

Linguistic Features of Donald Trump's Political Discourse: Focusing on Genre Differences and Change over Time

Woonhyung Chung*

1. Introduction

Donald Trump's victory in the United States 2016 presidential election could not have been foreseen. Most political pundits, critics, and even public opinion polls did not anticipate his win. While a number of factors have been mentioned as potential reasons for this unexpected election win, his unique use of language is said to be one of the most critical factors contributing to his victory (Sclafani, 2017). Trump describes himself as someone who knows how to use language and maintains that his unconventional language use greatly contributed to his being elected. His linguistic features are generally known for a narrow range of vocabulary, repetition of expressions, short sentences, and somewhat distracting compositions (Golshan, 2016; Liberman, 2016; Sclafani, 2017). These characteristics of his language use are often criticized for being unpresidential and a reflection of his low intellectual level and erraticism as well as his inability to focus (Hanson, 2018; Nerdwriter1, 2015;

* Doctoral Student, Department of English Linguistics and Language, Yonsei University

Pullum, 2016). In contrast, those characteristics are sometimes considered to make it easy and natural for the public to understand and remember his words, and could therefore be regarded as a clever rhetorical strategy (Colvin, 2016; Kayam, 2018; Sclafani, 2017; Simms, 2018; Swaim, 2015). On the other hand, occasionally he appears to follow the traditional norms of a president's discourse in official speeches, such as high-level vocabulary and sufficiently long sentences. This has often been accredited to speechwriters (or ghostwriters) and thus has been criticized (Lakoff, 2016; Savoy, 2017b), but his delivery is so natural that it is difficult to assess whether these speeches of higher quality are solely the work of speech- or ghostwriters. His delivery often gives the impression that he fully understands the scripts and that this mixture of easy and high-quality speeches is part of his strategy. Furthermore, some genre-specific studies (Savoy, 2017b; Wang & Liu, 2018) have revealed that unlike interviews, in campaign speeches, which are classified as the written communication genre, Trump has occasionally employed richer vocabulary and more complex sentences than other candidates or his predecessor. Furthermore, it has been suggested that his speeches have tended to become more sophisticated over time depending on circumstances (Lakoff, 2016; Wang & Liu, 2018).

The aim of this study was to provide a clearer basis for the controversy over Trump's language. By using various corpus analysis tools such as AntConc and WordSmith tools, the study analyzed and compared his political discourse with that of his three predecessors: Bill Clinton, George W. Bush, and Barack Obama. This study was conducted with three main goals. The first was to analyze his presidential discourse from three dimensions - lexical, syntactic, and cohesion levels - unlike previous studies that only evaluated the lexical dimension. So far, corpus

studies on his language have mostly concentrated on his use of specific words such as 'will,' 'going,' 'tremendous,' and 'millions and millions,' and only a few have analyzed his overall use of language. Thus, this study quantitatively analyzed the overall linguistic characteristics of his presidential political discourse with three different dimensions: lexical complexity, syntactic complexity, and text cohesion. The indices chosen were three lexical complexity indices (standardized type/token ratio, lexical density, and the proportion of big words), two syntactic complexity indices (Flesch-Kincaid grade level, and number of clauses per sentence), and two indices for text cohesion (semantic coreferentiality, and given/new information).

The second goal of the study was to analyze Trump's political discourse by genre by dividing it into interviews and formal speeches. Most of the prior corpus studies of Trump's political discourse consist of studies in one domain - such as interviews or campaign speeches - so there is little extant research which analyzes and compares his political discourse from different domains. Only a few studies have analyzed his political discourse by dividing it between an oral or a written communication genre (Savoy, 2017b; Wang & Liu, 2018). Since different language uses are observed by genre, the present work took the genre-specific approach by comparing Trump's interviews and formal speeches; through this, the study attempted to verify whether the prevailing assessment on Trump's language use is in line with both communication genres.

The final goal was to track the changes in Trump's rhetoric prior to and during his presidency. Since most of the previous studies analyzed his language use before his presidency, there is a research gap in the analysis of his language use during his term and a comparison between

the two periods. Previous corpus studies analyzing Trump's discourse have been limited mostly to his TV debates, campaign speeches (e.g., Dalman, 2017; Savoy, 2017b; Sclafani, 2017; Wang & Lui, 2018), or his inaugural address (Al-Saeedi, 2017; Chen, 2018; Tchapanian, 2017). Only a few prior studies have even mentioned changes in Trump's language use in any context. Lakoff (2016) and Savoy (2017b) briefly speculated that the way Trump uses language depends on the situation, while Wang and Lui (2018) merely confirmed that his language in the second half of his campaign tended to be more complicated. Therefore, this research takes his presidential speeches as the object of its study and quantitatively explores whether his language has changed over time.

In summary, this research divided his presidential discourse into interviews and formal speeches, compared them to that of his three predecessors from three language dimensions, and detected changes of his language use before and during his presidency. Through a large amount of data analysis, the present work intended to analyze his overall linguistic use, which are difficult to capture through qualitative research and present them objectively.

2. Literature Review

2.1. Previous Analyses

Since Donald Trump announced his candidacy for the presidency (16 June 2015), a wide range of corpus-based studies on his language have been conducted. In particular, there have been several studies comparing the political discourses of Hillary Clinton and Donald Trump in the runup

to the 2016 presidential election. Nerdwriter1 (2015) analyzed Trump's TV show interview and calculated what portion of monosyllabic words were in his one-minute answer. It found that among 220 words of his answer, 172 words (around 78%) consisted of one syllable. He further criticized his language in the interview as being at only 4.1 of the FK grade level, which is far lower than other politicians such as Hillary Clinton (7.7), and Bernie Sanders (10.1). In addition, a substantial number of studies revealed that Trump deployed a less diverse vocabulary during the campaign than Clinton (e.g., Savoy, 2017a; Aswad, 2019). Furthermore, Savoy (2017b) compared the relative frequency of part-of-speech (POS) between Trump and Clinton. He discovered that compared to Clinton's interviews, Trump's interviews deployed more verb phrases (verb and adverbs) which are more oriented toward action. In a similar vein, Aswad (2019) demonstrates that Trump significantly used more action-oriented constructs (or more call-to-action terms) than Clinton.

There are a substantial number of studies examining Trump's use of pronouns. Tchaparlan (2017) investigated his use of the personal pronoun in his inaugural speech, in which the personal pronouns 'we' and 'us,' and their possessive form 'our' were far more frequently used than other personal pronouns. Kvítková (2018) found not only the frequent use of 'we,' but also his predominant use of 'you.' The pronoun 'you' directly points to the audience which renders his speeches more personal. Nerdwriter1 (2015) indicated that Trump enjoys using imperative sentences such as, 'Look at Paris,' and 'Look at what happened in Paris,' to address the audience, and also prefers to use the second person pronoun 'you' on the premise that the audience already agrees with him, which can be seen in clauses such as 'you watch[ed] last night,' and 'you

see people talking.’ Aswad (2019) examined Hilary Clinton and Donald Trump campaign speeches and found that the latter preferred collective terms (e.g., ‘we,’ ‘us,’ ‘together,’ ‘our’) and words related to cooperation such as ‘teamwork’ and ‘sharing’ compared to Clinton, which gave a sense of unity and shared interests to his audience. Furthermore, Trump uses a communal charismatic rhetoric in his speech which forms a contrasting structure of ‘we’ versus ‘they’ in order to create a negative image of other parties such as ‘illegal immigrants,’ ‘refugees,’ and ‘adversarial media.’

A good deal of research on his frequent words and phrases has been conducted. The modal verb ‘will’ is one of his most favored words. Dalman (2017) analyzed Trump’s campaign speeches and found that the word ‘will’ had the highest frequency, and the word ‘going,’ which is a part of ‘be going to,’ was also frequently used. He explained that his repeated use of the words ‘will’ and ‘going’ is because he wants to “express strong aims, conviction, determination and promise” and to “persuade the electorate to vote for him.” Tchaparlan (2017) also explained that his excessive use of ‘will’ shows that he focuses on his future plans and his intention of connecting Americans to his plans. In addition, Kvítková (2018) discovered that in his speeches, the cluster ‘going to’ was a highly frequent collocation of the pronouns ‘you’ and ‘we.’ Simms (2018) contended that the lexical cluster ‘we’re going to’ plays the role of making him sound like a man of action rather than thought; of sharing his self-belief with his audience; and of giving the sense of a common undertaking. Simms also discovered other favorable words of Trump’s. He maintained that Trump uses a strategy of exaggeration using words such as ‘bigly,’ ‘huge,’ ‘millions and millions,’ and ‘a lot’ without actually providing any specific information. In

addition, Nerdwriter1 (2015) contended that Trump used only a small number of three-syllable words in a talk show, among which the word 'tremendous' accounted for a large portion. In his semantic-based analysis of campaign speeches Savoy (2017b) found that while Hillary Clinton prefers words in the category of "tenacity" such as 'is,' 'was,' 'has,' 'must,' and 'do,' which indicates the speaker's persistence, Trump employs more negative emotional words or expressions such as 'hate,' 'fear,' 'sad,' and 'war,' indicating that he does not follow the general assumption that a leader's charisma improves by showing positive emotions and a bright future.

In her qualitative study, Sclafani (2017) identifies the words that Trump seldom uses. She discovered that in the 2016 Republican Party presidential primaries, Trump rarely starts his sentence with the word 'well' compared to other Republican rivals, which makes him sound less devious, more straightforward, and more direct. Furthermore, in an interview with PBS NewsHour (Hendry, 2018), Sclafani noted the change in his speech style, declaring that his sentence structures were more complicated and the words he used were more abstract and metaphoric in the Union Speech in February 2018. However, in the same interview, Mark Liberman pointed out that those kinds of speeches are not necessarily written by Trump, so they might not fully show how he uses language.

Savoy (2017b) also mentions the possibility of the influence of ghostwriters in Trump's electoral speeches. Savoy divided the political discourse during the 2016 presidential election into two genres: the oral communication form (interviews and TV debates) and the written communication form (electoral speeches), revealing that the difference between genres was much larger than for Hillary Clinton. He argues that

this considerable difference indicates that Trump's speeches were written by a ghostwriter or a team of ghostwriters without him having close control over his writers. In short, Trump's oral form is less complex than Clinton's, with shorter sentences (Mean Sentence Length, MSL), lower informativeness values (Lexical Density, LD), less rich vocabulary (Type/Token Ratio, TTR), and less complicated words (i.e., 'big' words composed of six letters or more). In contrast, he found that unlike in the oral genre, Trump's written genre had higher values of LD and TTR and more big words than Clinton's written genre. Furthermore, when it comes to the stylistic or topical affinities between the two genres, the textual distance of Trump is clearly far larger than Clinton's.

Wang and Lui's study (2018) also analyzes Trump's political discourse during the 2016 election by genre: oral genre (debates) and written genre (campaign electoral speeches), and compared his TV debates and campaign speeches with those of Barack Obama and Hillary Clinton. Beyond investigation on the use of specific words and phrases, they conducted quantitative analyses on Trump's political discourse from three dimensions - lexical and syntactic complexities, along with the analysis of text cohesion. This resulted in a distinctive genre variation in lexical and syntactic complexities compared to Obama and Hillary. To be specific, his lexical diversity as measured by the moving-average type-token ratio and his readability level as measured by the Flesch-Kincaid (FK) test was lower than the other two in the case of debates, but a little higher than the other two in the case of campaign speeches. His campaign speeches especially, showed a higher level of complexities towards the end of the election campaign. However, interestingly enough, at a cohesion level they investigate thematic concentration (TC) and the result was that Trump does not show obvious genre differences. Moreover, whereas in

the case of debates there was no statistical difference among the three politicians, Trump's campaign speeches had more central themes than the others whose themes were more diverse. Wang and Lui explain that his high level of TC might be interpreted as his tendency toward authoritarianism, and his concentration on themes such as 'immigration,' 'terrorism,' 'security,' and 'job,' might be an important factor contributing to his election win.

Although there has been a wide range of research on the characteristics of Trump's language use as such, most studies have concentrated primarily on Trump's use of specific words or expressions, so there are relatively few studies addressing genre variation or with diverse dimensions. Therefore, there is a need for a more diversified index. Moreover, most of the research has only focused on Trump's discourse during the campaign period. Assumptions presented by previous genre-specific studies that Trump's discourse has become more linguistically complex depending on the situation or later in the election left the need for follow-up research to objectively measure his post-election language use.

To make up for this paucity, this study conducted corpus-based research at three different levels - the lexical, syntactic, and cohesion levels - on his political discourse after his election win, and compared it with his pre-presidency discourse. Based on prior genre-specific studies (e.g., Savoy, 2017b; Wang & Lui 2018), this study diversified measurements by adding some indices that have not been investigated in prior studies as new factors in order to gain a deeper understanding of his political discourse. More specifically, this study included the number of clauses per sentence (C/S) as a new index for measuring syntactic complexity to complement the shortcomings of Flesch-Kincaid (FK) grade

level, which identifies syntactic complexity based only on the number of syllables, words, and sentences. Furthermore, for the text level analysis, whereas the previous two genre-specific studies, Savoy (2017b) and Wang and Lui (2018), examined stylistic or topical affinities between the two genres and thematic concentration (TC), respectively, this study selected semantic coreferentiality (SC) and given/new information (GNI) as new factors to gauge cohesiveness between adjacent sentences, in order to add diversity to the analytical elements of existing studies, and to gain a more accurate measure to assess the strongest criticism of his speech, low cohesion. Through this investigation, this study intended to verify whether the prevailing criticism on Trump's use of simple words and sentence structures, and less cohesive content is in line with all communication genres by using objective figures derived from various index analyses. The study also sought to numerically verify the observations of previous studies that the complexity of his language has increased over time by comparing his discourses between before and during his presidency.

2.2. Theoretical Background

2.2.1. Lexical Complexity

Lexical complexity can be measured for various aspects of word properties, including lexical variation, density, and sophistication. This study gauges the standardized type/token ratio to measure lexical variation, the proportion of lexical words to measure lexical density, and the proportion of big words to measure lexical sophistication.

1) Standardized Type/Token Ratio (STTR)

The STTR was chosen as the first index in order to assess how diverse

words were used in political texts. The STTR is a concept that developed from TTR, which was originally devised by Scott (1997). The TTR is the percentage of the number of different words (types) divided by the number of running words (tokens). A higher value of the TTR indicates that more diverse words are used in a text. However, this value can contain a distortion because the number of running words - or text length - which is unrelated to lexical diversity, can substantially affect it. To remedy this issue, Scott devised the STTR in which each text is divided into a certain number of words, i.e. 1,000 words, and then the average value of TTRs for each text is calculated so that the STTR can be an independent value from the text length.

2) Lexical Density (LD)

Lexical density was measured as the second index to estimate the informativeness of political texts from the aspect of lexicon. The LD is the percentage ratio between the number of lexical words and the total number of word tokens. Lexical words, which are also referred to as content words, are composed of nouns, adjectives, verbs, and adverbs. Unlike function words (e.g., determiners, modals, auxiliaries), whose role is to indicate grammatical relationships with and links to other words, lexical words play the main role in providing information and enriching the meanings in a text. Therefore, a higher ratio of content words in a text, or higher value of the LD, indicates that the text contains more information and its contents are more complex.

3) Big Words (BWs)

The third index is the proportion of big words (denoted BWs) which are words with six letters or more. BWs calculation is done by dividing

the number of big words by the total number of words. Savoy (2017b) selected the proportions of BWs as an index in comparing Trump's and Clinton's rhetoric during the 2016 president election under the theory that basic words are relatively short and complex words are the opposite so that the high percentage of big words indicates a higher lexical difficulty and complexity of a text.

2.2.2. Syntactic Complexity

This study chose two indices to gauge syntactic complexity. One is the FK grade level, which has frequently been used in previous studies (e.g., Nerdwriter1, 2015; Wang & Lui, 2018); the other is C/S, which measures language use not considered by the FK grade level.

1) Flesch-Kincaid Grade Level (FK Grade Level)¹⁾

The Flesch-Kincaid readability test (Kincaid, Fishburne, Rogers & Chissom, 1975) was conducted to gauge the FK grade level of each political text. The value of the FK grade level relies on two factors; the mean length of words and sentences. The following is the formula for calculating the FK grade level:

$$\text{FK grade level} = 0.39 \times \frac{\text{the total number of words}}{\text{the total number of sentences}} + 11.8 \times \frac{\text{the total number of syllables}}{\text{the total number of words}} - 15.59$$

The reason why the FK grade level was chosen as a syntactic index is that various research and articles have assessed Trump's political

1) Although there was no complete agreement on whether the FK grade level fell into the category of the syntactic complexity, the previous study that was used as the model for this study (Wang & Lui, 2018) included this index under the syntactic complexity category, so the study followed this categorization. Therefore, the more general syntactic complexity indices need to be included in further studies.

discourse using this grade level (e.g., 'his speech is the fourth grade') so that the value of the FK grade level enables us to compare previous assessments with the results of this study.

2) The Number of Clauses per Sentence (C/S)

Unlike the FK grade level, which simply shows the number of words within a sentence, C/S gauges the number of clauses per sentence to show the length of different types of clauses including coordination and subordination, providing a more detailed measurement of the syntactic complexity of a text. Previous research on text complexity corroborates that this index is one of the most reliable predictors of syntactic complexity in learner texts (e.g., Lu, 2011).

2.2.3. Text Cohesion

Text cohesion indicates how naturally sentences are connected to each other in a given text. It is an important device for a writer or speaker (in the case of a transcript) to deliver the main message. A low level of text cohesion reflects a lack of proficiency of the writer/speaker as well as shallow understanding of the given subject. Moreover, a text lacking cohesiveness makes it less persuasive, or worse, difficult to understand. In this study, two aspects of text cohesion were measured: semantic coreferentiality and given/new information.

1) Semantic Coreferentiality (SC)

Semantic coreferentiality refers to how often a text contains words with semantic similarity. Cohesive text tends to show greater semantic similarity between words. This is because the sudden occurrence of a word that is not meaningfully related to adjacent words, especially

without justifiable reason, may reduce text cohesion and reflect a writer/speaker's scattered thought process.

Semantic coreferentiality can be measured by using Latent Semantic Analysis (LSA) (Landauer, McNamara, Dennis & Kintsch, 2007), which is a mathematical and statistical technique for gauging semantic overlap between words, sentences, and paragraphs. It calculates how frequently each word appears within the text, thereby producing weighted vectors of each word within the text and forming a basis on which the relationship between the vectors represents semantic similarity between each word. For example, for the word 'flower,' the word 'tree' has a higher LSA value than the word 'robot.' The Coh-Metrix tool (McNamara, Graesser, McCarthy & Cai, 2014) automatically calculates the SC figure of a text using LSA and represents the figure from 0 (low cohesion) to 1 (high cohesion).

2) Given/New Information (GNI)

Given/new information reflects one aspect of text cohesion. The appropriate use of given information lessens the cognitive load of the audience which helps them to easily understand the message. By using the same algorithm for obtaining SC, LSA achieves a weighted vector value for each sentence and calculates the proportion of new information. The Coh-Metrix tool automatically calculates this value from 0 (low cohesion) to 1 (high cohesion).

As such, the study sought to verify criticism of Trump's linguistic use from three different dimensions, attempting to make more detailed quantitative observations by adding additional indices to each dimension from previous studies (e.g. Savoy, 2017b; Wang & Lui, 2018).

3. Method

3.1. Data Collection

The data was collected mainly from the American Presidency Project (www.presidency.ucsb.edu) and the Miller Center (www.millercenter.org). The American Presidency Project is a free online resource which contains various types of transcripts ranging from presidential addresses, to interviews, to campaign speeches. The Miller Center posts transcripts of formal presidential speeches of both previous and current American presidents. Moreover, since these two websites do not provide transcripts of interviews prior to presidency, this type of transcript was extracted from the websites of broadcasting companies such as CNN, the Washington Post, and the New York Times.

The total number of corpora used in this study was 575,608 words. First, the first target corpus and the reference corpus were constructed, with a total number of word tokens of approximately 43 million words, with the aim of comparing Trump's presidential discourse with that of his predecessors. The first target corpus is a collection of Donald Trump's thirty-two transcripts, with a total of 91,715 words. The reference corpus is a collection of ninety-six transcripts of the latest three previous American Presidents: Bill Clinton, George W. Bush, and Barack Obama. Its composition is thirty-two transcripts per each president and the total number of tokens is 338,143 words. Each corpus consists of two sub-corpora for investigating the differences in oral (interviews) and written (presidential speeches) genres. Second, in order to examine the diachronic change of Trump's linguistic features, an additional target corpus which contains Trump's political discourse before his presidency

was built (145,750 words). The organization of the corpus is shown in Table 1.

Table 1. The composition of the corpus

		# of texts		# of tokens	
		oral	written	oral	written
Reference Corpus	Bill Clinton	16	16	72,653	51,239
	George W. Bush	16	16	57,729	40,679
	Barack Obama	16	16	59,032	56,811
Target Corpus 1	Donald Trump	16	16	51,745	39,970
Target Corpus 2	Donald Trump before president	16	16	78,888	66,862
Total number		80	80	320,047	255,561
		160		575,608	

As shown in Table 1, an equal number of texts (16) was extracted from the corpus of each president for a better comparison. The total numbers of tokens per president were not statistically different. The one-way ANOVA for interviews does not reveal any significant differences among the text lengths of the four presidents, $F(3, 60) = 1.762$, $p = .164$, $\eta_p^2 = 0.081$, while the one-way ANOVA for speeches reveals a slight difference among text lengths of the four presidents, $F(3, 60) = 2.775$, $p = .049$, $\eta_p^2 = 0.122$, but post-hoc tests (Bonferroni correction) did not find any significant difference among the four presidents (Clinton vs. Bush, $p = 0.817$; Obama vs. Bush, $p = 0.147$; Obama vs. Clinton, $p > .1$; Trump vs. Bush, $p > .1$; Trump vs. Clinton, $p = 0.673$; Trump vs. Obama, $p = 0.115$).

When this study constructed the corpus, complete texts were used. Some noise such as descriptions of audience's reactions, for example, applause and laughter were erased. Moreover, remarks from interviewers

and other interviewees were also erased.

3.2. Data Analysis

To distinguish Trump's linguistic features from that of his predecessors, this study conducted the analyses from three different dimensions: lexical, syntactic, and cohesion levels. Various quantitative measurements were employed related to these three levels: STTR, LD, and BWs at the lexical level; FK grade level and C/S at the syntactic level; and SC and GNI at the cohesion level. Several tools were utilized in a complementary manner. WordSmith Tools 6.0 (Scott, 2011) was used to extract the STTR and the portions of BWs. AntConc 3.5.8 (Anthony, 2019) was used to calculate LD. A number of web-based tools were also utilized to achieve numeral values from several aspects: syntactic complexity analyzer (<https://aihaiyang.com/software/l2sca/single/>) for C/S; readability test tool (www.webfx.com/tools/read-able) for the FK grade level; and Coh-Metrix (<http://cohmetrix.com/>) for SC and GNI. A one-way analysis of variance (ANOVA) test was used to inspect statistical differences among the texts from the presidents. In the case of a significant difference, pairwise comparisons using Bonferroni tests were conducted to pinpoint the locus of the difference. All statistical analyses were conducted using SPSS.

4. Results and Discussion

4.1. Lexical Complexity

4.1.1. Standardized Type/Token Ratio (STTR)

As shown in Table 2 and Figure 1, results of the descriptive statistics in interviews show that Trump has the lowest value ($M = 32.31$, $SD = 1.32$) and Obama has the highest value ($M = 40.06$, $SD = 1.39$). The one-way ANOVA for interviews revealed a significant difference among the four presidents in the STTR, $F(3, 60) = 113.001$, $p < .001$, $\eta_p^2 = 0.850$. As shown in Table 3, post-hoc tests (Bonferroni correction) showed that Trump has significantly lower values in comparison with other presidents, indicating that his interviews contained a significantly lower number of words.

In speeches, results of the descriptive statistics show that Trump has the highest value ($M = 45.07$, $SD = 5.55$), but the one-way ANOVA did not reveal a significant difference among these four presidents, $F(3, 60) = 2.263$, $p = .090$, $\eta_p^2 = 0.102$. This result suggests that unlike in the interviews, Trumps used as many words as the other presidents did in speeches.

When comparing Trump's interviews and speeches between the two periods, no significant difference existed in the STTR value of Trump's interviews for the two time periods ($p = .620$), but there was a significant difference in his speeches ($p = .004$) (see Table 3). These results indicate that Trump's speeches contained a greater number of words during his presidency compared to during the campaign.

Table 2. Descriptive statistics of the STTR of presidents

Presidents	Interviews		Speeches	
	Mean	Std. Deviation	Mean	Std. Deviation
Clinton	36.60	1.04	43.21	1.94
Bush	38.21	1.20	42.32	1.51
Obama	40.06	1.39	44.21	1.86
Trump	32.31	1.32	45.07	5.55
Trump before	32.02	1.85	39.71	3.90

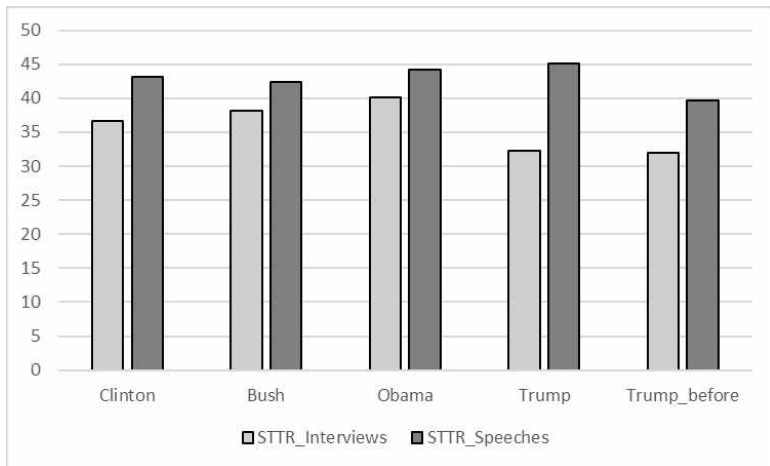


Figure 1. Mean values of the STTR of presidents (%)

Table 3. P-values from the comparisons of the STTR between Trump and other presidents

		Clinton	Bush	Obama	Trump before presidency
Trump	Interviews	< .001	< .001	< .001	.620
	Speeches	.618	.103	> .1	.004

4.1.2. Lexical Density (LD)

As shown in Table 4 and Figure 2, results of the descriptive statistics in interviews show that Trump has the lowest value in LD ($M = 36.72$, $SD = 3.29$). The one-way ANOVA revealed a significant difference

among the four presidents, $F(3, 60) = 20.152$, $p < .001$, $\eta_p^2 = 0.502$. As shown in Table 5, post-hoc tests (Bonferroni correction) showed that Trump had significantly lower values than other presidents.

In speeches, Bush has the highest value ($M = 47.80$, $SD = 3.04$), followed by Trump ($M = 46.65$, $SD = 3.55$), and then by Obama ($M = 46.02$, $SD = 1.94$). Clinton has the lowest value ($M = 45.08$, $SD = 1.96$). The one-way ANOVA revealed a significant difference among the four presidents, $F(3, 60) = 2.818$, $p = .047$, $\eta_p^2 = 0.124$. Post-hoc tests (Bonferroni correction) showed that the difference was found between Clinton and Bush ($p = .038$), not between Trump and any other presidents (See Table 5). These results indicate that the lexical diversity was lower for Trump than the other presidents in the interviews, but not in the speeches.

For the comparison of Trump's interviews and speeches between the two periods, there was no significant difference in LD between before and during his presidency in both interviews ($p = .248$) and speeches ($p = .246$) (see Table 5).

Table 4. Descriptive statistics of LD of presidents

Presidents	Interviews		Speeches	
	Mean	Std. Deviation	Mean	Std. Deviation
Clinton	38.83	1.32	45.08	1.96
Bush	41.89	1.43	47.80	3.04
Obama	40.51	1.07	46.02	1.94
Trump	36.72	3.29	46.65	3.55
Trump before	37.75	1.15	45.30	2.90

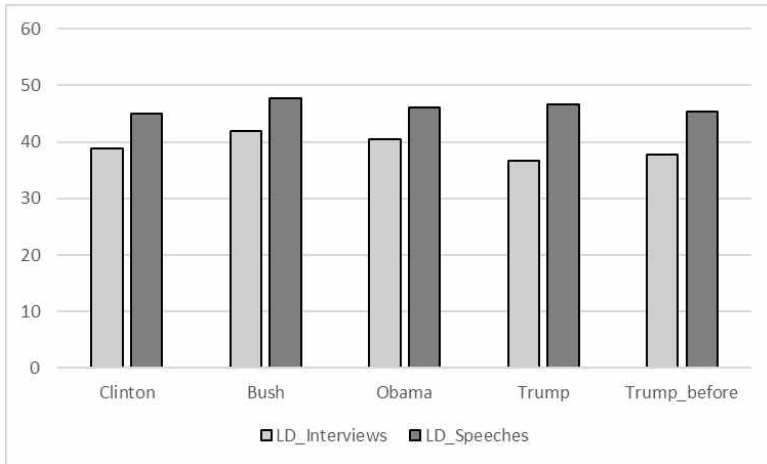


Figure 2. Mean values of LD of presidents (%)

Table 5. P-values from the comparisons of LD between Trump and other presidents

		Clinton	Bush	Obama	Trump before presidency
Trump	Interviews	.023	< .001	< .001	.248
	Speeches	.641	> .1	> .1	.246

4.1.3. Big Words (BWs)

Similar to results of the STTR and LD, the results of the descriptive statistics in interviews show that Trump has the lowest value ($M = 20.86$, $SD = 1.44$) (See Table 6 and Figure 3). The one-way ANOVA for interviews revealed a significant difference among the four presidents, $F(3, 60) = 50.214$, $p < .001$, $\eta_p^2 = 0.715$. As shown in Table 7, post-hoc tests (Bonferroni correction) showed that Trump has significantly lower values than other presidents.

For speeches, Obama has the highest value ($M = 32.22$, $SD = 3.03$), followed by Trump ($M = 32.18$, $SD = 4.16$). However, the one-way

ANOVA revealed no significant difference among the four presidents, $F(3, 60) = 0.422$, $p = .738$, $\eta_p^2 = 0.021$. These results suggest that for the use of big words, Trump was indistinguishable from the other presidents in speeches.

As shown in Table 7, there was no significant difference in the portion of big words of Trump's interviews for the two time periods ($p = .695$) but there was a significant difference in the portion of big words of Trump's speeches for the two time periods ($p = .047$), indicating a more use of big words after than before election.

Table 6. Descriptive statistics of BWs of presidents

Presidents	Interviews		Speeches	
	Mean	Std. Deviation	Mean	Std. Deviation
Clinton	24.61	1.15	31.03	2.40
Bush	27.16	2.12	31.67	3.85
Obama	26.47	1.51	32.22	3.03
Trump	20.86	1.44	32.18	4.16
Trump before	21.88	1.89	29.06	4.35

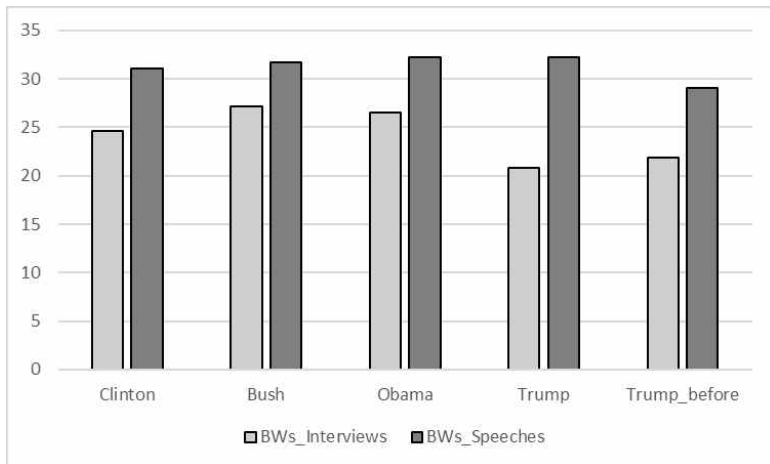


Figure 3. Mean values of BWs of presidents (%)

Table 7. P-values from the comparisons of BWs between Trump and other presidents

		Clinton	Bush	Obama	Trump before presidency
Trump	Interviews	< .001	< .001	< .001	.095
	Speeches	> .1	> .1	> .1	.047

The of lexical complexity revealed a clear difference between Trump's interviews and speeches. His lexical complexity in the interviews turns out to be significantly simpler than his three predecessors in all three measurements. The low STTR and LD values validate that he uses a lot of repetitive vocabulary and uses fewer content words, which indicates that he conveys less amount of information during interviews. This is in line with the existing research findings and the public assessment of Trump's language. Moreover, Trump is known to repeatedly use his favorite big words such as 'tremendous,' 'beautiful,' and 'incredible'; nevertheless, his overall use of BWs was still limited. Example (1) shows that Trump rarely uses words that contain big words except for 'tremendous.' This is in line with Nerdwriter1's findings (2015) that about 78% of the total words were one-syllable words in one of his interviews.

- (1) First thing I brought up was North Korea. Can he help us with North Korea? Because we can't allow it and it's no good for you and you have a *tremendous* power because of train. He then explains thousands of years of history with Korea. Not that easy. In other words, it's not as simple as people would think. They've had *tremendous* conflict with Korea over the years. Now his father was in China four times. He was never in China. So I said look you have a *tremendous* power because of trading through the border.

(Trump_Interviews_April 11, 2017_Interview with Maria Bartiromo on Fox Business Network)

However, his speeches were different from interviews. In all three measures, there was no significant difference from his predecessors. He had the highest value (in the STTR) or second highest values (in LD and BWs). These results demonstrate that he employs a wider range of vocabulary, content words, and big words in speeches than in interviews. These results are in line with prior findings suggesting Trump's use of more tokens of words with greater diversity than those of Hillary Clinton and Obama, and more use of big words than Hillary Clinton in the written form (Savoy, 2017b; Wang & Lui, 2018). Moreover, in all three measurements, he showed the difference between before and during his presidency, supporting previous findings that Trump uses different language depending on the situation, and also his language use has changed over time (Lakoff, 2016; Savoy, 2017b; Wang & Lui, 2018).

4.2. Syntactic Complexity

4.2.1. Flesch-Kincaid Grade Level (FK Grade Level)

Results of the descriptive statistics in interviews show that Trump has the lowest value ($M = 4.66$, $SD = 0.87$) (See Table 8 and Figure 4). The one-way ANOVA revealed a significant difference among the four presidents in the FK grade level, $F(3, 60) = 56.508$, $p < .001$, $\eta_p^2 = 0.739$. As shown in Table 9, post-hoc tests (Bonferroni correction) showed that Trump had significantly lower values than other presidents.

Results of the descriptive statistics in speeches show that Clinton has the highest value ($M = 10.63$, $SD = 0.70$), followed by Obama ($M =$

10.11, $SD = 1.58$), and then by Bush ($M = 9.54$, $SD = 1.22$). Trump has the lowest value ($M = 9.09$, $SD = 1.71$). The one-way ANOVA revealed a significant difference among these four presidents, $F(3, 60) = 3.860$, $p = .014$, $\eta_p^2 = 0.162$, but post-hoc tests (Bonferroni correction) showed there was only the difference between Trump and Clinton ($p = .013$) (See Table 9).

As shown in Table 9, there was no significant difference in Trump's FK grade level for interviews between before and during his presidency ($p = 0.598$) but there was a significant difference in Trump's FK grade level for speeches between before and during his presidency ($p = .048$).

Table 8. Descriptive statistics of the FK grade level of presidents

Presidents	Interviews		Speeches	
	Mean	Std. Deviation	Mean	Std. Deviation
Clinton	8.68	0.98	10.63	0.70
Bush	8.11	1.02	9.54	1.22
Obama	9.65	1.60	10.11	1.58
Trump	4.66	0.87	9.09	1.71
Trump before	4.86	1.16	7.72	2.05

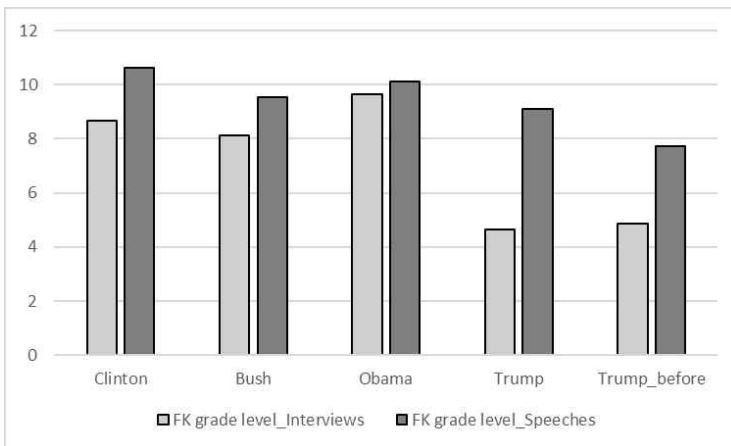


Figure 4. Mean values of FK grade level of presidents

Table 9. P-values from the comparisons of the FK level between Trump and other presidents

		Clinton	Bush	Obama	Trump before presidency
Trump	Interviews	< .001	< .001	< .001	.598
	Speeches	.013	> .1	.234	.048

4.2.2. The Number of Clauses per Sentence (C/S)

As shown in Table 10 and Figure 5, results of the descriptive statistics in interviews show that Obama has the highest value ($M = 2.83$, $SD = 0.34$), followed by Clinton ($M = 2.61$, $SD = 0.18$), and then by Bush ($M = 2.31$, $SD = 0.28$). Trump has the lowest value ($M = 2.03$, $SD = 0.27$). The one-way ANOVA revealed a significant difference among these four presidents, $F(3, 60) = 26.342$, $p < .001$, $\eta_p^2 = 0.568$. As shown in Table 11, post-hoc tests (Bonferroni correction) showed that Trump had significantly lower values than other presidents.

In speeches, results of the descriptive statistics show that Trump has the lowest value ($M = 1.61$, $SD = 0.20$) and Obama has the highest value ($M = 2.00$, $SD = 0.20$). The one-way ANOVA revealed a significant difference among these four presidents, $F(3, 60) = 16.939$, $p < .001$, $\eta_p^2 = 0.459$. As shown in Table 11, post-hoc tests (Bonferroni correction) showed that Trump had significantly lower values than the other presidents. Overall, Trump used a fewer number of clausal coordination and subordination in both interviews and speeches compared to the other presidents.

As shown in Table 11, there was no significant difference in Trump's C/S in both interviews ($p = .109$) and speeches ($p = .146$) between the two time periods.

Table 10. Descriptive statistics of C/S of presidents

Presidents	Interviews		Speeches	
	Mean	Std. Deviation	Mean	Std. Deviation
Clinton	2.61	0.18	1.20	0.19
Bush	2.31	0.28	1.81	0.14
Obama	2.83	0.34	2.00	0.20
Trump	2.03	0.27	1.61	0.20
Trump before	2.25	0.45	1.71	0.19

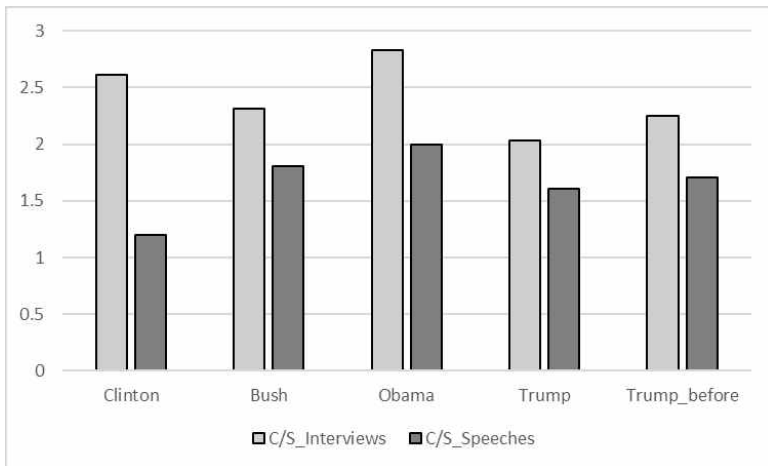


Figure 5. Mean values of C/S of presidents

Table 11. P-values from the comparisons of C/S between Trump and other presidents

		Clinton	Bush	Obama	Trump before presidency
Trump	Interviews	< .001	.032	< .001	.109
	Speeches	< .001	.014	< .001	.146

According to the analyses of results of lexical and syntactic complexity - STTR, LD, BWs, and FK grade level, the general values were all lower in interviews than in speeches. This contrast may stem from the difference in genre. The oral communication genre, i.e. interview, is

generally less complicated than the written communication genre, i.e. speech. However, in the case of C/S, the values in interviews were higher than in speeches. This opposite pattern may be attributed to the frequent use of prefabricated chunks such as ‘I think,’ ‘I believe,’ ‘I mean,’ and ‘you know’ during interviews. Example (2) illustrates this tendency: Trump’s interview (2a) has lower FK grade level than Trump’s speeches (2b) (5.0, 10.3, respectively), but the C/S value of (2a) is higher than that of (2b) (3.5, 1.67, respectively).

- (2) a. No, but it’s getting close. They want, I feel, *you know*, I have great feeling for DACA. *I think* that we should be able to do something with DACA. *I think* it’s foolish if we don’t, they’ve been here a long time, they’re no longer children, *you know*. People talk of them as children, *I mean* some are 41 years old and older. But some are in their teens, and late teens, but nevertheless *I think* we should do something with DACA and *I think* we should do something to help people. (Trump_Interviews_January 11, 2018_Interview with the Wall Street Journal)
- b. Tonight I ordered a targeted military strike on the airfield in Syria from where the chemical attack was launched. It is in this vital national security interest of the United States to prevent and deter the spread and use of deadly chemical weapons. There can be no dispute that Syria used banned chemical weapons, violated its obligations under the Chemical Weapons Convention and ignored the urging of the U.N. Security Council. (Trump_Speeches_April 6, 2017_Remarks on United States Military Operations in Syria from Palm Beach, Florida)

For the syntactic complexity of interviews, both Trump's FK grade level and C/S were found to be lower than his predecessors. This is consistent with the result of lexical complexity. For the syntactic complexity of speeches, however, the FK grade level of Trump was on par with Bush and Obama, with only Clinton having a significantly higher value. In the case of C/S, Trump had a significantly lower value than his three predecessors. These results suggest that his speeches may be less detailed in the use of clausal units than the other presidents, but not in terms of the number of syllables, words, and sentences. This pattern is reflected in Example (3) where Trump's FK grade level is far higher than Obama's (14.1 vs. 9.1), but Trump's C/S is lower than Obama's (1.25 vs 1.71).

- (3) a. As we lift restrictions and expand authorities in the field, we are already seeing dramatic results in the campaign to defeat ISIS, including the liberation of Mosul in Iraq. Since my inauguration, we have achieved record-breaking success in that regard. We will also maximize sanctions and other financial and law enforcement actions against these networks to eliminate their ability to export terror. When America commits its warriors to battle, we must ensure they have every weapon to apply swift, decisive, and overwhelming force. (Trump_Speeches_August 21, 2017_Address to the Nation on United States Strategy in Afghanistan and South Asia From Joint Base Myer-Henderson Hall, Virginia)
- b. Around the world, young people are moving forward hungry for a better world. Around the world, in small places, they're overcoming hatred and bigotry and sectarianism. And they're learning to respect each other, despite differences. The people

of the world now look to us, here, to be as decent and as dignified and as courageous as they are trying to be in their daily lives. And at this crossroads, I can promise you that the United States of America will not be distracted or deterred from what must be done. We are heirs to a proud legacy of freedom, and we're prepared to do what is necessary to secure that legacy for generations to come. I ask that you join us in this common mission, for today's children and tomorrow's. (Obama_Speeches_September 24, 2014_Remarks to the United Nations General Assembly in New York City)

Combining the results of lexical and syntactic complexity, we can interpret that the level of Trump's language complexity is significantly higher in speeches than in interviews in general. Furthermore, the greater complexity in Trump's language use after than before election, which was more prominent in speeches than in interviews, supports the previous results that he changes his language use depending on situations, and that his language use has changed over time (Lakoff, 2016; Savoy, 2017b; Wang & Lui, 2018).

4.3. Syntactic Complexity

4.3.1. Semantic Coreferentiality (SC)

As shown in Table 12 and Figure 6, results of the descriptive statistics in interviews show that Trump has the highest value ($M = 0.13$, $SD = 0.02$) with Clinton ($M = 0.13$, $SD = 0.01$). The one-way ANOVA revealed a significant difference among the four presidents in SC, $F(3, 60) = 6.747$, $p = .001$, $\eta_p^2 = 0.252$. Post-hoc tests (Bonferroni correction) showed that Trump's value is significantly higher than

Obama's value only ($p = .007$) (See Table 13).

Similar to interviews, results of the descriptive statistics in speeches show that Trump has the highest value ($M = 0.17$, $SD = 0.07$). The one-way ANOVA revealed a significant difference among these four presidents, $F(3, 60) = 3.174$, $p = .031$, $\eta_p^2 = 0.137$, but no significant difference was found in pairwise comparisons (See Table 13).

As shown in Table 13, there was no significant difference in Trump's SC between in both interviews ($p = .334$) and speeches ($p = .099$) between the two time periods.

Table 12. Descriptive statistics of SC of presidents

Presidents	Interviews		Speeches	
	Mean	Std. Deviation	Mean	Std. Deviation
Clinton	0.13	0.01	0.13	0.02
Bush	0.11	0.02	0.16	0.03
Obama	0.10	0.02	0.13	0.03
Trump	0.13	0.02	0.17	0.07
Trump before	0.11	0.04	0.14	0.02

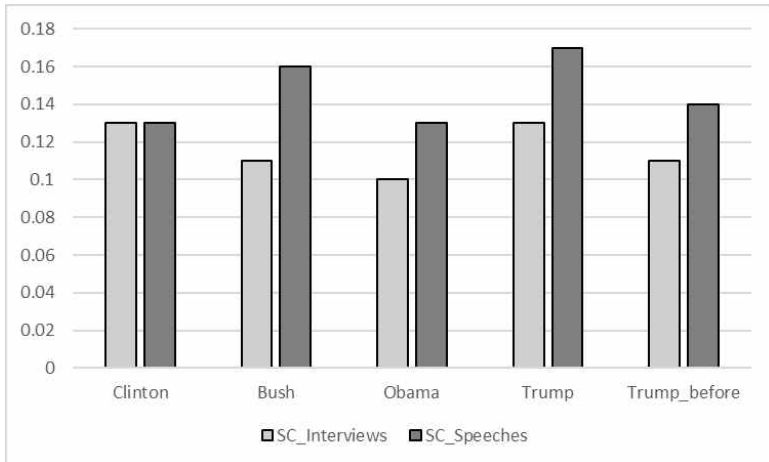


Figure 6. Mean values of SC of presidents

Table 13. P-values from the comparisons of SC between Trump and other presidents

		Clinton	Bush	Obama	Trump before presidency
Trump	Interviews	> .1	.508	.007	.334
	Speeches	.121	> .1	.066	.099

4.3.2. Given/New Information (GNI)

Results of the descriptive statistics in interviews show that Trump has the highest value ($M = 0.34$, $SD = 0.02$), followed by Clinton ($M = 0.32$, $SD = 0.01$), and then by Bush ($M = 0.30$, $SD = 0.01$). Obama has the lowest value ($M = 0.29$, $SD = 0.01$). The one-way ANOVA revealed a significant difference among these four presidents in GNI, $F(3, 60) = 51.476$, $p < .001$, $\eta_p^2 = 0.720$. As shown in Table 15, post-hoc tests (Bonferroni correction) showed that Trump had significantly higher values than the other presidents.

In speeches, results of the descriptive statistics that Trump has the highest value ($M = 0.31$, $SD = 0.02$) with Bush ($M = 0.31$, $SD = 0.02$). However, the one-way ANOVA revealed no significant difference among these four presidents in GNI, $F(3, 60) = 2.326$, $p = .084$, $\eta_p^2 = 0.104$.

As shown in Table 15, there was no significant difference in Trump's GNI in both interviews ($p = .052$) and speeches ($p = .783$) between the two time periods.

Table 14. Descriptive statistics of GNI of presidents

Presidents	Interviews		Speeches	
	Mean	Std. Deviation	Mean	Std. Deviation
Clinton	0.32	0.01	0.30	0.01
Bush	0.30	0.01	0.31	0.02
Obama	0.29	0.01	0.30	0.01
Trump	0.34	0.02	0.31	0.02
Trump before	0.35	0.01	0.32	0.02

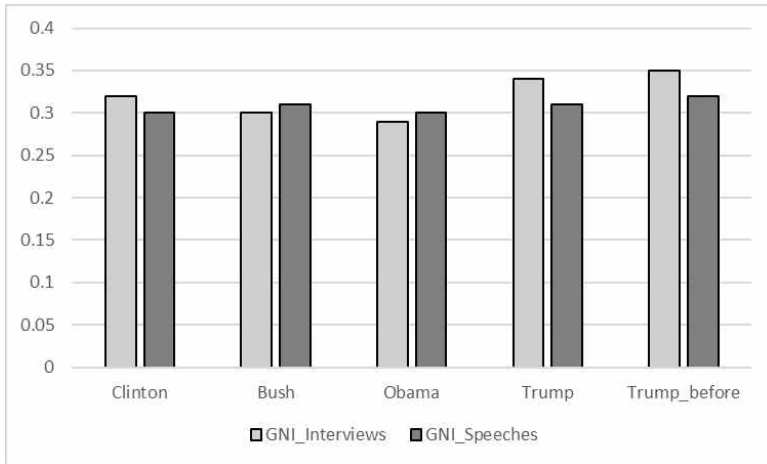


Figure 7. Mean values of GNI of presidents

Table 15. P-values from the comparisons of GNI between Trump and other presidents

		Clinton	Bush	Obama	Trump before presidency
Trump	Interviews	.001	< .001	< .001	.052
	Speeches	.529	> .1	.190	.783

The analyses of interviews showed that Trump's text cohesion measured by SC was not different from that of Clinton and Bush, and was higher than that of Obama. Moreover, his value of GNI was significantly higher than those of all three predecessors. These results indicate that even though his language is lexically and syntactically simple in interviews, it shows higher text cohesions than his predecessors. Example (4) is parts of Trump's interview and speech that showed high levels of SC and CS.

- (4) a. Well, it's going to affect everybody. It's going to, actually, we're getting rid of a lot of the big deductions for business,

and we're simplifying, we're bringing down the rate, but the rate is actually for rich people, about the same. The big beneficiaries are the middle class; I call it the working people. But the big beneficiaries are the middle class. (Trump_Interviews_October 17, 2017_Interview with David Webb of Sirius XM Patriot Radio)

- b. He trained terrorist armies, including Hezbollah, launching terrorist strikes against civilian targets. He fueled bloody civil wars all across the region. He viciously wounded and murdered thousands of U.S. troops, including the planting of roadside bombs that maim and dismember their victims. Soleimani directed the recent attacks on US personnel in Iraq that badly wounded four service members and killed one American, and he orchestrated the violent assault on the US embassy in Baghdad. In recent days, he was planning new attacks on American targets, but we stopped him. (Trump_Speeches_January 8, 2020_Statement on Iran, Washington, D.C.)

This result goes against previous criticisms regarding inconsistency of his language. Although he appears to be commanding a less detailed language, his sentences are cohesively connected, which could have played an important role in attracting the audience. As Sclafani (2017) argued, Trump uses everyday language in interviews, and makes his sentences flow naturally, which probably attracts the audience's attention in a friendly and engaging manner. Moreover, the high level of text cohesion in Trump's language use is consistent with Wang and Lui's findings (2018) that the thematic concentration of his discourse is intensive.

One may not rule out the possibility that his repeated use of words

(i.e., his low vocabulary diversity) may have distorted the figures and led to the high level of text cohesion in Trump's discourse. Nevertheless, his speeches, which show high values in STTR and LD indices, indicate that his overall use of language is far from being incohesive. Moreover, Obama's low SC and CS figures require deep qualitative research as that may indicate that Obama deals with more diverse topics in interviews and speeches than other presidents.

5. Conclusion

This study began with an attempt to verify the evaluations of Trump's political discourse through corpus analysis and make more accurate assessments through a genre-specific analysis. The main criticism of Trump's political discourse has been that its lexicon, syntax, and cohesiveness are low-level. In order to assess whether this criticism can be applied to different communication genres, this study analyzed Trump's political discourse by genre - interviews vs. formal speeches. To verify all components of the criticism - lexicon, syntax, and cohesiveness, the present work investigated Trump's political discourse along three different dimensions: lexical complexity, syntactic complexity, and text cohesion. Furthermore, previous research was limited to his pre-presidential discourse, so this study expanded the analysis to his discourse after his election and examined how it has evolved.

The results of the analyses showed that Trump utilizes two different language styles, one for interviews versus another for formal speeches. His interviews are simple both lexically and syntactically, which is consistent with existing research results and evaluations. However, the

level of complexity of Trump's speeches is on par with that of his predecessors. There was also a notable finding with regard to syntactic complexity. Trump had a similar FK grade level as Bush and Obama, with only Clinton having a significantly higher value, while Trump's C/S value was lower than that of his three predecessors. This indicates that the lexical difficulty and sentence length of Trump's speeches are comparable to those of his predecessors; only his use of clauses is plainer. In sum, criticism of Trump's low level of political discourse only applies to his interviews and not his formal speeches, while it can be said that his formal speeches show less diversity only in clauses.

Another finding of note is that while his formal speeches do not differ from those of other presidents in cohesion, his interviews show the highest cohesion in both SI and GNI. Unlike the general assessment that his use of language is distracting, the text cohesion of his formal speeches was no different from that of other presidents, and the text cohesion of his interviews was higher. We cannot rule out the possibility that these high cohesion values are the result of his repeated use of certain vocabulary. However, another feasible interpretation is that Trump excels in presenting given/new information and maintaining high coherency in his interviews, while showing low levels of lexical and syntactic difficulty.

Finally, when it comes to temporal changes in Trump's political discourse, there was no change in his interviews, but his formal speeches show that his language use has gained in complexity during his tenure. This could mean that he has made efforts to speak more presidentially after his election. It would be difficult to make an assessment of whether this is just the impact of speechwriters, as has been indicated in preceding studies, or instead reflects a strategic use of different language styles for

interviews and official speeches. Though some would like to give credit to his speechwriters, one might instead say that he is politically brilliant in mixing and matching different styles and intentions. In other words, Trump believes that the repetition of a simple message is enough to persuade citizens in interviews (Savoy, 2017b); he does not depart from this communication style, but follows the same practices at appropriate times. This is also supported by the findings of this study that the standard deviations of his speeches are higher than those of his predecessors, which is in line with Wang and Lui's finding (2018) that he is adept at stylistic changes depending on the audience and subject matter.

In sum, this study is significant in that it has redefined the characteristics of Trump's political discourses by applying genre analysis and diversifying the indices used. This study confirms Savoy's (2017b) and Wang and Lui's (2018) findings that Trump's lexical and syntactic complexities vary from genre to genre and finds new features of his political discourse: less use of clausal units in his formal speeches and high cohesiveness in his interviews. Furthermore, this study examined the temporal changes of his political discourse to find that his formal speeches have become more complex since his election. One last caveat to note is that this study verified and presented the characteristics of Trump's overall language use in numerical values. Quantitative research such as this does have its limitations, as it still does not give enough explanation about the causes of the numerical verification. Therefore, supplementing this study with a qualitative analysis would help re-examine and capture Trump's language in a more holistic manner.

References

- Ahmadian, S., Azarshahi, S. and Paulhus, D. L. (2017). "Explaining Donald Trump via communication style: Grandiosity, informality, and dynamism", in *Personality and Individual Differences*, 107, 49-53.
- Al-Saedi, H. M. A. (2017). "The function of repetition in Trump's inaugural address: A discourse analysis study", in *Journal of College of Education/Wasit*, 1(28), 709-732.
- Anthony, L. (2019). AntConc version 3.5.8 [Computer Software]. Tokyo, Japan: Waseda University. Available from <https://www.laurenceanthony.net/software>
- Aswad, N. G. (2019). "Exploring charismatic leadership: A comparative analysis of the rhetoric of Hillary Clinton and Donald Trump in the 2016 presidential election", in *Presidential Studies Quarterly*, 49(1), 56-74.
- Chen, W. (2018). "A critical discourse analysis of Donald Trump's inaugural speech from the perspective of systemic functional grammar", in *Theory and Practice in Language Studies*, 8(8), 966-972.
- Colvin, G. (2016). Explaining Donald Trump's massive branding power. *Fortune*. Retrieved July 30, 2020, from <http://fortune.com/2016/04/28/donald-trump-branding-power/>
- Dalman, A. S. (2017). "A corpus analysis of Donald Trump's political communications", in *International Journal of Current Research*, 9(11), 61802-61807.
- Graesser, A. C., McNamara, D. S., Louwse, M. M. and Cai, Z. (2004). "Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, 36(2), 193-202.
- Golshan, T. (2016). "Donald Trump's strange speaking style, as explained by linguists", in *Vox*. Retrieved July 30, 2020, from <https://www.vox.com/2016/8/18/12423688/donald-trump-speech-style-explained-by-linguists>
- Hanson, M. (2018). "Does Trump have a low verbal ability? Let's look at his language from a linguistic standpoint", in *Medium*. Retrieved July 30, 2020, from <https://medium.com/@madelainehanson/is-trump-verbally-impaired-lets-look-at-his-language-from-a-linguistic-standpoint-a906daa28aaf>
- Hendry, E. (2018). "Trump's language shifts from 'I' to 'we' in State of the Union address", in *Public Broadcasting Service*. Retrieved August 28, 2019, from

<https://www.pbs.org/newshour/politics/trumps-language-shifts-from-i-to-we-in-state-of-the-union-address>

- Kayam, O. (2018). "The readability and simplicity of Donald Trump's language", in *Political Studies Review*, 16(1), 73-88.
- Kincaid, J. P., Fishburne, R. P. Jr., Rogers, R. L. and Chissom, B. S. (1975). "Derivation of new readability formulas (Automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel", in *Research Branch Report no. 8-75*, February. Millington, TN: CNTECHTRA.
- Kvítková, A. (2018). *Textual patterns and idiolect: A corpus-assisted study of individual textual profiles of Hillary Clinton and Donald Trump*. Doctoral dissertation. Prague: Charles University in Prague.
- Lakoff, G. (2016). *Understanding Trump*. Retrieved July 30, 2020, from <https://georgelakoff.com/2016/07/23/understanding-trump-2>.
- Landauer, T., McNamara, D. S., Dennis, S. and Kintsch, W., eds. (2007). *Handbook of latent semantic analysis*. Mahwah, NJ: Erlbaum.
- Liberman, M. (2016). "The rhetorical style of spontaneous speech", in *Language Log*. Retrieved July 10, 2020, <http://languagelog.ldc.upenn.edu/nll/?p=27515>
- Lu, X. (2011). "A corpus-based evaluation of syntactic complexity measures as indices of college level ESL writers' language development", in *TESOL Quarterly*, 45(1), 36-62.
- McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge: Cambridge University Press.
- Nerdwriter1 (2015, Dec, 31). *How Donald Trump answers a question*. YouTube. Retrieved August 30, 2019, from <http://digg.com/video/donald-trump-linguistics-answer-question>
- Pullum, G. (2015). *Trump's aphasia*. *Language Log*. Retrieved August 29, 2019, from <http://languagelog.ldc.upenn.edu/nll/?p=20490>
- Savoy, J. (2017a). "Analysis of the style and the rhetoric of the 2016 US presidential primaries", in *Digital Scholarship in the Humanities*, 33(1), 143-159.
- Savoy, J. (2017b). "Trump's and Clinton's style and rhetoric during the 2016 presidential election", in *Journal of Quantitative Linguistics*, 25(2), 168-189, DOI: 10.1080/09296174.2017.1349358
- Scott, M. (1997). *Wordsmith: Software Tools for Windows*. Oxford: Oxford University Press.
- Scott, M. (2011). *WordSmith Tools version 6 [Computer Software]*. Liverpool: Lexical Analysis Software.

- Sclafani, J. (2017). *Talking Donald Trump: A sociolinguistic study of style, metadiscourse, and political identity*. London and New York: Routledge.
- Simms, K. (2018). "One year of Trump: Linguistics expert analyses US President's influence on language", in University of Liverpool. Retrieved August 1, 2020, from <https://news.liverpool.ac.uk/2018/01/19/one-year-trump-linguistics-expert-analyses-us-presidents-influence-language/>
- Swaim, B. (2015). How Donald Trump's language works for him. *The Washington Post*. Retrieved August 6, 2020, from <https://www.washingtonpost.com/news/the-fix/wp/2015/09/15/how-trump-speak-has-pushed-the-donald-into-first-place>
- Tchaparian, V. (2017). "Linguistic analysis of the two speech functions in President Donald Trump's inaugural speech", in *Armenian Folia Anglistika*, 111-118.
- Wang, Y. Q., and Liu, H. T. (2018). "Is Trump always rambling like a fourth-grade student? An analysis of stylistic features of Donald Trump's political discourse during the 2016 election", in *Discourse and Society*, 29(3), 299-323.

Woonhyung Chung

Doctoral Student, Department of English Linguistics and Language, Yonsei University.
[03722] 50, Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea
E-mail: woonka82@hanmail.net

원고 접수일 : 2020년 11월 10일

원고 수정일 : 2021년 1월 24일

게재 확정일 : 2021년 2월 10일

Abstract

**트럼프의 정치적 담화와 언어적 특성:
장르별 차이와 시간에 따른 변화를 중심으로**

정운형
(연세대학교)

미국 대통령 트럼프의 언어적 특징은 쉽고 반복된 어휘와 짧은 문장의 사용, 그리고 낮은 텍스트 응집성이라고 알려져 있다. 그러나 몇몇 연구들은 그의 언어가 항상 그런 특징을 보이지는 않는다는 것을 보여주었다. 이 연구는 트럼프의 언어 사용을 장르별(인터뷰 대 스피치)로 다양한 언어적 차원에서 분석하고 대통령 당선 전과 후로 나누어 살펴봄으로써 그동안의 그의 언어 사용 특징에 대한 평가와 연구에 대해 양적으로 검증하기 위한 시도에서 시작되었다. 분석 결과, 그의 인터뷰는 기존의 평가와 연구와 마찬가지로 어휘적, 구문적으로 단순한 것으로 나타났지만 그의 스피치의 경우에는 어휘적, 구문적으로 세 명의 전임 대통령들과 유의미한 차이를 보이지 않았다. 더 나아가, 스피치에서는 대통령 당선 이후가 당선 전보다 그의 언어가 어휘적, 구문적으로 복잡해지는 것으로 나타났다. 텍스트 응집성에 있어서도 기존의 평가와는 달리 인터뷰에서 트럼프가 세 명의 전임 대통령들보다 더 높은 응집성 수치를 보여주었다. 이러한 결과는 기존의 트럼프의 정치적 언어 사용에 대한 평가와 연구가 다양한 언어적 차원에서 장르별, 시기별로 다시 검토될 필요가 있음을 보여주는 것이다.

Keywords: Donald Trump, political discourse, corpus-based study, linguistic features, genre differences

핵심어: 도널드 트럼프, 정치적 담화, 코퍼스 기반 연구, 언어적 특성, 장르별 차이

Abstract

Linguistic Features of Donald Trump's Political Discourse: Focusing on Genre Differences and Change over Time

Woonhyung Chung
(Yonsei University)

US President Donald Trump's language has been characterized by easy and repeated words, short sentences, and incohesive content. However, some studies have suggested that his language does not always show those linguistic features. This study addressed this unresolved issue by analyzing Trump's language use by genre (interviews vs. presidential speeches) from various linguistic dimensions, before and during his presidency. Results showed that while his interviews were generally simple in terms of both lexicon and syntax, his presidential speeches showed similar degrees of lexical and syntactic complexity with his three predecessors. Furthermore, his speeches have tended to become more lexically and syntactically complicated since his election. In terms of text cohesion, his interviews showed higher cohesion than those of his three predecessors, although text cohesion of his speeches was similar with them. These results contest the existing assessment that Trump's language is linguistically simple and incohesive, suggesting that the characteristics of his political discourse should be redefined by genre in terms of language complexity and text cohesion.

핵심어: 도널드 트럼프, 정치적 담화, 코퍼스 기반 연구, 언어적 특성, 장르별 차이

Keywords: Donald Trump, political discourse, corpus-based study, linguistic features, genre differences