Szczygłowska, Tatiana

Exploring Obama's and Trump's political discourse through the lens of wordlists, keywords and clusters

Brno studies in English. 2022, vol. 48, iss. 1, pp. 93-116

ISSN 0524-6881 (print); ISSN 1805-0867 (online)

Stable URL (DOI): https://doi.org/10.5817/BSE2022-1-5
Stable URL (handle): https://hdl.handle.net/11222.digilib/digilib.76859
License: CC BY-NC-ND 4.0 International
Access Date: 17. 02. 2024
Version: 20221109

Terms of use: Digital Library of the Faculty of Arts, Masaryk University provides access to digitized documents strictly for personal use, unless otherwise specified.

MUNI ARTS Masarykova univerzita Filozofická fakulta

Digital Library of the Faculty of Arts,
Masaryk University
digilib.phil.muni.cz

Exploring Obama's and Trump's Political Discourse through the Lens of Wordlists, Keywords and Clusters

Brno Studies in English
Volume 48, No. 1, 2022

ISSN 0524-6881 | e-ISSN 1805-0867
https://doi.org/10.5817/BSE2022-1-5

Tatiana Szczygłowska

Abstract
Using corpus linguistics analysis techniques, this study compares salient lexical and phraseological features of the political discourse of Barrack Obama and Donald Trump to reveal the presidents' individual linguistic styles and preferred discursive themes. For this purpose, the WordSmith Tools software is used to extract frequency wordlists, keyword lists and clusters from two corpora of political speeches, and select the most frequent units for further identification of their patterns of use and function. The findings show that the speakers convey their intent to the listeners relying on their own idiolects. Differences are found in thematic concerns and in the use of those items that figure high in both corpora, building the image of Obama as a serious, objective and organized speaker, and that of Trump, as emotion-driven, subjective and fraternizing.

Key words
Political discourse; wordlists; keywords; clusters; Obama; Trump

1. Introduction

Politics is an issue that almost everyone has an opinion about, which comes as no surprise, since "politics is power over man", as Morgenthau (1946: 195) asserted, and power brings control, money, authority and other advantages. Politics gives rise to debates in the public sphere, hosting voices of journalists and commentators as well as those directly involved in state governance activities. It also dominates private conversation, dividing friends and families into the proponents and opponents of a given parliamentary faction or its individual members. Similarly, politics pervades scholarly research, especially in the humanities and the social sciences which look at it from a variety of perspectives, including those concentrated on political discourse and its actors. Politics and language are closely linked, because "language is vital to the process of transforming political will and power into social governance and all political actions are prepared, accompanied, and controlled by language" (Partington 2013: 1299).

Seemingly, any type of political discourse can potentially be of interest to a researcher but the fact is that "the vast bulk of studies of political discourse is about the text and talk of professional politicians" (Van Dijk 1998: 12). The motivation to explore the language of specific speakers is even stronger if the political actors are themselves considered as an anomaly in the world of politics. Capitalizing on these premises, this corpus linguistic study focuses on the political language of two American presidents, Barack Obama and Donald Trump, offering insight into prominent lexical and phraseological features found in a sample of their spoken texts to reveal their individual linguistic styles and preferred discursive themes. Both politicians have been voted America's most admired men in 2019 and each has been regarded as controversial when elected president – Obama largely for his biracial origin and Trump for his status of a billionaire TV celebrity and real estate mogul. As political actors, they represent two different parties, personalities and styles of leadership, which is likely to be reflected in the speech patterns of their idiolects.

2. Using corpus linguistics to explore political discourse

Political discourse has been the focus of study in many disciplines but its definition is still vague. Graber (1981: 196) sees it as a situation in which "political actors, in and out of government, communicate about political matters, for political purposes". Van Dijk (1998) conceptualizes political discourse in terms of its main participants, practices, functions, implications, context and events. According to Ädel (2010), the concept has three definitional scopes, determined respectively by the political genre, topic and underlying issue. Fetzer (2013) characterizes political discourse as institutional, public, mediated, dynamic, linked with the media, bound by culture and ideology. Burkhardt (1996 in Wodak 2009: 6) instead proposes the term 'political language' to refer to "all types of public, institutional and private talks on political issues, all types of texts typical of politics as well as the use of lexical and stylistic linguistic instruments characterizing talks about political contexts". What seems to complete the above picture is Ralph Waldo Emerson's famous quote: "Speech is power: speech is to persuade, to convert, to compel", as it epitomizes the essence of political discourse.

Considering the variety of aspects involved in what can be essentially seen as "the verbal interaction between political actors in a policy domain", political discourse appears as a complex phenomenon (Leifeld 2014: 1). This complexity can be partly revealed by applying the methods of corpus linguistics which facilitate insight into the linguistic practices and resources employed in political communication. Writing about corpus studies of political discourse, Ädel (2010) mentions political genres as one of the popular research subjects, as exemplified by Bevitori's (2006) study of debates in the UK House of Commons. Additionally, Ädel (2010) discusses various, sometimes overlapping, corpus analysis techniques for exploring political discourse. One of them involves looking at how something is talked about, as in Willis's (2017) analysis of how UK politicians discuss climate change.

Another focus of analysis can constitute sets of linguistic features characterizing a particular style (e.g. Pearce's (2005) study of informalisation in UK party election broadcasts) or corpus comparisons of different speakers (e.g. Milizia's (2010) and Milizia and Spinzi's (2010) comparisons of Bush's and Blair's political language). Other phenomena that have attracted researchers' attention include (un)favourable representation of concepts (e.g. Koteyko's (2007) study of English loanwords in texts by the Russian pro-Communist community), metaphors (e.g. Partington 2003) and reported speech (e.g. Garretson and Ädel 2008). A popular approach is also the analysis of keywords, as has been done by Bachmann (2011) in his study of the civil partnership debates in both Houses of the UK Parliament. An interesting direction in corpus studies of political discourse involves cross-discourse and cross-linguistic comparisons, such as, respectively, Duguid's (2009) exploration of how voices about the Iraq war in 2003 are presented in political and news discourse and Pan's (2019) study of policy speeches in the United Kingdom and Hong Kong.

Regarding the studies that have dealt with how language is used by the two politicians discussed in the present paper, some of them also rely on systematic corpus data. Trump has been the focus of, for instance, Ahmadian et al.'s (2017) analysis of traces of populism in the president's communication style, Sclafani's (2018) study of discourse-marking and interactional devices in his idiolect, Savoy's (2018) comparison of his style and rhetoric with that of Hilary Clinton's, or Homolar and Scholtz's (2019) investigation of linguistic manifestations of 'crisis talk' in what they call "Trump-speak". In turn, Obama has been the object of Boyd's (2009) research into those elements that deracialize the president's discourse or Shelly's (2009) exploration of the practices and techniques contributing to his oratorical strength. Additionally, attention has been devoted to comparisons of both presidents' use of language, as exemplified by Hunston's (2017) examination of their inaugural speeches or Wang and Liu's (2018) discussion of Obama's, Trump's and Clinton's debates and campaign speeches.

The key advantage of using corpus techniques for the study of political discourse is that they allow the analysis of many authentic texts at once, facilitating the identification of linguistic patterns and trends that might be overlooked in an individual text. Quantitative data can be collected quickly and objectively, which lays solid foundations for further qualitative analysis and reduces arbitrariness in the researcher's analytical process. According to Kutter (2017: 183), such "a bird's eye view of the use of specific words across all the texts contained in a corpus" combined with "the display of word clusters" reveal "patterned semantic, syntactic, and so on, relationships between words", helping to understand the subtle ways in which political language is used.

From among the different procedures of corpus analysis, a comprehensive overview of which can be found in Baker (2006), some have been selected for use in this study. The first consists in the generation of wordlists which, as Szudarski (2010: 23) claims, "are a powerful tool for making comparisons between corpora that represent different language uses", since they reveal the general lexical profile of the political language of the considered speakers and help envisage which of its aspects are worth further attention. Another technique is keyword analysis

that focuses on single-word units of outstanding frequency in a text or corpus by comparison to a reference corpus. Keywords are "strongly associated with the content of texts in a target discourse domain" (Egbert and Biber 2019: 78) and thus reveal the "aboutness" (Scott and Tribble 2006) and style of the analyzed political texts. Ädel (2010: 597) also notes that keyword analysis of political discourse shows the "recurrent ways of talking about concepts and ideas", indicating how the discourse users "think about the social world". Following Halliday's (1994) functional view of language, it can be said that the majority of content items reveal something about the ideational focus of the texts, personal pronouns or modality markers as keywords indicate the preferred interpersonal meanings, whereas keywords in the form of contractions can disclose the speakers' textual choices.

Another means of gaining insight into the political language of Obama and Trump involves a qualitative analysis of the concordances of selected lexical items. As Grabowski (2015: 213) explains, the exploration of such typical co-occurrence patterns of potentially interesting words helps classify them "into semantic or functional categories reflecting their various aspects (e.g. a type of information they convey, role in the organization and structure of particular discourse, evaluative charge, semantic prosody etc.)". Added to that is the analysis of recurrent word clusters, also termed 'lexical bundles', 'chunks' or '*n*-grams', that are likely to perform different discourse functions, the identification of which can contribute to a better understanding of Obama's and Trump's linguistic styles.

3. Material and method

The corpus data used in this study consist of 214,683 words from 169 speeches by Barack Obama (selected over a period from 2008 to 2016) and 174,306 words from 65 speeches by Donald Trump (from 2016 to 2019). The texts were downloaded from the official websites of the Presidents and from two collections of political speeches available on the website http://www.thegrammarlab.com: the Corpus of Presidential Speeches (Brown 2016) and the Clinton-Trump Corpus (Brown 2017). The target corpora, hereafter referred to as Obama and Trump, include speeches in the proper sense as well as weekly addresses, debates, press conferences and briefings, which were filtered to remove unnecessary headings, items like <applause> or <laughter> and fragments spoken by other people.

The diversification of spoken political language for inclusion in the corpora was motivated by the desire to reduce potential bias that may creep in the research results if only formal speeches were considered. As Milizia and Spinzi (2010: 60-61) claim, such speeches are usually "pre-prepared, written-to-be-spoken" and thus "cannot be regarded as fully representative of spoken language". To minimize the effects they can possibly have on the collected data, also "hybrid forms, such as the written text of oral performance" were taken into account, owing to their potential for displaying characteristics unique to spoken discourse (Sauer 2002: 115). The same procedure was applied to compile the Referential Corpus (RC), which comprises 1,074,187 words from 340 speeches collected from

the mentioned website by Brown and dating back to President Eisenhower. The design of the RC was guided by Berber-Sardinha's (2000: 12) claim that a reliable reference corpus should be around five times the size of the study corpus and Scott and Tribble's (2006: 58) suggestion that it "should be an appropriate sample of the language which the text we are studying (the 'node-text') is written in".

The main investigative technique was a quantitative and qualitative comparison between the corpora made with the use of *WordSmith Tools* 6.0 (Scott 2012) to capture the main lexical and phraseological aspects of the political language used by each speaker. To this end, lists of top-frequency words, keywords and clusters were generated and subsequently subjected to a more fine-grained analysis supported by concordance lines of selected items.

4. Results

In the following sections, there are reported the corpus analysis results for the aspects of Obama's and Trump's speeches considered in the study. Unless stated otherwise, statistical significance, determined by means of a contingency table $\chi 2$ statistical test, is reported at the α level of 0.05.

4.1 Wordlists

In the first stage, two wordlists were generated from the corpora and compared to determine the type, range and distribution of the high-frequency vocabulary. Considering the top 25 items in both lists, the majority were closed-set items, as is usually the case with most corpora (see Baker 2006: 53, Scott and Tribble 2006: 23). Admittedly, there were verbs such as *is, are* or *have*, but a careful scrutiny of their concordance lines revealed that they usually served as function words. Hence, generalizations were made about such cases, based on concordancing results. The only clear lexical items were found in Trump: the noun *people*, ranking 21st, and the verb *know*, ranking 25th.

Subsequently, with the help of the concordancing function of WordSmith Tools, 25 top-frequency content words were selected from the corpora. As shown in Table 1, 13 italicized words overlap, indicating tentatively the presidents' shared concerns with the country and its citizens (e.g. *people, American, country, jobs*). Still, a closer analysis of the concordances of the overlapping items revealed differences in their use by each speaker. For instance, the adverb *here* was most often embedded within the cluster *right here in* (22 examples) in Trump and *here at home* (28 examples) in Obama, where it referred to America. In Trump *right here in* was followed by *America* only two times and in one case it was followed by *the USA*. In Obama, the clusters of *here* focused explicitly on America were more numerous and also included *(right) here in America* (18 examples) and *(right) here in the (United States)* (8 examples).

Table 1. 25 top-frequency content words in Obama and Trump

Obama			Trump		
Word	**Freq.**	**Rank**	**Word**	**Freq.**	**Rank**
people	798	36	*people*	1256	21
do	698	40	*know*	1042	25
new	632	43	said	779	34
America	623	44	*do*	764	35
make	550	52	great	718	39
American	507	56	*want*	633	42
years	495	58	*country*	567	48
Americans	494	59	thank	498	59
time	489	60	*get*	492	60
work	459	63	think	481	63
country	448	64	say	414	71
world	441	65	*American*	403	74
jobs	418	71	good	375	79
get	405	74	look	372	80
know	403	75	got	366	82
economy	391	77	way	348	85
here	379	79	*jobs*	340	88
year	350	83	*years*	338	89
help	317	90	go	337	90
congress	302	93	*make*	334	91
need	300	95	see	328	93
families	298	96	never	324	95
states	298	98	*time*	311	99
want	290	100	*new*	309	100
care	284	102	*here*	308	101

Regarding other differences between the corpora, Obama willingly employs words which allude to the country, its citizens and affairs, such as *American(s), country, jobs, congress, families, economy, states*, from among which only the first three are also frequent in Trump. Trump, in turn, more readily uses evaluative words like *new, great, good*, from among which only *new* can be found in Obama's list of frequent content words. In Obama, the adjective most often refers to *job(s)* (92 examples, 14.55%) but this combination is less frequent in Trump (17 examples, 5.5%).

Trump uses more verb forms (13 vs. 9 in Obama), two of which reveal something about the speaker's preferred type of knowledge. The first is *know*, a factive verb expressing knowledge about the surrounding world and, when used in the phrase *I know*, implying someone's firm belief in their claims (Van Dijk 2003: 107). This use of *know* was significantly [$\chi2(1)=64.229$, $p<.001$, $d=0.43$] more frequent in Obama (112 examples, 27.79%) than in Trump (87 examples, 8.34%), indicating that the former speaker is likely to base his knowledge on objective "evidence or a reliable source" which "is not a matter of doubt or controversy" (Zheni 2019: 47). In Trump *know* was most often directly preceded by the pronoun *you* (652 examples, 62.57%), which was significantly [$\chi2(1)=172.294$, $p<.001$, $d=0.73$] less popular in Obama (22 examples, 5.45%). Trump's excessive exploitation of the discourse marker *you know* suggests that his individual discursive style is straightforward, direct or even brush (see Sclafani 2018: 29-30), as by using the phrase "the speaker strives towards getting the addressee to cooperate and/or to accept the propositional content of his utterance as mutual background knowledge" (Östman 1981: 17). Fox Tree and Schrock (2002: 738) also suggest that "*you know* is used when speakers are having extra trouble expressing themselves, to encourage addressees to infer the intentions". Similarly, Trump relies heavily on the verb *think*, which was not found among Obama's top-frequency content words. Considering that the verb's typical left-hand collocate was the pronoun *I* (240 examples, 49.89%), it seems that Trump prefers to sell "personal, subjective opinions as objective, reliable judgments", as the verb "expresses attitudes, hence the speaker's ideology" (Zheni 2019: 47). Also, as Bramley (2001: 260) suggests, in sentences like (1), Trump wants to distance himself from saying something about himself that he does not "want to claim knowledge of or responsibility for".

(1) *I think* my message is resonating because they have confidence on me at the border.

Since *people* is the first shared top-frequency content word, ranking 21st in Trump and 36th in Obama, it is worth closer attention, especially that it often ranks high in wordlists generated from political speeches corpora (see Milizia 2010). Another word which is of interest in a study focused on the speeches of American politicians is the proper noun America, ranking 44th in Obama and 231st in Trump. To compare the relative frequencies of both words in the corpora, the log-likelihood test (*LL*) was performed using Paul Rayson's online calculator available on the website http://ucrel.lancs.ac.uk/llwizard.html. As shown in Table 2, in Obama, as compared to Trump, *people* is underrepresented and *America* is overrepresented, with both differences found to be statistically significant at p-value <.001 (i.e. 99.9% confidence level).

Table 2. Distribution of *people* and *America* in Obama and Trump

Word	Observed freq. in Obama	Relative freq. %	Observed freq. in Trump	Relative freq. %	*LL*[1]	*ELL*[2]
people	798	0.37	1256	0.73	-220.66	0.00008
America	623	0.29	247	0.14	+99.07	0.00004

In each corpus, the words did not occur at all in five texts. The dispersion plot of *people* indicated an over 10-hit-per-text concentration of the word in 33.84% of Trump's speeches (22 texts), ranging from 12 to even 117 occurrences in a single text, which is significantly more [$\chi 2(1)=5.771$, p=.016, d=0.35] in comparison with 14.79% of Obama's speeches (25 texts), ranging from 11 to only 48 occurrences in a single text. The difference in the dispersion of *America* proved insignificant, as an over 10-hit-per-text concentration of the word was found in 6.15% of Trump's speeches, ranging from 14 to 18 occurrences in a single text, and in 8.87% of Obama's speeches, ranging from 11 to 30 occurrences in a single text.

A comparison of the 3-word clusters of *people* revealed that in Obama, the most frequent cluster is *the American people* (Freq. 107), whereas *lot of people* ranks 1st (Freq. 47) in Trump. Obama appears as more cosmopolitan, as he also speaks of *the Iraqi/Palestinian/Jewish/Afghan/Cuban/young people*, none of which occurs in Trump's 3-word clusters. Trump refers only to *people of Florida/Ohio*, which indicates a local focus of his main concerns that may be due to the fact that some of his speeches come from the presidential campaign. Yet, both presidents make a relatively frequent use of the phrase *our people* (47 tokens in Obama, 36 tokens in Trump), by which they refer to American citizens. As for the 3-word clusters of *America*, *States of America* ranks 1st in Obama (Freq. 48), whereas *make America great* (Freq. 37) is the most popular in Trump but absent in Obama. Trump's tendency to talk about greatness when referring to America and its citizens is also reflected in the frequency of the phrase *great people* (42 tokens), which was used to talk about a different nation in only one case and which was not identified in Obama at all.

Obama's frequent use of the noun *America* may be due to the exceptional discursive position that the country occupies in his speeches, testifying to the president's rhetorical identification with the place where he was born, educated and elected president. In turn, the popularity of noun *people* in Trump's speeches, also reported by Homolar and Scholtz (2019), may be due to the fact that the word is so general that it has "become a political category" – especially if not preceded by a nationality adjective – a passive mass necessary "so that the power of the capitalist oligarchy can be considered democratically legitimate" (Badiou 2016: 25, 29). Considering Trump's tendency to employ many small and non-substansive words (see Sclafani 2016), the prevalence of the word *people* is his political discourse is not surprising.

4.2 Keywords

In the next stage, two keyword lists were generated by individually comparing Obama's and Trump's wordlists with the Reference Corpus wordlist, with the minimum frequency cut-off point set at 25 and the probability value set at 0.000001, which was aimed at "controlling the quantity of keywords derived, and thus the number of keywords a researcher must interpret" (Culpeper 2009: 36). Overall, 376 keywords in Obama (326 positive) and 500 in Trump (454 positive) were identified. 50 top keyness terms were selected for further analysis and functionally classified, following Goźdź-Roszkowski's (2011) and Grabowski's (2015) approach to keyword lists.[3] The categories in Table 3 were developed intuitively after examining the concordances for each keyword to identify their dominant meanings. Contractions, such as *it's* or *we're* were treated as one orthographic word whose function was determined by the initial element (see Allwood et al. 2010: 11).

Table 3. Functional classification of top 50 keywords in Obama and Trump

Functional category	**Obama**	**Trump**
Keywords denoting the nation	*everybody*, *folks*	*folks*
Keywords related to citizen/social concerns	*class*, *families*, *kids*, *care*, *middle*	*people*, *guy*
Central pronouns as keywords	*it's*, *our*, *she*, *we*, *her*, *their*	*you*, *it's*, *they*, *we're*, *she*, *they're*, *it*, *I'm*, *you're*, *she's*, *he's*, *her*, *he*
Personal names as keywords		*Hillary*, *Trump*, *Clinton*, *Obama*
Inserts as keywords	*hi*, *thanks*	*ok*, *right*, *thank*
Intensifiers/Evaluative keywords	*more*	*lot*, *great*, *very*, *really*, *love*, *incredible*, *big*, *bad*
Miscellaneous keywords	*get*, *just*, *rules*, *keep*, *sure*, *make*, *week*, *innovation*, *weekend*[4]	*know*, *said*, *got*, *get*, *look*, *mean*, *doing*
Functors as keywords	*'ve*, *'re*, *'ll*, *that's*, *that*, *why*, *who*, *and*, *because*, *like*, *to*, *what*, *don't*, *so*, *how*	*don't*, *so*, *that's*, *like*, *because*, *what's*, *what*, *was*
Keywords related to employment concerns	*jobs*, *workers*, *wages*, *manufacturing*, *working*	*jobs*

Functional category	Obama	Trump
Keywords related to economic concerns	*businesses*, *economy*, *insurance*, *companies*	
Keywords related to political concerns	*Republicans*	
Keywords signaling intentions and desires		*going*, *want*, *gonna*

As shown in Table 3, Obama's speeches are marked by the presence of keywords that are indicative of his involvement in economic and political concerns. The matters of employment and citizen/social concerns also feature prominently in the corpus, building the image of Obama as a politician who is preoccupied with weighty issues. The same cannot be said about Trump, whose top-frequency keywords do not reveal growing interest in any serious matters except for his fellow politicians, which may suggest the speaker's "general oratorical lack of [...] substance" that according to Sclafani (2018: 1) has been often imputed to Trump. This seems to be further corroborated by the relatively low standardized type/token ratio of 34.72 compared to 43.11 for the Obama corpus. As Baker (2006: 52) claims, "a low type/token ratio is likely to indicate that a relatively narrow range of subjects are being discussed, which can sometimes [...] suggest that the language being used is relatively simplistic". Added to that is Trump's preference for uttering shorter sentences whose average length is 11.17 words in comparison with Obama's average of 18.6, which correlates with the results reported by Savoy (2018: 175), according to whom "the presence of long sentences indicates a substantiated reasoning or specifies the presence of a more detailed explanation". Regarding Trump's use of names such as *Hillary*, *Clinton* and *Obama*, the dispersion plot shows their greatest concentration in the speeches delivered during the presidential campaign when frequent reference to other politicians can be a persuasive means of defining oneself in terms of what one is or is not (see Van Hout and Van Praet 2011: 102).

Trump's speeches are also characterized by the frequent use of intensifying and evaluative keywords. As for intensifiers such as *very, really* or *so*, Sojda (2019: 61) argues that speakers employ them to make their "statements expressive, stronger in reception, sometimes blunter or marked, e.g. emotionally", which correlates with the finding of several studies (see e.g. Wang and Liu's 2018, Hololar and Scholz 2019) that Trump's idiolect is marked by an appeal to people's emotions. Intensifiers are often used with "adjectives expressing opinion or appreciation", which was also observed in Trump (e.g. *a very big trade deal* or *really bad people*), reflecting a subjective point of view and attitude (Bordet 2017).

This conclusion seems even more valid when the distributional context of *so* is taken into consideration. Based on the word's predominant meanings identified in the corpora, in both Obama (795 tokens) and Trump (1331 tokens) it was

classified as a functor, serving mainly as a conjunction (e.g. *so that; so tonight, I propose*). Still, it has to be noted that in Trump 40.12% (534 tokens) of the word's uses were in the role of an intensifying adverb, as in (2), which is by around half more in comparison with 25.28% (201 tokens) in Obama. Trump's special liking for this use of *so* further confirms him as a speaker who prefers "a trendy expressive style" (Tagliamonte 2008: 371) and "elements that carry intense semantic contents" (Bordet 2017). Sclafani (2016) also suggests that Trump "often introduces topics abruptly with non-substantive words", such as *so* or the already mentioned *you know*. These "Trumpisms" are very characteristic of his style and help him to organize talk as well as create the impression he is speaking for himself and intimately to individual people, which his supporters consider to be a sign of authenticity (Simms 2018).

(2) I think it's *so* sad and *so* disgusting.

What stands out in Trump as well is the strong use of keywords signaling intentions and desires, such as *going, gonna* and *want*. Lorenz (2013: 134), quoting several studies, argues that semi-modals such as *be going to* or *want to* "are currently on the rise in many contexts and across varieties, partly displacing the central modals (will, must, can, etc.)", which has led to the occurrence of contractions like *gonna* or *wanna*. The present and past forms of *want* (683 tokens), especially if preceded by *I* (248 tokens) or the inclusive pronoun *we* (127 tokens), express "the speaker's private state of desire" (Polge 2015: 32), reflecting their attitude or point of view, whereas *be going to/gonna* (1460 tokens: 137 preceded by *I*, 661 preceded by *we*) serve to mark personal volition and, less frequently, also prediction (Biber et al. 1999: 496). The three semi-modals are considered informal and, given their unusually high incidence in Trump, suggest that the speaker is "making a deliberate slide down the formality/informality cline, so as to reach a more colloquial, peer-to-peer level of communication", probably to make his 'voice' "more persuasive and more easily shared with the reader/audience" (Belladelli 2009: 310). Yet, as noted by Kerner (2018), such use of language makes that Trump "does not sound presidential".

Differences between the presidents' speeches can also be noticed in the categories of inserts and functors as keywords. Regarding the use of inserts, which "carry emotional and interactional meanings" and "make an important contribution to the interactive character of speech", it can be seen that polite formulas (*thank, thanks*) are the only class of inserts that the speakers share (Biber et al. 2002: 16, 449). Additionally, Obama uses the greeting *hi*, whereas Trump prefers the response getter *right* and the discourse operator *OK*, which despite its diverse interactional functions, was usually used as a confirmation marker (3) or as an attention-getter (4).

(3) We're not gonna have much of a country left. Believe me, *OK*?

(4) It's really disgusting, *OK*? It's really disgusting.

Regarding the use of functors, which "glue texts together" (Scott and Tribble 2006: 23) and may indicate their "characteristic style" (Groom 2010: 62), Obama (15 items) uses half as many of them as Trump (8 items). This may suggest the speaker's "general oratorical lack of coherence" for which Trump has been criticized (Sclafani 2018: 1). Yet, following Groom's (2010: 63) claim, if function words are to reveal something important about the meanings expressed in the investigated texts, they should be looked at in their typical contexts of occurrence. Here, the focus will be on *don't*, as it is the function keyword that is the highest ranking one in Trump (752 tokens) but significantly less popular [$\chi 2(1)=515.856$, $p<.001$, $d=0.07$] in Obama (162 tokens). In both corpora it was most often directly preceded by personal pronouns (59.87% of the concordance sample in Obama and 80.18% in Trump), from among which *I* was the most frequent in Trump (238 tokens) and *we* in Obama (41 tokens).

The phrase that is worth considering is *I don't*, as it "conveys conviction and represents the firm stance of the speaker" and is thus quite persuasive (Patrick and Hagtvedt 2012: 393). An interesting finding is that in contrast to Obama, whose use of *I don't* does not reveal any marked preferences of the speaker in terms of the choice of words that follow the phrase, Trump has a special liking for such sequences as *I don't know* (73 tokens) and *I don't think* (42 tokens). Regarding *I don't know*, Pichler (2013: 69) claims it is "the most frequent negative collocation in spoken corpora". Essentially, the phrase serves as a declaration of insufficient knowledge, which is often pragmatically motivated by "a concern to save the face of self and other" that is achieved by averting potential contradictions from addressees (Tsui 1991: 607). For instance, Trump uses it to signal uncertainty about the accuracy of some proposition, as in (5), but also to minimize impolite beliefs, as in (6).

(5) And we also raised, from other people, about – *I don't know*, the total was about $51 million, OK.

(6) You know, *I don't know* why we're going to lose.

Additionally, Pichler (2013: 148) argues that *I don't think* "serves to signal various degrees of confidence speakers attach to their propositions" as well as "to signal politeness and promote good social rapport", which is possible owing to the phrase's subjective epistemicity, whereby the speaker draws conclusions based on their "own knowledge of the state of the world at the time of speaking" (Watts 1984: 131), as illustrated by example (7).

(7) *I don't think* it matters from my standpoint. I really don't.

Other frequent sequences with *I don't* in Trump include *I don't want* (34 tokens), *I don't care* (17 tokens), *I don't get it* (11 tokens) and *I don't even* followed by a verb (11 tokens), the last three of which were not identified in Obama at all.

The functional category of central pronouns, as proposed by Quirk et al. (1985: 346), encompasses personal, possessive and reflexive pronouns. Despite being

formally categorized as function words, pronouns are considered separately in the study, since they carry important information about the speaker, the listeners and other entities referred to in the analyzed speeches, giving an idea of whom the speaker identifies with. The pronominal choices made by the two speakers differ markedly, as Obama shows a statistically significant preference [$\chi 2(1)=69.856$, $p<.001$, $d=0.14$] only for the pronoun *we* (Obama: 4297 tokens, Trump 3888 tokens) and its derivatives (Obama: 3995, Trump: 1933), which corroborates earlier findings on Obama's usage of pronouns (see e.g. Boyd 2009, Smith 2009) and reveals him as a speaker who wants to send the message that he is "on the same team, in the same boat, facing the same fate" as his listeners (Shelly 2009: 78). Obama employs *we* to construct institutional identity with the party he represents (see Wilson 1990: 63), as in (8); to refer to all Americans, invoking a sense of collectivity and solidarity (see Boyd 2009: 86), as in (9); to indicate interpersonal involvement with the audience, strengthening his own trustworthiness in their eyes (see Karapetjana 2011: 41), as in (10). Yet, in sentences similar to (11), Obama excludes the interlocutors but includes other figures, possibly to emphasize his authority over political actions and signal willingness to share responsibility for these undertakings (see Karapetjana 2011: 41).

(8) You see, *we* Democrats have a very different measure of what constitutes progress in this country.

(9) When it comes to education, we are not a collection of states competing against one another; *we* are a nation competing against the world.

(10) [...] a recognition that *we* are all in this together, and when fortune turns against one of us, others are there to lend a helping hand, [...]

(11) And I invited the Prime Minister to come to the White House in December, as *we* plan for all the important work that *we* have to do together.

Trump's pronominal choices show him as a speaker who emphasizes his individual perspective when speaking (4482 tokens of *I* and 988 of its variants vs., respectively, 3201 and 689 in Obama), which correlates with Ahmadian et al.'s (2017) finding that first-person pronouns dominate in Trump's speeches, as illustrated by (12). He also willingly divides people in groups (see Bramley 2001: 259-266), as in (13), and refers to third parties (5212 tokens of *he, she, it* and 744 of their variants vs., respectively, 2597 and 708 in Obama; 2789 tokens of *they* and 983 of its variants vs., respectively, 1144 and 1562 in Obama), which as Lyons (1977: 638) observes, "does not correlate with any positive participant role", as illustrated by (14). Trump's preference for exclusive pronouns is manifested not only in the frequent use of *I*, which, as Halamari (2008: 260) argues, "is more transparently audience-exclusive than *we*", but also *you* (3665 tokens of *you* and 380 of its variants vs., respectively, 1276 and 393 in Obama), which apart from being used generically to implicate everyone (15), may also be used to address the nation (16) or other individuals (17), without indicating that one is speaking on their behalf.

(12) Believe *me*, folks. But *I* had a feeling. *I* said *I* think [...]

(13) Now, just so *you* understand, when *we* send things to China *they* charge us a tax.

(14) *They* have taken advantage. *She* doesn't care at all about the hurting people of this country or the suffering *she* has caused *them*, and *she* meaning *she* and *her* party officials.
(15) *You* can almost say, you know, politics aside, whether *you* are a Democrat, whether *you* are a liberal, a Republican, a conservative; what are they doing?

(16) And all of *you, as Americans*, you find a home for the ones that you love the most.

(17) Are *you* allowed to set up a super-PAC, *Mike*, if *you* are the president, to fight somebody?

Other assertions that can be made about the investigated political speeches stem from an analysis of negative keywords, that is, "those items which occur significantly less often in the target corpus than in the reference one" (Evison 2010: 128). Obama's most significantly underrepresented words are *which, Mr, peace, freedom, very*, whereas the words that are rare in Trump include *itself, ourselves, future, resources* and *basic*. The unusually low frequency of *very* in Obama adds to his image of a speaker whose oratorical strength does not lie in the use of intensifiers that carry no semantic contents (see Lorenz 2002: 146). What may come as surprising is Obama's infrequent use of words like *peace* or *freedom*, yet they at least appear among his keywords, which cannot be said about Trump. The word that surprises most among Trump's negative keywords is *future*, to which the speaker refers very infrequently compared with the reference corpus.

4.3 Clusters

The final stage of the study focuses on 4-word clusters which occur at least 5 times in the corpora, totalling 732 in Obama and 726 in Trump. Goźdź-Roszkowski (2011: 110) argues that such "sequences appear to have a more readily recognizable range of structures and functions than" 3-word ones. To limit the scope of the study, 50 top-frequency clusters were subjected to a closer functional analysis, modelled on the one proposed by Biber et al. (2004), Goźdź-Roszkowski (2011) and Grabowski (2015). Similarly as in the case of keywords, the clusters were assigned to the functional categories after a close reading of the relevant concordances. Therefore, *make America great again* in Trump, for instance, was classified as a stance: intention/prediction cluster, since it was typically preceded by *going* or *will*. Still, there were clusters, such as Obama's *in my State of*, which proved difficult to categorize. The initial preposition *in* seems to indicate it is a referential: location cluster, yet the State of the Union Address is not a place but a document, so the cluster was classified as a referential: document-related one.

The data show there is little overlap among the most frequent clusters in both corpora. As summarized in Tables 4-6, Obama's speeches are dominated by referential (35) and discourse-organizing (9) clusters, which express, respectively, ideational and textual meanings, and stance clusters (5), which convey interpersonal meanings, are marginal (see Dontcheva-Navratilova 2012: 40–41). This suggests that Obama concentrates on ideas, concepts, entities, and their important attributes, which he introduces, clarifies and elaborates on in his speeches (Biber 2006: 142–145). These findings explain why Obama is often characterized as a speaker who, as Shelly (2009: xii) notes, "successfully drove his points home, fused the best of rhetoric and substance, focused on a powerful message, and delivered it with great effectiveness". Trump, in turn, whose speeches have a large number of stance clusters (35), only a few referential clusters (13) and one discourse organizer, is concerned with expressing own attitudes and assessments of information conveyed in the speeches (Biber 2006: 139–142). What both corpora have in common is that each has only one cluster that serves special conversational functions (Biber et al. 2004: 388), namely, the politeness expression *thank you very much*, ranking 2nd in Trump and 39th in Obama. It should be noted, however, that the clusters which in Obama were categorized as topic closure additionally play the role of politeness routines. Indeed, the sequence *thanks everybody and have a good/great/wonderful weekend* was used as a form of a polite farewell, but because it occurred exclusively at the end of selected speeches, its main function was identified as discourse-organizing one.

Overall, the functional classification of clusters reflects some of the trends observed in Obama's and Trump's dominant categories of keywords as well as adds details to the description of frequent and typical patterns of language use in the analyzed political speeches. As Table 4 shows, Obama's interest in such matters as economy, citizens and employment is further accentuated by various recurrent multi-word clusters, some of which reveal new aspects of the speaker's main concerns (e.g. *the men and women, and women in uniform*) rather than repeat the ideas conveyed by the keywords. This, however, does not apply to Trump, whose top-frequency clusters, or rather their lack in some of the functional categories and accumulation in the other ones, reveal him as a speaker who is less sophisticated in the choice of topics to speak about (see his referential clusters in Table 4) and less articulate in developing his argumentation in a fluent and coherent manner (see his discourse-organizing clusters in Table 5) than he is in expressing evaluations, attitudes and intentions (see his stance clusters in Table 6). A possible explanations for this, as stems from the research invoked by Begley (2017), is the gradual deterioration of Trump's cognitive functions due to aging as well as stress, anger, frustration or fatigue. Yet, as McWhorter (2018) argues, Trump's linguistic incoherence may be simply due to his preference for using casual speech, which is rapid, spontaneous as well as "fundamentally subjective rather than objective, [...] decorated with what linguists call *pragmatic* words and constructions, expressing attitude rather than content".

Table 4. Referential clusters in Obama and Trump

Category	Obama	Trump
State-related	*United States of America, the United States of, of the United States, the United States and*	*in the history of, the history of our, history of our country, of the United States*
Temporal	*now is the time, for the first time, at a time when, is the time to, by the end of, the first time in, in the #[5] century, over the past #, and a half years, at the same time*	*for a long time*
Attributive	*more than # million*	*a lot of money, a lot of people, have a lot of, and a lot of*
Location	*in the United States, the Middle East and, in the Middle East, all across the country, East and North Africa, Middle East and North*	*all over the world, in the United States*
Identification/ focus	*one of the most*	*one of the most*
Economic focus	*our businesses have created*	
Citizen--related	*men and women in, and women in uniform, of the American people, the men and women, men and women who*	
Document--related	*the Affordable Care Act, State of the Union, of the Union Address, my State of the, in my State of*	
Employ-ment-related	*# million new jobs, raise the minimum wage*	
Private focus		*a friend of mine*

Table 5. Discourse-organizing clusters in Obama and Trump

Category	Obama	Trump
Topic intro-duction/focus	*when it comes to, in the face of*	*and by the way*
Clarification/ topic elabora-tion	*and that is why, in a way that*	

Category	**Obama**	**Trump**
Closing a topic	*have a great weekend, and have a great, thanks and have a, everybody and have, thanks everybody and have a*	

Table 6. Stance clusters in Obama and Trump

Category	Obama	Trump
Desire	*to make sure that, I want to thank, and I want to*	*I don't want to, and I want to, I want to thank*
Obligation/ directive	*bless the United States, God bless the United*	*take a look at, and God bless America, you and God bless*
Intention/ prediction		*we are going to, and we're going to, we're going to have, to take care of, it's going to be, we're going to make, we're going to win, make America great again, we're going to take, we're not going to, going to take care, going to have a, we're going to d, take care of our, going to be a, going to make America, we're going to be, we're going to build, not going to happen, so we're going to, we're going to bring, we're going to get, is going to be, going to pay for, are going to make, to make America great, going to be very*
Epistemic (personal)		*and you know what, but you know what*

Regarding certain detailed findings, an interesting phrase identified as key in both corpora is *one of the most*, ranking 40th in Trump and 45th in Obama. There is a difference in the way each speaker uses this referential cluster, as Obama employs it solely to focus on positive things, such as *beautiful places* or *significant contributors*, whereas Trump uses the cluster also to emphasize something negative, such as *depressing stats*, *one-sided documents* or *dishonest papers*. Two other clusters in Trump that are worth mentioning are *and by the way* and different variants of the phrase *a lot of*. The first one is a discourse organizer which, as Sclafani (2018: 32–33) notes, is a characteristic feature of Trump's idiolect, who employs it to shift the topic of discussion to an issue he himself would like to discuss, as in (18), where he is trying to slip in a private comment on another politician. The second

one, according to Kerner (2018), is "a vague phrase that is coherent with Trump's predilection for speaking in generalities and at the same time it shows the critical importance he gives to money". *Money* was found to be the most frequent right-hand collocate of *a lot of* (300 tokens) with 50 instances (16.66%) of such use in the corpus, followed by *people* with 44 instances (14.66%).

(18) But I have great support from Israel. [...] But I have great relationships as you know, to the people in Israel. *And by the way*, Obama in my opinion is the single worst thing politically speaking that's ever happened to Israel.

The clusters that are worth exploring in Obama are those which refer to citizens, and specifically the three that incorporate the nominal binomial *men and women* (see Table 4), attested 91 times in the corpus and found to be "the most frequent binomial in both the BBC and British Books sub-corpora of the Bank of English" (Bastow 2010: 146, quoting Hatzidaki 1999) as well as the most frequent binomial in the BNC (Mollin 2014: 46). Actually, the only word with which the noun *men* coordinates in Obama is *women*. Concordances of this binomial reveal the repetitive presence of *our* in 1st and 2nd left as well as 2nd right position, and the frequent presence of *in* in 1st right as well as *uniform* in 2nd right position. These findings corroborate those of Bastow (2010: 146), who observed exactly the same trends in a corpus of geopolitical speeches delivered by senior defense officials between the years 1995–2001. This suggests that the recurrent multi-word sequence *(our) men and women in uniform* constitutes an important domain-specific topic in Obama's speeches. The men and women who are evoked by means of the discussed binomial are also frequently preceded by such evaluative adjectives as *brave* and *young*, as well as followed by the relative pronoun *who* and a verb form that refers to the actions undertaken by them, for instance, *acted so boldly, have served, fought, helped, serve our country*.

Interestingly, *men and women* was attested in Trump only 27 times, where it was most often preceded by *incredible* (6 tokens) and *forgotten* (5 tokens), as well as followed by *who* (7 tokens), but rarely collocated with *our* (1 token) and *in uniform* (1 token). Other frequent binomials in Trump, namely, *law and order* (14 tokens) and *millions and millions* (7 tokens), were not found in Obama at all, who instead willingly employed the binomial *Democrats and Republicans* (29 tokens) as well as *Israelis and Palestinians* (13 tokens) that were not attested in Trump. These findings further illustrate the differing concerns of each speaker, testifying especially to the already mentioned special importance that Trump attaches to money as well as to Obama's preoccupation with such matters as politics and other nations.

5. Conclusion

The aim of the study was to compare selected political speeches of Obama and Trump in terms of the patterns of use of key vocabulary and phraseology to reveal the discursive themes and linguistic strategies prioritized by each speaker.

The comparison of wordlists showed the strong presence of function words among high-frequency lexis in both corpora. It was also revealed that the overlapping content words, due to the differences in their use by each speaker, cannot be unequivocally interpreted as indicating shared thematic interests. In Obama, the top-frequency words indicate an increased focus on America and domestic issues as well as a cosmopolitan mind-set, which is not the case in Trump, where preference is given to evaluative words and verb forms. The concordances of selected mental verbs showed that Obama relies on objective evidence which induces him to hold trust in own claims, whereas Trump attempts to persuade the audience into accepting his personal opinions.

The comparison of keywords demonstrated that Obama's speeches are dominated by lexis indicating commitment to such weighty matters as economy and politics, which are alluded to in a controlled and orderly manner. Many keywords also show the speaker's inclination for phatic expressions that help develop bonds of sociality with the listeners (e.g. *hi, thank/s, weekend*), solidarity with whom is emphasized by the frequent use of the pronoun *we*. By comparison, Trump's speeches were found to have more intensifying and evaluative keywords as well as items signaling intentions and desires. Keywords suggesting the speaker's interest in serious matters were rare, unlike markers of subjectivity, self-reference, emotions and directness.

The comparison of clusters revealed further differences, showing Obama's preference for referential and discourse-organizing clusters and Trump's, for stance clusters. The image of each politician that emerges from the data is basically that Obama appears as the orator who relies on eloquent and orderly speech – one who appeals to his audience, delivering messages that are impartial and simultaneously filled with substance. In turn, Trump seeks the strength of his persuasion in those arguments that are emotional and quite personal in their nature, which harmonizes well with his fraternizing and camaraderie attitude to listeners. Overall, the speakers' liking for specific linguistic exponents of meaning construction as well as their unique use in the speeches, as reflected in the concordances, testify to the differing concerns of the politicians both in terms of the discursive themes discussed by them and the favoured linguistic conventions.

Inevitably, this study has its limitations. First, the investigated speeches were not delivered during the same period of time, which may partly explain differences in their thematic focus. Second, some of the texts could have been prepared by professional speechwriters whose linguistic style may not necessarily be inherent in the speakers' own idiolects. Third, the genres of political discourse included in the corpora are not equally represented in each of them, which may cause an unequal distribution of certain lexical features. Fourth, the proposed functional classifications of keywords and clusters may seem somewhat subjective as they were developed intuitively. Finally, although the paper revolves around the words and phrases that are considered to be the most distinctive ones, the limitations relating to its size restricted the analyzed sample of top-frequency items. Yet, this shortcoming can be overcome in the future by investigating other salient lexical and phraseological features of Obama's and Trump's political language and

devoting more attention to a comprehensive interpretation of additional concordances showing how the speakers use these key items.

Notes

1 *LL* = log-likelihood value.
2 *ELL* = effect size for log likelihood.
3 In their studies of keywords – in legal discourse by Goźdź-Roszkowski (2011) and in pharmaceutical discourse by Grabowski (2015) – both scholars classify these words into a variety of functional categories aimed at reflecting their numerous fine-grained aspects, such as a type of information they convey, evaluative charge or role in the organization and structure of particular discourse.
4 76.19% of the word's occurrences were within the phrase *have a good/great/wonderful weekend*, which the speaker used as a form of polite farewell at the end of his speeches.
5 The hash (#) symbol replaces numbers used in the analyzed texts.

References

Ädel, Annelie (2010) How to use corpus linguistics in the study of political discourse. In O'Keeffe, Anne and Michael McCarthy (eds.) *The Routledge Handbook of Corpus Linguistics*. London: Routledge, 591–604.

Ahmadian, Sara, Sara Azarshahi and Delroy L. Paulhus (2017) Explaining Donald Trump via communication style: Grandiosity, informality, and dynamism. *Personality and Individual Differences* 107, 49–53.

Allwood, Jens, A. P. Hendrikse and Elizabeth Ahlsén (2010) Words and alternative basic units for linguistic analysis. In: Henrichsen, Peter Juel (ed.) *Linguistic Theory and Raw Sound*. Copenhagen: Samfundslitteratur, 9–26.

Bachmann, Ingo (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.

Badiou, Alain (2016) Twenty-four notes on the uses of the word "people". In: Badiou, Alain, Pierre Bourdieu, Judith Butler, Georges Didi-Huberman, Sadri Khiari and Jacques Rancière (eds.) *What is a People?* New York: Columbia University Press, 21–31.

Baker, Paul (2006) *Using Corpora in Discourse Analysis*. London: Continuum.

Bastow, Tony (2010) *Friends and allies:* The rhetoric of binomial phrases in a corpus of U.S. defense speeches. In: Okulska, Urszula and Piotr Cap (eds.) *Perspectives in Politics and Discourse.* Amsterdam: John Benjamins, 143–154.

Begley, Sharon (2017) Trump wasn't always so linguistically challenged: What could explain the change? *STAT* May 23rd. Accessed on 22 October 2019 from https://www.statnews.com/2017/05/23/donald-trump-speaking-style-interviews/.

Belladelli, Anna (2009) The interpersonal function of going to in written American English. In: Renouf, Antoinette and Andrew Kehoe (eds.) *Corpus Linguistics: Refinements and Reassessments*Amsterdam: Rodopi, 309–325.

Berber-Sardinha, Tony (2000) *Comparing corpora with WordSmith tools: How large must the reference corpus be?* In Kilgarriff, Adam and Tony Berber-Sardinha (eds.) Proceedings of the Workshop on Comparing Corpora – Volume 9. Stroudsburg: Association for Computational Linguistics, 7–13.

Bevitori, Cinzia (2006) Speech representation in parliamentary discourse. Rhetorical strategies in a heteroglossic perspective: A corpus-based study. In: Flowerdew, John and Maurizio Gotti (eds.) *Studies in Specialized Discourse.* Bern: Peter Lang, 155–180.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan (1999) *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education Limited.

Biber, Douglas, Susan Conrad and Geoffrey Leech (2002) *Longman Student Grammar of Spoken and Written English*. Harlow: Pearson Education Limited.

Biber, Douglas, Susan Conrad and Viviana Cortes (2004) *If you look at...*: Lexical bundles in university teaching and textbooks. *Applied Linguistics* 25 (3), 371–405.

Biber, Douglas (2006) *University Language: A Corpus-Based Study of Spoken and Written Registers*. Amsterdam: John Benjamins.

Bordet, Lucile (2017) From vogue words to lexicalized intensifying words: the renewal and recycling of intensifiers in English. A case-study of *very*, *really*, *so* and *totally*. *Lexis* 10. Accessed on 26 September from https://journals.openedition.org/lexis/1125.

Boyd, Michael S. (2009) De-constructing race and identity in US Presidential discourse: Barack Obama's speech on race. *Atlantis. Journal of the Spanish Association of Anglo-American Studies* 31 (2), 75–94.

Bramley, Nicolette Ruth (2001) *Pronouns of Politics: The Use of Pronouns in the Construction of 'Self' and 'Other' in Political Interviews*. Australia: Australian National University. Doctoral dissertation. Accessed on 21 October 2019 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.125.4780andrep=rep1andtype=pdf.

Brown, David W. (2016) *Corpus of Presidential Speeches*. Accessed on 2 September from http://www.thegrammarlab.com.

Brown, David W. (2017) *Clinton-Trump Corpus*. Accessed on 2 September from http://www.thegrammarlab.com.

Culpeper, Jonathan (2009) Keyness. Words, parts-of-speech and semantic categories in the character-talk of Shakespeare's *Romeo and Juliet*. *International Journal of Corpus Linguistics* 14(1), 29–59.

Dontcheva-Navratilova, Olga (2012) Lexical bundles in academic texts by non-native speakers. *Brno Studies in English* 38(2), 37–58.

Duguid, Alison (2009) Insistent voices: Government messages. In Morley, John and Paul Bayley (eds.) *Corpus-Assisted Discourse Studies on the Iraq Conflict: Wording the War.* London: Routledge, 234–260.

Egbert, Jesse and Douglas Biber (2019) Incorporating text dispersion into keyword analyses. *Corpora* 14(1), 77–104.

Evison, Jane (2010) What are the basics of analysing a corpus? In: O'Keeffe, Anne and Michael McCarthy (eds.) *The Routledge Handbook of Corpus Linguistics.* London: Routledge, 122–135.

Fetzer, Anita (2013) The multilayered and multifaceted nature of political discourse. In Anita Fetzer (ed.) *The Pragmatics of Political Discourse.* Amsterdam: John Benjamins, 1–18.

Fox Tree, Jean E. and Josef C. Schrock (2002) Basic meanings of *you know* and *I mean*. *Journal of Pragmatics* 34, 727–747.

Garretson, Gregory and Annelie Ädel (2008) Who's speaking?: Evidentiality in US newspapers during the 2004 Presidential campaign. In: Ädel, Annelie and Randi Reppen (eds.) *Corpora and Discourse: The Challenges of Different Settings.* Amsterdam: John Benjamins, 157–187.

Goźdź-Roszkowski, Stanisław (2011) *Patterns of Linguistic Variation in American Legal English: A Corpus-Based Study*. Frankfurt am Main: Peter Lang.

Graber, Doris (1981) Political Language. In: Nimmo, Dan and Keith R. Sanders (eds.) *Handbook of Political Communication.* Beverly Hills, CA: Sage, 195–224.

Grabowski, Łukasz (2015) Towards teaching English for pharmaceutical purposes: An attempt at a description of key vocabulary and phraseology in clinical trial protocols and European public assessment reports. In: Piasecka, Liliana, Małgorzata Adams-Tukiendorf and Przemysław Wilk (eds.) *New Media and Perennial Problems in Foreign Language Learning and Teaching.* Berlin: Springer, 209–228.

Groom, Nicholas (2010) Closed-class keywords and corpus-driven discourse analysis. In Bondi, Marina and Mike Scott (eds.) *Keyness in Texts*. Amsterdam: John Benjamins, 59–78.

Halliday, M. A. K (1994) *An Introduction to Functional Grammar*. London: Edward Arnold.

Halmari, Helena (2008) On the language of the Clinton-Dole presidential campaign debates: General tendencies and successful strategies. *Journal of Language and Politics* 7(2), 247–270.

Homolar, Alexandra and Ronny Scholz (2019) The power of Trump-speak: Populist crisis narratives and ontological security. *Cambridge Review of International Affairs* 32(3), 344–364.

Hunston, Susan (2017) Donald Trump's inaugural speech – and Barack Obama's. *Language in the News*. Accessed on 1 April 2022 from https://blog.bham.ac.uk/susanhunston/donald-trumps-inaugural-speech-and-barack-obamas/20170130.

Karapetjana, Indra (2011) Pronominal choice in political interviews. *Baltic Journal of English Language, Literature and Culture* 1, 36–45.

Kerner, Milena (2018) Rambling like a president: Lose the grammar and Trump it up! *America in the Time of Trump* March 21st. Accessed on 22 October 2019 from https://crl.parisnanterre.fr/medias/fichier/america-in-the-time-of-trump-magazine-c1-_1531723140728-pdf.

Koteyko, Nelya (2007) A diachronic approach to meaning: English loanwords in Russian opposition discourse. *Corpora* 2 (1), 65–95.

Kutter, Amelie (2017) Corpus analysis. In: Wodak, Ruth and Brnhard Forchtner (eds.) *The Routledge Handbook of Language and Politics*. London: Routledge, 169–186.

Leifeld, Philip (2014) Polarization of coalitions in an agent-based model of political discourse. *Computational Social Networks* 1 (7), 1–22.

Lorenz, Gunter (2002) Really worthwhile or not really significant? A corpus-based approach to the delexicalization and grammaticalization of intensifiers in Modern English. In: Wischer, Ilse and Babriele Giewald (eds.) *New Reflections on Grammaticalization*. Amsterdam: John Benjamins, 143–161.

Lorenz, David (2013) From reduction to emancipation: Is *gonna* a word? In: Hasselgård, Hilde, Jarle Ebeling and Signe Oksefjell Ebeling (eds.) *Corpus Perspectives on Patterns of Lexis*. Amsterdam: John Benjamins, 133–152.

Lyons, John (1977) *Semantics*. Vol. 2. Cambridge: Cambridge University Press.

McWhorter, John (2018) What Trump's speech says about his mental fitness. *The New York Times* Feb 6th. Accessed on 24 October from https://www.nytimes.com/interactive/2018/02/06/opinion/trump-speech-mental-capacity.html.

Milizia, Denise (2010) Keywords and phrases in political speeches. In: Bondi, Marina and Mike Scott (eds.) *Keyness in Texts*. Amsterdam: John Benjamins, 127–145.

Milizia, Denise and Cinzia Spinzi (2010) The 'terroridiom' principle between spoken and written discourse. In: Römer, Ute and Rainer Schulze (eds.) *Patterns, Meaningful Units and Specialized Discourses*. Amsterdam: John Benjamins, 57–85.

Morgenthau, Hans J. (1946) *Scientific Man vs. Power Politics*. Chicago: University of Chicago Press.

Östman, Jan-Ola (1981) *You Know: A Discourse Functional Approach*. Amsterdam: John Benjamins.

Pan, Jun (2019) The pragmatics of political discourse: An analytical framework and a comparative study of policy speeches in the United Kingdom and Hong Kong. *Bandung* 6 (2), 252–284.

Partington, Alan (2003) *The Linguistics of Political Argument: The Spin-Doctor and the Wolf-Pack at the White House*. London: Routledge.

Partington, Alan (2013) Corpus analysis of political language. In: Chapelle, Carol (ed.) *The Encyclopedia of Applied Linguistics*. Oxford: Blackwell Publishing, 1299–1307.

Patrick, Vanessa M. and Hagtvedt, Henrik (2012) How to say "no": Conviction and identity

attributions in persuasive refusal. *International Journal of Research in Marketing 29* (4), 390–394.
Pearce, Michael (2005) Informalization in UK party election broadcasts 1966–97. *Language and Literature* 14 (1), 65–90.
Pichler, Heike (2013) *The Structure of Discourse-Pragmatic Variation*. Amsterdam: John Benjamins.
Polge, Olivier (2015) *I want you to* and *I need you to* in modern and present-day English. In: Zamorano-Mansilla, Joan Rafael and Carmen Maíz, Elena Domínguez and Victoria Martín de la Rosa (eds.) *Thinking Modally: English and Contrastive Studies on Modality*. Cambridge: Cambridge Scholars Publishing, 23–45.
Quirk, Randolph and Sidney Greenbaum, Geoffrey Leech Jan and Svartvik (1985) *A Comprehensive Grammar of the English Language*. London and New York: Longman.
Sauer, Christoph (2002) Ceremonial text and talk: A functional-pragmatic approach. In: Chilton, Paul and Christina Schäffner (eds.) *Politics as Text and Talk: Analytic Approaches to Political Discourse*. Amsterdam: John Benjamins, 111–141.
Savoy, Jacques (2018) Trump's and Clinton's style and rhetoric during the 2016 presidential election. *Journal of Quantitative Linguistics* 25 (2), 168–189.
Sclafani, Jennifer (2016) The idiolect of Donald Trump. *Scientific American* March 25th. Accessed on 22 October 2019 from https://blogs.scientificamerican.com/mind-guest-blog/the-idiolect-of-donald-trump.
Sclafani, Jennifer (2018) *Talking Donald Trump. A Sociolinguistic Study of Style, Metadiscourse, and Political Identity*. London: Routledge.
Scott, Mike and Christopher Tribble (2006) *Textual Patterns. Key Words and Corpus Analysis in Language Education*. Amsterdam: John Benjamins.
Scott, Mike. (2012) *WordSmith Tools version 6*. Stroud: Lexical Analysis Software.
Shelly, Leanne (2009) *Say it Like Obama and Win! The Power of Speaking with Purpose and Vision*. US: McGraw-Hill.
Simms, Karl (2018) One year of Trump: Linguistics expert analyses US President's influence on language. *University of Liverpool News* January 19th. Accessed on 24 October 2019 from https://news.liverpool.ac.uk/2018/01/19/one-year-trump-linguistics-expert-analyses-us-presidents-influence-language/.
Smith, Zadie (2009) Speaking in tongues. *New York Review of Books* 56(3). Accessed on 21 October 2019 from http://www.nybooks.com/articles/22334.
Sojda, Sylwia (2019) Adverbial intensifiers in contemporary Polish and Slovak. *Jazykovedný časopis* [Journal of Linguistics] 70 (1), 55–72.
Szudarski, Paweł (2018) *Corpus Linguistics for Vocabulary: A Guide for Research*. London: Routledge.
Tagliamonte, Sali A (2008) So different and pretty cool! Recycling intensifiers in Toronto, Canada. *English Language and Linguistics* 12 (2), 361–394.
Tsui, Amy B. M. (1991) The pragmatic functions of *I don't know*. *Text* 11, 607–622.
Van Dijk, Teun (1998) What is political discourse analysis? In: Blommaert, Jan and Chris Bulcaen (eds.) *Political Linguistics*. Amsterdam: John Benjamins, 11–52.
Van Dijk, Teun (2003) Knowledge in parliamentary debates. *Journal of Language and Politics* 2 (1), 93–129.
Van Hout, Tom and Van Praet, Ellen (2011) Competence on display: Crafting stories during newsroom editorial conferences. In: Pelsmaekers, Katia, Craig Rollo, Tom Van Hout and Priscilla Heynderickx (eds.) *Displaying Competence in Organizations. Discourse perspectives*. Basingstoke, Hampshire, UK: Palgrave Macmillan, 111–130.
Wang, Yaquin and Haitao Liu (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. *Discourse and Society* 29 (3), 299–323.

Watts, Richard J. (1984) An analysis of epistemic possibility and probability. *English Studies* 65 (2), 129–140.
Willis, Rebecca (2017) Taming the climate? Corpus analysis of politicians' speech on climate change. *Environmental Politics* 26 (2), 212–231.
Wilson, John. 1990. *Politically Speaking: The Pragmatic Analysis of Political Language*. Oxford: Blackwell.
Wodak, Ruth (2009) *The Discourse of Politics in Action: Politics as Usual*. Basingstoke: Palgrave Macmillan.
Zheni, Thouraya (2019) Factive vs. ideological knowledge in political discourse. *International Journal of English Linguistics* 9 (1), 36–53.

Tatiana Szczygłowska is Assistant Professor at the Institute of Neophilology, University of Bielsko-Biala, Poland, where she lectures on translation theory and text analysis to students of English philology. Her main research and teaching interests lie in linguistics and translation with a special focus on academic discourse and corpus linguistics. She has also conducted studies in the area of students' attitudes to target language culture and pronunciation.

Address: Tatiana Szczygłowska, Institute of Neophilology, University of Bielsko-Biala, ul. Willowa 2, 43-309 Bielsko-Biała, Poland. [email: tszczyglowska@ath.bielsko.pl]

This work can be used in accordance with the Creative Commons BY-NC-ND 4.0 International license terms and conditions (https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode). This does not apply to works or elements (such as image or photographs) that are used in the work under a contractual license or exception or limitation to relevant rights.