Transformation, Faculty of Social Sciences and Humanities, UKM. Her areas of interest are semantics, pragmatics and corpus linguistics.

Azhar Jaludin is a Senior Lecturer at Center for Literacy and Socio Cultural Transformation, Faculty of Social Sciences and Humanities, UKM and his area of expertise is corpus linguistics.

Anis Nadiah Che Abdul Rahman is a PhD candidate Center for Literacy and Socio Cultural Transformation, Faculty of Social Sciences and Humanities, UKM. She is currently pursuing PhD research in semantics and corpus linguistics.

Imran Ho Abdullah (Ph.D), ia Professor at Center for Literacy and Socio Cultural Transformation, Faculty of Social Sciences and Humanities, UKM. His areas of expertise are corpus linguistics and cognitive semantics and is currently the Deputy Vice-Chancellor (Industry & Community Partnerships), UKM.