# การวิเคราะห์คลังภาษาเพื่อการศึกษาวาทกรรมทางการเมือง

# ลลนา ปฐมชัยวัฒน์

สาขาวิชาภาษาอังกฤษ ครุศาสตรบัณฑิต คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยราชภัฏนครปฐม linchie2018@gmail.com

### บทคัดย่อ

จากการเปลี่ยนแปลงไปของศตวรรษ วิธีการวิจัยด้านคลังข้อมูลคอมพิวเตอร์ได้รับการยอมรับว่าเป็นการศึกษาใน รูปแบบใหม่ของงานวิจัยสาขามนุษยศาสตร์และสังคมศาสตร์ (HSS) (Baker et al, 2008, 221)(Caldas-Coulthard and Moon 2010. 99) (Mayaffre & Poudat, 2013, 65). การศึกษาวาทกรรมคลังภาษาหรือ CADS นี้ได้รับการพัฒนาในทั่วทุก มุมโลกอย่างต่อเนื่องเป็นเวลาหลายทศวรรษ โดยเฉพาะอย่างยิ่งวาทกรรมทางการเมืองของประเทศสหรัฐอเมริกา

การวิจัยในครั้งนี้ มีจุดประสงค์ดังนี้ คือ เพื่อศึกษาวาทกรรมของประธานาธิบดีประเทศสหรัฐอเมริกาอย่างเป็นระบบ โดยวาทกรรมสำคัญที่กล่าวขึ้นโดยประธานาธิบดีของสหรัฐอเมริกาผู้มีชื่อเสียงแห่งทศวรรษที่ 21 ในวันรับตำแหน่ง ประธานาธิบดี 3 ท่านเป็นข้อมูลทางวาทกรรมหลักงานของวิจัยนี้ ประธานาธิบดีทั้งสามท่านได้แก่ ประธานาธิบดีโอรัก โอ บามา ประธานาธิบดีดอนัลด์ ทรัมป์ และประธานาธิบดีโจ ไบเดน โดยมีคำถามวิจัย 2 ประการคือ 1) ลักษณะทางภาษาใดที่ ปรากฏในวาทกรรมทางการเมือง 2) อำนาจทางวาทกรรมเรื่องใดถูกสะท้อนให้เห็นได้จากบริบททางการเมือง โดยวิธีการ วิเคราะห์ข้อมูลเชิงปริมาณของคลังข้อมูลภาษาจากคอมพิวเตอร์ ทำให้ลักษณะทางภาษาที่สำคัญหลายประการรวมทั้งวจนกรรมได้ถูกอธิบายอย่างเป็นรูปธรรม

ในบรรดาการวิเคร<sup>า</sup>ะห์และการรวบรวมวรรณกรรม การอภิปรายผลการวิจัย ในครั้งนี้ ได้จากการพิจารณาผลลัพท์ที่ ได้ทางสถิติ และในแง่มุมของการพรรณนาจากอากัปกิริยาและพลังของคำพูดในขณะที่พูด ผลการวิจัยนี้ยังมีบทบาทสำคัญต่อ การศึกษาทั้งทางด้านภาษาศาสตร์เกี่ยวกับวาทกรรมทางการเมือง (PDA) และงานวิจัยที่เกี่ยวกับวิเคราะห์คลังภาษาต่อไป

คำสำคัญ: การวิเคราะห์คลังภาษา, ความถี่, สุนทรพจน์ทางการเมือง, การวิเคราะห์วาทกรรมทางการเมือง

## A Corpus-Driven Analysis for Political Discourse Analysis

#### Lalana Pathomchaiwat

English Education Department, Social Sciences and Humanities Faculty

Nakhon Pathom Rajabhat University

linchie2018@gmail.com

#### **Abstract**

According to a turn of the century, corpora are established as a modern platform of Humanities and Social Sciences (HSS) empirical researches (Baker et al, 2008, 221)(Caldas-Coulthard and Moon 2010. 99) (Mayaffre & Poudat, 2013, 65). A corpus-analysis discourse studies or CADS research have increasingly developed in several parts of the world for decades particularly in the US political scenarios.

The research major purpose is to explore the potential US presidential speeches systematically. Three essential residents' victory speeches conducted by three famous presidents in the 21<sup>st</sup> century were included for the political discourse analysis (PDA); the presidents of the United States, "Barack Obama", "Donald Trump", and "Joe Biden". The investigation was carried out serving two crucial research questions; 1) what lexical features are emerged from PDA and 2) What kinds of powers are reflected from the US political discourses. By means of the quantitative data analyzing, potential linguistic aspects and speech acts were explained concretely.

Among those corpus analysis and findings, the research discussions are instituted basing on the statistical results and descriptive aspects; speech acts and powers. The implications of this research play an important role on both the PDA in linguistic experiments and the corpus-driven analysis works extensively.

Keywords: Corpus Analysis, Frequency, Political Speech, Political Discourse Analysis, PDA

#### Introduction

A corpus analysis is known as 'Digital Corpora' that several linguistic scholars indicate it as an essential device in discourse analysis work, linguistic studies, corpus linguistics and political discourse (Mayaffire & Poudat, 2013, 66). Due to the corpus methods, the digital nature of the data is created and expressed effective explanation on language which complementing traditional reading (Mayaffire & Poudat (2013, 67).

In the concrete terms, Mayaffire and Poudat (2013, 67) point out that this digital technology is interpreted by the linkage of ideas and echoes of the language relevance. The corpus evidence is shown from word-for-word, linear or systematic reading like tables, indexes and list of words (Mayaffire & Poudat (2013, 67). A notable merit of political corpus can be specified in concise multidisciplinary dimensions regarding on linguistic materials like words, sentences, documents, and discourse (Mayaffire & Poudat, 2013, 65-66). It is said that the digital revolution of corpora in language studies provide a precious approach especially for the political corpora interpretation and evaluation (Mayaffire & Poudat (2013, 67). In a deeper focus, the corpus analysis in political discourse works is relied on the statistical techniques nevertheless it also requires the researchers' judgments and decisions for the text-input, the taken-

analysis and the cut-off points (Baker et al, 2008, 221-222). The qualitative discourse explanation is called for other essential details and considerations in the social, cultural, and political context (Weiss & Wodak, 2002, 223) which Sinclair (1994, 7) notes that "Language users cannot accurately report language usage, even their own. Consequently, such a form of PDA, the corpus study is crucial advantages that Willis (2017, 217) and Sinclair (1994, 7) pinpoint them as an uncover patterns in speech styles allowed to be analyzed via the corpus statistical technology. The value of using corpora in political context experiments can be indicated by several linguistic scholars like Johnson, Culpeper and Suhr (2003) whose work on the political terms were applied in three corpora of UK newspapers during 1994-1999 and the findings pointed out the noteworthy of Labour Party. Willis (2017) used corpus analysis to study about the UK politicians' talks about climate change under the topic; 'the 2008 Climate Change Bill', and his results showed the highlights of economic and technical issues going beyond the human and social dimensions. Moreover, Chen et al. (2019, 13) examined the linguistic styles in the presidential campaigns expressed by Hilary Clinton and Donald Trump's. And the textual analysis of corpora via AntConc expressed a significant diversity of lexical uses in both politicians' speeches. Similarly, Hamed (2020, 137) investigated the keywords and collocations in US presidents' discourses which expressed by Clinton, Bush, Obama and Trump. His corpus analysis data insisted for several similarity and differences among the pontifical speeches variously.

The value of PDA work relieves political phenomena be more explicable which is the same line as Pelinka (2007, 129), Chilton and Schaffer (1997, 206) insisted. The data found and expressed by means of political discourse can be substituted for the politicians' brightly foregrounds. In this study, the researcher also includes the descriptive analysis of speech acts and power for a higher level of succinct explanation in PDA. The speech acts proposed in this work is 'an illocutionary act' which is valued for the speakers' communicative purposes; representatives, expressives, verdictives, directives, commisives, and declarations. With a high regard to these pragmatic elements, the utterances in political conventions can be intensively explained (Hudri and April, 2018, 1). As the connection of language and contexts, the research data can be construed pragmatically based on linguistic forms and the users of them (Yule, 1996, 49), the literal meaning and pragmatic acts are essentially expressed via the political speeches purposively and functionally Yule, 1996, 49). Similarly, 'power' in PDA is independently associated and is proposed in this work. The power is as domination urged in a certain political situation which Karlberg (2005, 2) notes that it plays a significant role in human society in which the ways people think and talk as a result learning about 'power' is very meaningful for the peaceful cultivation (Karlberg (2005, 1). In this view, Wartenberg (1990, 5) insists for the PDA is necessary to pin audience down to the definition of 'power to' and 'power over'. Both distinct terms describe a set of phenomena of social powers which needs to be comprehended past political speech analysis.

Substantively, this PDA work revealed significant verification of how corpus-driven techniques and pragmatic descriptive analysis are beneficial in the political discourse. It also proclaimed displaying and analyzing statistically in order to answer two research questions; 1) what lexical features are emerged from PDA and 2) What kinds of powers are reflected from the US political discourses.

#### Methodology

Considering the research inquiries, the corpus-driven investigations were conducted using the quantitative analysis and the qualitative descriptions which were respectively explored.

This study employs corpus-driven approach to systematize the political speech data of the US leaders during the two decades, Barack Obama (2009-2017), Donald Trumps (2017-2021) and Joe Biden

(2021 up to present). The comprising of 12,232 words of the Victory Speech was purposively selected on account of the sharing of primary stage for their presidential appearance. It was the first day for the 44<sup>th</sup>, 45<sup>th</sup> and 46<sup>th</sup> US presidents delivered the foremost orations to American people. To complete the research investigation, the corpus of *LancsBox5.0* was utilized. The *Words* in *Lancsbox5.0* was considered to identify data relevant to the PDA. Interestingly, the corpus-driven analysis revealed the significant frequency toward the key notes as well as the remarkably concordances. The various dimensions of collocations in the statistic results imparted diverse political powers via the three presidential speeches.

#### Results and Discussions

In what follows, corpus linguistics and text statistics were carried out and the results pointed out several dimensions of the three US presidential speeches diversely. The frequencies and collocations below indicated distinct and similarity among the US politicians.

Joe Biden				
the	101			
and	96			
we	88			
of	77			
to	65			
а	46			
our	43			
in	42			
this	39			
i	33			

Barak Obama			
the	238		
and	204		
of	143		
to	127		
we	109		
our	94		
а	90		
that	89		
in	60		
is	55		

Donald Trump			
the	236		
and	196		
of	153		
to	147		
our	110		
in	98		
will	96		
i	82		
we	75		
have	69		

Figure 1 Top 10 Frequency Results of three US presidents' speeches analyzed by corpus

According to the corpus data, the top frequencies among the three presidents' speeches illustrate ten word frequencies they had uttered differently. For Joe Biden, the recent US president used 'the' as the most frequent for 101 times, then 'and', 'we', 'of', 'to', 'a', 'our', 'in', 'this' and the word 'I' was the least. Meanwhile Barack Obama, the 44<sup>th</sup> President's Victory Speech revealed 'the' as the most frequent, then 'and', 'of', 'to', 'we', 'our', 'a', 'that', 'in', and 'is' as the least. And the 45<sup>th</sup> President expressed 'the' as the most', then 'and', 'of', 'to', 'our', 'in', 'will', 'I', 'we' and 'have' as the least. In a regard to the corpus analysis, the function words like nouns, verbs and adjectives usually convey the speakers' intentions and they should be involved in the data discussions. Obviously, among these diversities, the function word; 'we' and 'our' were represented the most sharing word uses for all speeches. However, they were uttered variously in frequencies. In this closer view, the present US president, Joe Biden said 'we' as the third often and 'our' for the seventh comparing to other frequencies in his speech. In the same way, Barack Obama used 'we' and 'our' for relatively similar frequencies as the fifth and the sixth frequent word lists. However, Donald Trump uttered 'we' seventy-five times as the ninth frequent word and 'our' for more than a hundred as the fifth in his lists.

The diverse 'we' expressions showed different presidency viewpoints as the figurative signs to their US audience. As the democrat party members, two US presidents, Obama and Biden uttered 'we' for the equality and respect. The word, 'we' crucially urged the listeners' several heartfelt feeling and this referred to the aspects of 'Representatives' for stating and claiming. It also arouses the acts of

'Commisives' for promising and 'Declarations' of blessing. These speech act elements vitally highlighted their speech into the unity of America. The use of 'we' was also seen in Trump's victory speech but it was not often as comparing to Biden and Obama. Meanwhile Trump uttered 'our' for more often due to the homogeneity and it also illustrated speech acts in the same way as 'we'. In this case 'our' was used as 'Verdictives' for assessing, ranking and thinking according to the sense of belonging. This concisely showed the speakers' possession nonetheless the word 'our' also reflected the personals' self-concern at the same time. These evidences indicated all presidents' view on the remarkable 'Powers' of 'Unification of American citizen' and 'Ownership of the land' which both valued 'Nationalism' as the supreme goal.

Joe Biden		
come	5.292949	
can	5.231548	
do	5.05591	
things	5.029915	
must	4.971021	
other	4.707987	
stand	4.707987	
here	4.515341	
with	4.444952	
if	4.292949	

Barak Obama			
far	5.660961		
yes	5.660961		
day	5.11664		
can	5.03091		
change	4.853606		
were	4.853606		
hope	4.660961		
will	4.653141		
as	4.647155		
tonight	4.531678		

Donald Trump			
6.865699			
6.61416			
6.128733			
5.543771			
5.450661			
5.391768			
5.391768			
5.254264			
4.99123			
4.927099			

Figure 2 Top 10 Collocation Results of three US presidents' speeches analyzed by corpus

In these statistic results, there are significantly expressions of the collocation in three leader's speeches. Generally, a collocation referred to the words occurred nearby the certain word. In the above data, 'can' was preciously sharing collocations among the three presidents' orations while 'must', and 'don't' were considered interesting collocations. Precisely, 'can' was used with the second strongest locate for Joe Biden, the fourth locate for Barack Obama, and the eighth for Donald Trump. In this case, the number of word data was also needed to be considered. In Biden's talk, it comprised 2,388 words while Obama's was 4,521 words, and Trump's was 5,323 words. And these related to the 45<sup>th</sup> President, Donald Trump for saying 'can' as the most frequent in collocations however comparing to the others' total number of words, his 'can' was not the highest frequency in top ten collocation results. On the other hand, for Joe Biden, 'can' was the second collocated with other words in this speech as same as Barack Obama who emphasized this word long side with 'change' which was resemble as his presidential promotion campaign.

For the further discussions, the researcher proposed the following figures for more concise explanations especially for 'can', 'must' and 'don't'.



Figure 3 Concordance Results of Joe Biden

The first word with highest collocation frequency 'can' was highlighted as uttered by all US Presidents' speeches in regard to the linguistic features which indicated the positive meaning and interpretations variously. The examples of 'can' and collocations were "...can and should be...", "...can be and we must be...", "...can do great things.", "...can do this if we open...", "...can join forces, stop the shouting...", and "....can right wrong...." etc. The mentions of 'can' for in these concordancing samples proclaimed several pragmatic communications of illocutionary acts; 'Representatives' for stating, 'Verdictives' for assessing, ranking, thinking, and estimating, and the last act was 'Directives' for requesting, inviting, insisting, and suggesting. These linguistic features were interpreted statistically in the similar way to Barack Obama for the same word, 'can'.

Search we	Occurrences	10/109 (22.12) Texts 1 ▼ Corpus	Barak	▼ Context 7 ▼ Display Text
Index	File	Left	Node	A Right
36	Barak Obama	a peophe that "we shahh overcome". Ye	s, we	can. A man touched down on the
37	Barak Obama	she knows how America can change. Yes	, we	can. America, we have come so far.
31	Barak Obama	pressed on with that American creed: Yes	s, we	can. At a time when women's voices
34	Barak Obama	greatness and a democracy was saved. Yes	s, we	can. She was there for the buses
47	Barak Obama	up the spirit of a peophe: ye	, we	can. Thank you, God bhess you, and
33	Barak Obama	a new sense of common purpose. Yes	, we	can. When the bombs fehh on our
32	Barak Obama	out and reach for the bahhot. Yes	, we	can. When there was despair in the
38	Barak Obama	America can change, Yes, we can. America	, we	have come so far. We have seen
70	Barak Obama	is dignified. Where the answer is ye	, we	in- tend to move forward. Where the
35	Barak Obama	from Athanta who tohd a peophe the	at "we	shahh overcome". Yes, we can. A man

Figure 4 Concordance Results of Barack Obama

Attentively, the 44<sup>th</sup> US President, Barack Obama used '*can*' collocations in the form of cohesions; 'repetition' like "*Yes, we can*" in order to arouse the audience's interests. This was sharing the same types of illocutionary or speech acts to Biden's previous remarks; 'Representatives' for stating, 'Verdictives' for assessing, ranking, thinking, and estimating, and the last act was 'Directives' for requesting, inviting, insisting, and suggesting. Repeatedly, both the Democrats persuaded his audience 'Bravery' and 'Hope' which were considered the great powers in their presidential conventions.

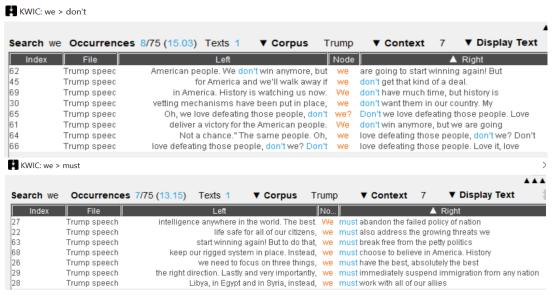


Figure 5 Concordance Results of Donald Trump

In focusing on this issue, the corpus data-driven interestingly identified the 45<sup>th</sup> US President speech as The Republican leader constantly delivered several denial views in his presidential speech. The communicative act of 'Verdictives' for assessing, thinking and estimating could be seen in his highest collocation word list like 'don't' and 'must'. According to the concordancing of 'don't' like "... don't get that kind of a deal...", "...don't have much time...", "...don't want them in our country....", "...don't win anymore...", and "Don't we love defending....". These sample concordances illustrated significant viewpoints of 'Verdictives', 'Directives', 'Commisives', and 'Declarations'. Beside the great expression of 'assessing, ranking, thinking, and estimating' of Verdictives, for 'Directives', they delivered 'commanding, daring, insisting, as well as suggesting'. Moreover, the 'Commisives Acts' could be related to 'threatening and promising'. The 'Declarations' comprised 'firing and arresting'. Crucially, many of word collocations in rejections occurred in Donald Trump's oration on his presidential speech day implied his concise ambitions ruling the country with the power of 'Authority'. According to the corpus analysis results, his oration urged the US citizens the appreciation and benefits of the US as the supreme goal. Again that the power of 'Patriotism' was aroused among the US similarly appeared in the speeches of the other presidents; Biden and Obama.

#### Conclusions

In regarding of the research inquiries, this study was involved both quantitative data via the corpora approach and the qualitative investigation in regard to the speech acts and powers. It significantly revealed in the important association among three political elements; actors, matters and purposes which 15 Graber (1981, 196) specially insists them as the essential parts of political discourses and the political communications.

The analysis of corpus frequencies in key words, collocations and concordances helped the research discussion be more concise. The descriptive data discussions were also systematically expressed and they indicated several interesting information via these three US leaders' speeches. The high frequency findings and discussions disclosed diversity between the two political parties in this county. While the Democrats value the mutual concessions with the diplomatically coaxing, the Republicans trended to be uncompromising with the appraising of the US accomplishment and victory. However, all US presidential

speeches were indistinguishably launched with a great attitude towards US powers like 'bravery', 'hope' and 'patriotisms'. These political powers were raised on account of the cooperation and the unity of America.

With a summary of this research, it was clear that the corpora analysis refined the PDA in the greater way. The frequencies of corpus-driven data strongly imparted statistic evidences for the effective political discourse exploration. Particularly, it was an important claim for implements of the corpora approach in the PDA as well as the other linguistic analysis works.

#### References

- Anton Pelinka. (2007) Language as a political category: the viewpoint of political science. Journal of Language & Politics. 6(1), 129-143.
- Carmen Rose Caldas-Coulthard and R, Moon. (2010) Curvy, hunky, kinky: using corpora as tools for critical analysis. Discourse & Society, 21(2), 99-133.
- Dalia Hamed. (2020). Keywords and collocations in US presidential discourse since 1993: a corpusassisted analysis. Journal of Humanities and Applied Social Sciences, 3(2), 137-158.
- Damon Mayaffre & Céline Poudat. (2013) Quantitative approaches to political discourse: corpus linguistics and text statistics. USA: John Benjamins Publishing Company.
- Doris A. Graber. (1981). Political Languages. In D. Nimmo, & K. Sanders (Eds.), Handbook of Political Communication (pp. 195-224). Beverly Hills, CA: Sage.
- Gilbert Weiss and Ruth Wodak. (2002) Critical Discourse Analysis: Theory and Interdisciplinary.

  London: Palgrave Macmillan.
- George Yule. (1996). Pragmatics. Oxford: Oxford University Press.
- Irwandi, Muhammad Hudri, Wanda April. (2018) An Analysis of Illocutionary Acts of Hillary Clinton's Concession Speech to Donald Trump in Presidential Election. Journal of English Language Teaching and Linguistics. 11(1), 1-4
- John Sinclair., Ed. (1994) Collins COBUILD English language dictionary. London: Collins.,7.
- Michael Karlberg. (2005) The Power of Discourse and the discourse of Power: Pursuing Peace Through Discourse Intervention. International Journal of Peace Studies. 10(1), 1-23.
- Paul Baker, et al. (2008) A useful methodological synergy? Combining Critical Discourse Analysis and Corpus Linguistics to examine discourses of refugees and asylum seekers in the UK press. Discourse Society, 19(3), 273-306.
- Paul Chilton and C. Schaffer. (1997) Discourse and politics. In T. van Dijk (ed.), Discourse as Social Interaction. Vol. 2. London: Sage, 206–31.
- Rebecca Willis. (2017) Taming the Climate? Corpus analysis of politicians' speech on climate change. Environmental Politics. 26(2), 212-231.
- Sally Johnson, Johnathan Culpeper and Stephanie Suhr. (2003) From "politically correct councilors" to "Blairite nonsense": Discourses of Political correctness in three British newspaper. Discourses and Society, 14(1), 29-47.
- Thomas E. Wartenberg. (1990) The Forms of Power: From Domination to Transformation. Philadelphia: Temple University Press, 1-23.
- Xueliang Chen, Yuanle Yan & Jie Hu. (2019). A corpus-based study of Hillary Clinton's and Donald Trump's linguistic styles. International Journal of English Linguistics. 9(3), 13-22.