Quantitative Analysis of US Presidential Inaugural Addresses

Miroslav Kubát, Radek Čech[1]

Abstract. The research aims to investigate several features of inaugural addresses of the presidents of the United States. The goal of the paper is to observe the presidential speeches from a viewpoint of stylometry indices and to discover whether political and historical circumstances (wars, financial crisis, ideology, etc.) influence the style of inaugural addresses, analogically to findings presented by Čech (2014). Specifically, vocabulary richness, thematic concentration and text activity are computed. These three indices were chosen especially due to (a) their high efficiency of automatic text classification (genre analysis, authorship attribution, etc.), (b) their independence on text length and (c) simple linguistic interpretation. The combination of the three methods allows both to investigate the style of the particular presidential speeches in powerful linguistically comprehensive view and to observe the development trends of the specific genre of inaugural addresses during the more than 200 years long history. The corpus comprises inaugural addresses of all US presidents from George Washington to Barack Obama (57 texts in total).

Keywords*: stylometry, presidential speeches, vocabulary richness, thematic concentration, activity*

1. Introduction

Political speeches are widely used in linguistic research, especially in discourse analysis (e.g. Lim, 2004; Carranza, 2008; Matić, 2012). Several quantitative analyses have dealt with this issue (e.g. Čech, 2014; Savoy, 2010; Tuzzi et al., 2010). It is not surprising therefore that addresses of the US presidents are frequently investigated because the American President can be ranked among the most powerful politicians of the contemporary world. In this study, we analyse all US presidential inaugural addresses. While most analyses deal with these data in terms of qualitative methods or content analysis, we focus on the issue from a viewpoint of stylometric indices of contemporary quantitative linguistics, particularly vocabulary richness, secondary thematic concentration, and text activity. These methods have proved to be an effective tool in political language research. Promising results were obtained by Čech (2014) who analysed an impact of ideology on a character of annual messages given by Czech and Czechoslovak presidents. Another related research was done by Tuzzi et al. (2010) who examined end-of-year speeches of Italian presidents.

The aim of this study is to analyse relationships between some characteristics of the style of US presidential speeches and certain pragmatic aspects which could have an impact on the addresses, specifically, historical development, ideology, financial crises, and wars. It is important to emphasise that this study is but a first insight into the issue and our approach is rather heuristic.

---

[1] University of Ostrava, Dept. of Czech Language. Czech Republic. Address correspondence to: cechradek@gmail.com

2. Language material

The inaugural address is a habitual part of the inauguration procedure. Except for con-stitutionally required presidential oath of office all parts of the inauguration procedure (including inaugural speech) are optional given by tradition. This is the reason why several presidents (particularly John Tyler, Millard Fillmore, Andrew Johnson, Chester A. Arthur, Calvin Coolidge) gave no address. In each of these cases, the incoming president substituted a president who had died.

This genre provides unique data for quantitative linguistic research because of homogeneity of the genre and its long tradition. In this study, 57 addresses were analysed. The list of all addresses with the results can be found in the appendix of the article. The data was collected by the American Presidency Project (Peters and Woolley, 2015).

3. Methodology

We use three methods to investigate some aspects of the style of US presidential addresses, mainly the vocabulary richness (*MATTR*), secondary thematic concentration (*STC*), and text activity (*Q*). These indices were chosen due to (a) high efficiency of automatic text clas-sification (genre analysis, authorship attribution, etc.), (b) their independence on text length, and (c) simple linguistic interpretation. The vocabulary richness was computed by *MaWaTaTaRaD* software (Milička, 2013); the thematic concentration and the activity were computed by *QUITA - Quantitative Index Text Analyzer* (Kubát et al., 2014).

3.1 Moving Average Type-Token Ratio (MATTR)

The measurement of vocabulary richness is one of the oldest quantitative methods in stylo-metry, with more than seventy years long history (cf. Popescu et al., 2009). A large number of indices of vocabulary richness has been set up in linguistics; however, almost all of them evidence an undesirable dependence on the length of the text. To avoid this dependence in our analysis, we use the moving average type-token ratio (*MATTR*), proposed by Covington and McFall (2010), which was experimentally proved to be independent of the text size (see Kubát, 2014).

The *MATTR* is defined as follows. A text is divided into overlapped subtexts of the same length (so called "windows" with arbitrarily chosen size *L*; usually, the "window" moves forward one token at a time), next, the type-token ratio is computed for every subtext and, finally, the *MATTR* is defined as a mean of particular values. For example, in the following sequence of characters: *a, b, c, a, a, d, f,* text length is 7 tokens (*N* = 7) and we choose the window size of 3 tokens (*L* = 3). We get subsequent 5 windows: *a, b, c | b, c, a | c, a, a | a, a, d | a, d, f,* and compute *MATTR* of the sequence as follows:

$$MATTR(L) = \frac{\sum_{i=1}^{N-L} V_i}{L(N-L+1)} = \frac{3+3+2+2+3}{3(7-3+1)} = 0.87$$

L…arbitrarily chosen length of a window, *L* < *N*
N…text length in tokens
V_i…number of types in an individual window

3.2 Secondary Thematic Concentration

The secondary thematic concentration (*STC*) is a method which measures the degree of intensity with which the author focuses on a topic (or topics) of a given text (cf. Čech et al., 2015). Specifically, the *STC* is based on two text characteristics: 1) the frequency distribution of words and 2) the so called *h*-point (cf. Popescu, 2007). The *h*-point is defined as a point where the frequency equals rank (see formula 1); it separates in a fuzzy way the most productive synsemantics from autosemantics in a rank frequency distribution of words or lemmas (for more details cf. Popescu et al., 2009, p. 17ff). Specifically,

$$(1) \quad h = \begin{cases} r_i, & \text{if there is } r_i = f(r_i) \\ \dfrac{f(r_i)r_{i+1} - f(r_{i+1})r_i}{r_{i+1} - r_i + f(r_i) - f(r_{i+1})} & \text{if there is } r \neq f(r) \end{cases} ,$$

where r_i is the rank and $f(r_i)$ is the respective frequency of this rank; given that r_i is the highest number for which $r_i < f(r_i)$ and r_{i+1} is the lowest number for which $r_{i+1} > f(r_{i+1})$. Thus, if no rank is equal to the respective frequency, one computes the lower part of formula (1) consisting of neighbouring values. Having stated the *h*-point, all autosemantics occurring at lower ranks are considered as thematic words because they signalize the frequent repetition of the given autosemantics.[2] (Čech et al., 2015). The *h*-point is multiplied by two in the concept of the *STC*, on reasons presented in Čech et al. (2015). The thematic weight (*TW*) of each thematic word can be computed and, finally, the *STC* is obtained as the sum of these weights (*TW*), specifically

$$(2) \quad STC = \sum_{r'=1}^{2h} TW = \sum_{r'=1}^{2h} \frac{(2h - r')f(r')}{h(2h-1)f(1)} ,$$

where *r'* is the rank of autosemantic word above *h*-point and *h* is the *h*-point. For illustration, we present here the computation of the *STC* of the Lincoln's inaugural address (see the Text 20 in Appendix and Table 1).

Table 1
The rank-frequency distribution of Text 20. *h* = 9.

Token	Rank	Average rank	Frequency	Token	Rank	Average rank	Frequency
the	1	1	58	for	11	10	9
to	2	2	27	with	12	12.5	8
and	3	3	24	be	13	12.5	8
of	4	4	22	this	14	14.5	7
it	5	5	13	a	15	14.5	7
war	**6**	**6.5**	**12**	by	16	17.5	6
that	7	6.5	12	we	17	17.5	6
all	8	8	10	is	18	17.5	6
in	9	10	9	**god**	**19**	**17.5**	**6**
which	10	10	9				

---

[2] It should be mentioned that not all autosemantics need be considered to express the thematic properties of the text; for instance Popescu et al. (2009) use only nouns and their predicates of the first order, i.e. adjectives and verbs. In this paper, this approach is followed.

$$STC_{Text\,20} = \sum_{r'=1}^{2h} \frac{(2h-r')f(r')}{h(2h-1)f(1)} = \frac{(2\cdot 9 - 6.5)12}{9(2\cdot 9 - 1)58} + \frac{(2\cdot 9 - 17.5)6}{9(2\cdot 9 - 1)58} = 0.0159$$

3.3 Activity

Each text focuses more intensively either on the action (plot) or on the description. For instance, travel books focus principally on description and, conversely, short stories concentrate on the plot. The concept of the activity and descriptiveness was introduced by Busemann (1925). Generally, the text activity is represented by verbs and the descriptiveness by adjectives. Index of activity Q is defined as a ratio of verbs V and the sum of verbs V and adjectives A in the text, see formula (3):

$$(3)\quad Q = \frac{V}{V+A}$$

For illustration, the activity Q of the Lincoln's inaugural address (Text 20 in Appendix) is

$$Q_{Text\,20} = \frac{V}{V+A} = \frac{102}{102+36} = 0.74\ ,$$

which expresses high activity of the text.

3.4 Statistical comparison

In this study, differences between results are tested by means of the u-test[3], see formula 4

$$(4)\quad u = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}\quad,$$

where, $\bar{X}_1$, $\bar{X}_2$...arithmetic mean of results in each group,
S_1, S_2...standard deviation,
n_1, n_2...number of results in each group.
  Since the threshold is 1.96, $u \geq 1.96$ means that the difference between two groups is statistically significant for the significance level $\alpha = 0.05$.

4. Results

4.1 Historical development

Firstly, we focus on the historical development of all US presidential inaugural addresses. The chronologically ranked resulting values are presented in Figure 1.

---

[3] In statistics, it is sometimes called z-test; here, we follow a convention used in the quantitative linguistics.
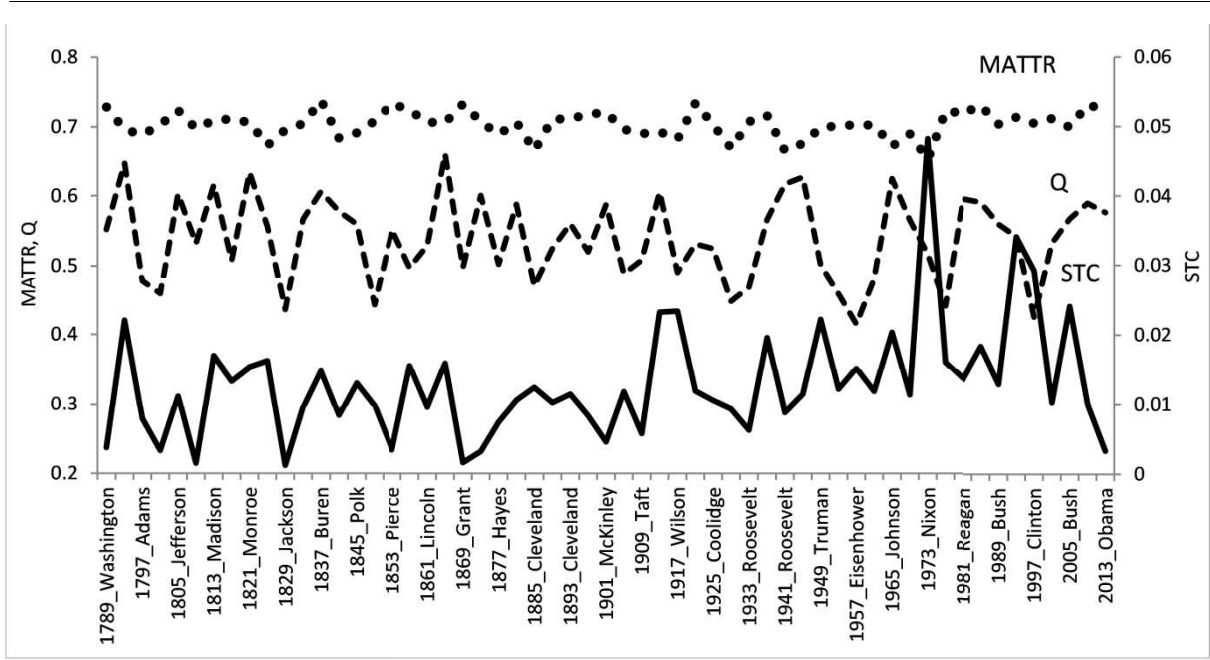
Figure 1. Chronologically ranked values of *MATTR*, *Q*, and *STC* of the US presidential inaugural addresses.

As can be seen in Figure 1, there is no tendency at first sight. The obtained values of the indices oscillate chronologically up and down without any obvious reason. The results seem to be a matter of individual style of each president rather than historical circumstances. Nevertheless, we try to find out whether the style of the addresses is influenced by some pragmatic causes, namely: political affiliation, war, and financial crisis.

4.2 Political affiliation

Throughout most of the American history, a two-party system dominated. Since 1852, every American president has been presented as a candidate either of Democratic or Republican political party. Before this date, the political affiliation of particular president was not so evident; consequently, we use only the addresses from 1852 for the analysis of the potential impact of political affiliation on the style. Theoretically, the political affiliation can influence political speeches because of different ideological basis (cf. Čech, 2014). Our aim is to discover whether inaugural addresses of democratic presidents differ from the republican ones. The resulting values are presented in Table 2.

Table 2
MATTR, *Q*, *STC* resulting values and statistical comparison of democrats and republicans

	democratic	republican	*u*
MATTR	0.70	0.70	0.13
Q	0.54	0.54	0.08
STC	0.015	0.012	0.82

The results in Table 2 show that there is no significant difference (at the significant level $\alpha = 0.05$). Surprisingly enough, the values of *MATTR* and *Q* display even no difference at all. Thus, we can state that political affiliation has no impact on the style of inaugural addresses in terms of the measured indicators. More detailed view of the issue is displayed in Figure 2 and 3 where the style of addresses is expressed as relation *MATTR-Q* and *MATTR-STC*.
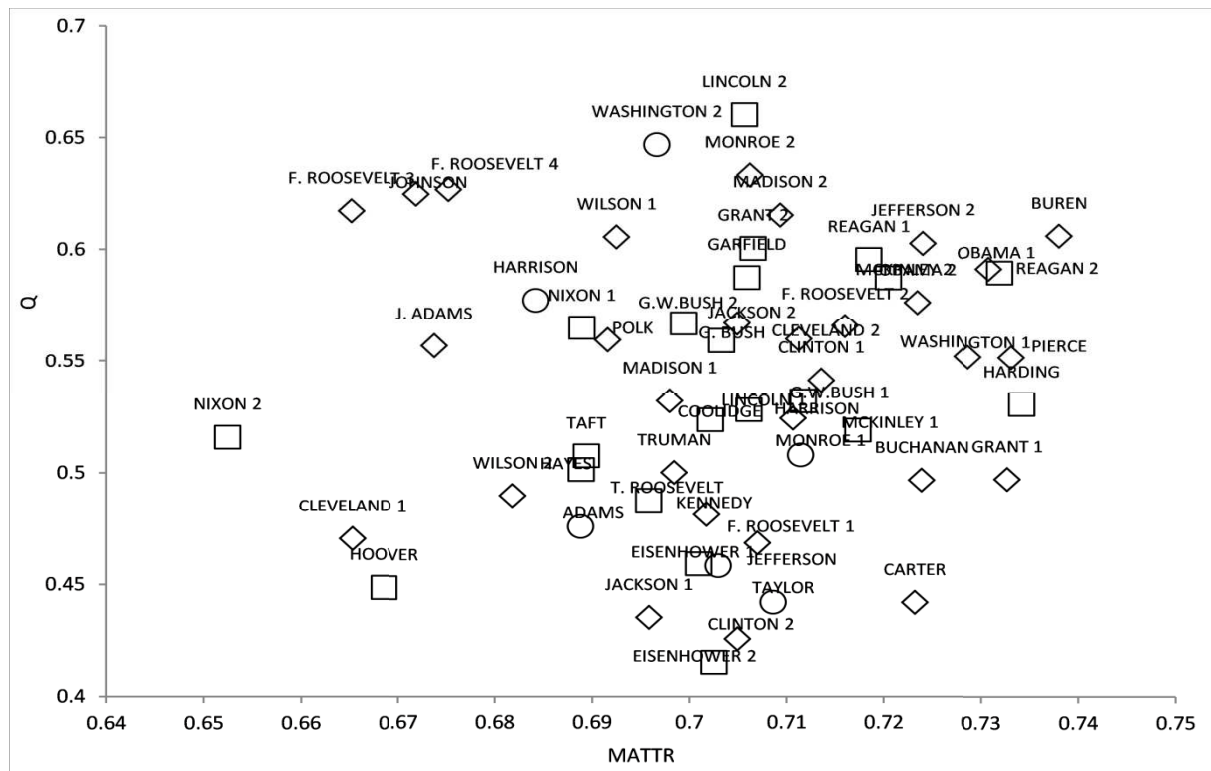
Figure 2. The relation between *MATTR* and *Q* in inaugural addresses;
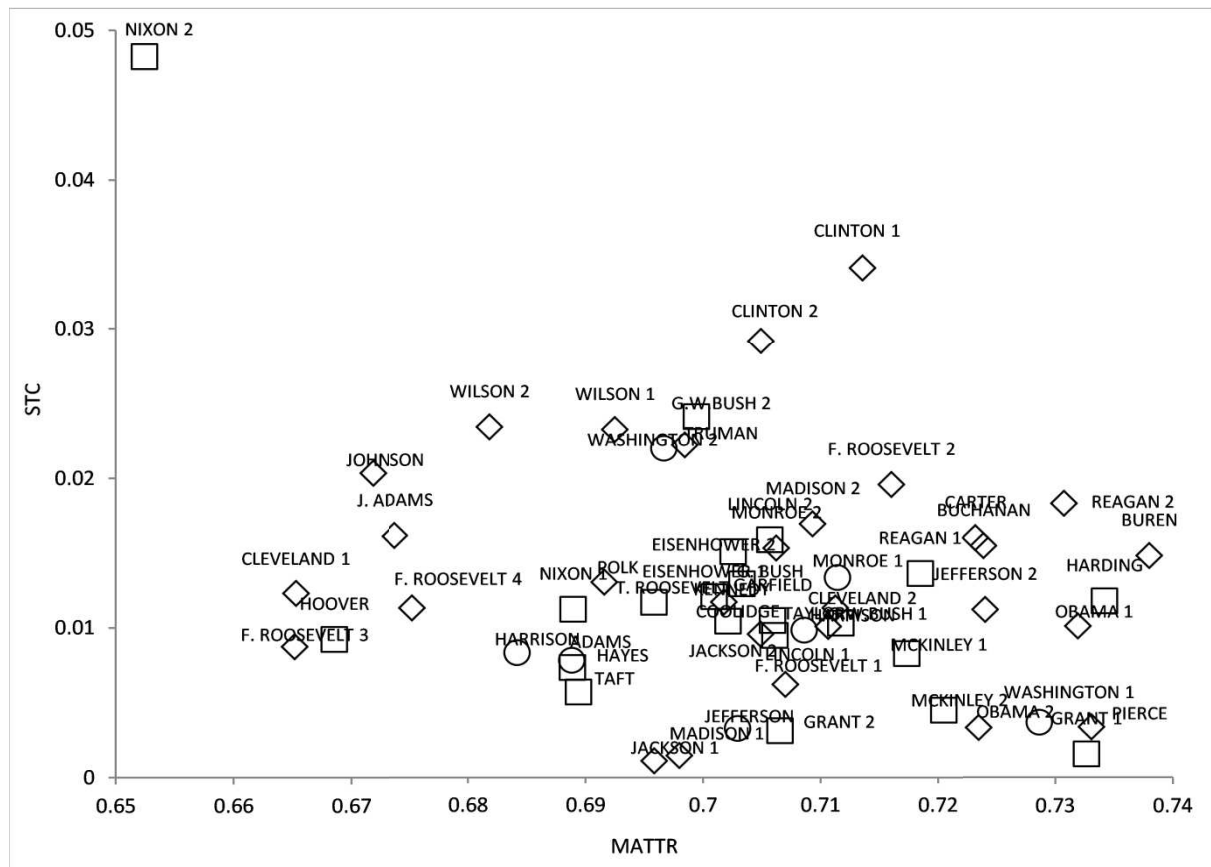square = republican, diamond = democratic, circle = others.

Figure 3. The relation between *MATTR* and *STC* in inaugural addresses;
square = republican, diamond = democratic, circle = others.

4.2 Wartime

A war affects a society in many ways, especially "big" ones such as the First and Second World War. For politicians, a war usually represents one of the most important topics in their political agenda and the wartime can be interpreted as an extraordinary era (in contrast to peacetime). This fact could be reflected by different style of wartime political speeches (in contrast to peacetime speeches).

However, the history of the USA, as of any other country, seems to be a series of various wars and it is difficult to decide which era can be assigned as the peacetime and which as the wartime. For example, let us consider the Cold War, the long era of strained and polarized relations between East and West. On the one hand, it was not a real war in fact; on the other hand, the cold war was one of the biggest wars in terms of number of arms, its impact to the particular societies, and danger of nuclear arms usage and so on. It is even hard to decide how long this war lasted.

Considering the aforementioned methodological problems, we decide to distinguish peacetime and wartime according to the US military expenditures (in percent of GDP). We choose 4% value as the border which seems to be suitable to distinguish the worst wars in US history (see Figure 4). Although this threshold is an arbitrary chosen value just for the purpose of this study, this method allows us to reasonably distinguish between wartime and peacetime.
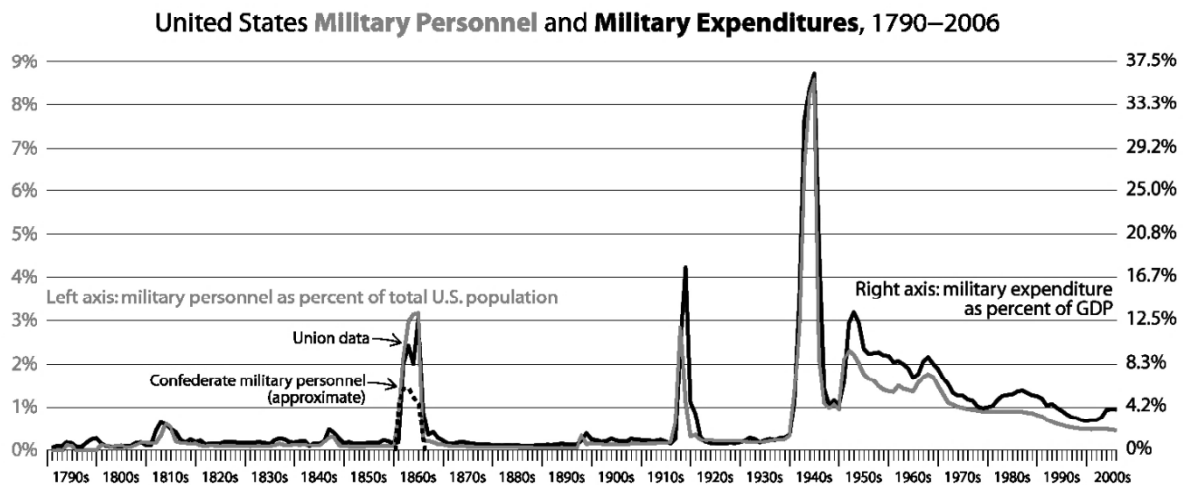
Figure 4. Military expenditures and military personnel in US history. Source: https://en.wikipedia.org/wiki/File:US_military_personnel_and_expenditures.png

As can be seen in Figure 4, only several wars are considered as wartime according to this criterion, namely the American Civil War, First World War, Second World War, Cold War, Korean War, Vietnam War, and partly ongoing War on Terror.

Table 3
MATTR, Q, STC resulting values and statistical comparison of wartime and peacetime.

	wartime	peacetime	*u*
MATTR	0.70	0.71	0.96
Q	0.55	0.53	0.94
STC	0.017	0.010	**2.98**

_____

The results in Table 3 show that wartime and peacetime addresses significantly differ only in the case of the *STC*. This result is in accordance with our assumption (see above) and is probably caused by the fact that in wartime era the war is really dominant topic whereas in peaceful era president tends to talk about more topics. This statement can be supported by the findings of totalitarian language (Čech, 2014). The *MATTR* and *Q* rather reflect the style of speeches; the results reveal that it (at least in the case of observed characteristics) is not influenced by the wartime.

4.3 Financial crisis

Aside from war, recession is one of the worst eras for people. The financial crises often trigger strikes, social unrests, and sometime even wars. There are several options how to determine financial crises through the history. We decide to use the unemployment rate which influences significantly the standard of living and is directly caused by recession. Since we do not have data before 1890, we must analyse only the period after this year; moreover, the values of unemployment between 1890 and 1940 are only estimated. Nevertheless, from the Figure 5 is obvious that there are two extraordinary periods where the unemployment exceeded 10%, particularly 1894-1898 and 1931-1939. Five hundred banks closed, 15,000 businesses failed, and the unemployment hit 35% in New York and even 43% in Michigan in the first serious economic depression starting in 1883, just thirteen days before the inauguration of G. Cleveland. The second financial crisis known as "The Great Depression" was the longest, deepest and most widespread depression of 20th century. This crisis started after the stock market crash of October 29, 1929 (known as Black Tuesday). The effect on people was enormous: more than 5000 banks failed, unemployment rate exceeded 20%, and hundreds of thousands found themselves homeless. The resulting values of *MATTR*, *Q*, *STC* and statistical comparison are displayed in Table 3.

Figure 5. Unemployment rate in US history. Source:
https://commons.wikimedia.org/wiki/File:US_Unemployment_1890-2009.gif

Table 3
MATTR, Q, STC resulting values and statistical comparison of normal and crisis

	normal	crisis	*u*
MATTR	0.70	0.71	**2.70**
Q	0.53	0.53	0.21
STC	0.016	0.011	1.35

As can be seen in Table 3, despite small difference in terms of *MATTR* resulting values (0.70 and 0.71), only vocabulary richness significantly distinguishes normal time and recession. Activity seems to be irrelevant in terms of crisis and despite some difference between thematic concentration results (0.016 and 0.011); the statistical test does not prove significant difference.

4.4 Thematic Words

The method of measurement of thematic concentration allows extraction of the so called thematic words, i.e. words which represent main topic(s) of text. The thematic words (*TW*) can be viewed as an alternative to keywords (cf. Čech et al. 2015). The advantage of *TW* lies in the fact that those words are based solely on the frequency structure distribution of the text; no reference corpus is needed for the analysis. The list of *TW* of all inaugural addresses is displayed in Table 4. The complete list of *TW* of each presidential speech can be found in the Appendix   .

Table 4
Frequency list of thematic words of all inaugural addresses (*f* ≥ 2).

#	Word	f	#	Word	f
1	have	46	23	freedom	5
2	government	29	24	other	5
3	people	28	25	citizens	4
4	has	22	26	war	4
5	been	20	27	law	3
6	world	13	28	let	3
7	who	11	29	what	3
8	country	11	30	nations	3
9	great	10	31	united	3
10	nation	10	32	own	3
11	states	10	33	had	3
12	shall	9	34	congress	2
13	more	9	35	power	2
14	America	7	36	free	2
15	peace	7	37	democracy	2
16	union	6	38	life	2
17	do	6	39	were	2
18	was	6	40	state	2
19	new	6	41	liberty	2
20	Public	6	42	time	2
21	constitution	6	43	spirit	2
22	such	5	44	justice	2

As can be seen in Table 4, most words are concentrated on the state and its citizens (e.g. *government, people, country, nation, America, union, public, citizens*). There are also several words connected to freedom such as *peace, freedom, free, democracy, or liberty* which comply with officially declared principles of USA. We can also see that adjectives among thematic words are positive (e.g. *great, new, free*); probably in order to ensure people that the new president will bring better future. There is just one word which expresses negative connotations – *war*.

Liu (2012) claims that the US presidential inaugural addresses consist of eight general parts. With the exception of salutation and other formalities such as announcing entering upon office or articulating sentiments on the occasion, Liu (2012) identifies following parts:

a) Making pledges – "The new president carries out this speech act to help the pubic with confidence in the new leader and his government." (Liu, 2012, p. 2410)
b) Arousing patriotism in citizens
c) Announcing political principles to guide the new administration – "The basic principles that all presidents swear to follow comprise American Constitution, union, freedom and democracy…" (Liu, 2012, p. 2410)
d) Resorting to religious power: "Every president will refer to God many times in his inaugural address as God is the common religious belief for nearly all Americans." (Liu, 2012, p. 2411)

As can be seen in Table 4, the thematic words comply with aforementioned themes. Only resorting to religious power do not fully correspond to *TW*, because *God* occurs only one time (Text 20, Lincoln).

5. Conclusion and Discussion

This study analyses the vocabulary richness (*MATTR*), text activity (*Q*), and secondary thematic concentration (*STC*) of US presidential inaugural addresses. We discovered that there is no obvious general tendency through the more than two centuries long history and the style of the speeches is rather influenced by personality of each president. We also found out that the aforementioned features are not relevant to the political affiliation. In these aspects our findings are different from those in Czech presidential speeches (cf. Čech 2014). On the other hand, we discovered that the addresses in wartime significantly differ in terms of secondary thematic concentration. Another difference was found in recession time where vocabulary richness is significantly higher.

To sum up, US presidential inaugural addresses seem to be mostly determined by individual style of each speaker but some important circumstances such as war or recession can affect the speech to some extent. Finally, it is necessary to say that this work is just a first attempt to analyse the US presidential addresses by the aforementioned indices. Therefore, more analyses must be done to support or reject our preliminary claims.

References

Busemann, A. (1925). *Die Sprache der Jugend als Ausdruck der Entwicklungsrhythmik.* Jena: Fischer.
Carranza, I. E. (2008). Strategic political communication: a leader's address to the nation. *Nueva Revista de Lenguas Extranjeras 10, 25–56.*

Čech, R. (2014). Language and ideology: Quantitative thematic analysis of New Year speeches given by Czechoslovak and Czech presidents (1949-2011). *Quality & Quantity 48, 899–910.*

Čech, R., Garabík, R., Altmann, G. (2015). Testing the thematic concentration of text. *Journal of Quantitative Linguistics 22, 215–232.*

Covington, M.A., McFall, J.D. (2010) Cutting the Gordian Knot: The Moving-Average Type–Token Ratio (MATTR). *Journal of Quantitative Linguistics 17, 94–100.*

Kubát, M. (2014). Moving window type-token ratio and text length. In: Altmann, G., Čech, R., Mačutek, J., Uhlířová, L. (eds.), *Empirical Approaches to Text and Language Analysis: 105–113.* Lüdenscheid: RAM.

Kubát, M., Matlach, V., Čech, R. (2014). *QUITA - Quantitative Index text Analyzer.* Lüdensheid: RAM.

Lim, E.T. (2004). Five Trends in Presidential Rhetoric: An Analysis of Rhetoric from George Washington to Bill Clinton. *Presidential Studies Quarterly 32, 328–348.*

Liu, F. (2012). Genre Analysis of American Presidential Inaugural Speech. *Theory and Practice in Language Studies 2, 2407–2411.*

Matić, D. (2012) Ideological discourse structures in political speeches. *Komunikacija i kultura online 3, 54-78.*

Milička. J. (2013) *MaWaTaTaRaD* (software). Available at http://milicka.cz/en/mawatatarad/

Peters, G., Woolley, J.T. (2015). *The American Presidency Project.* http://www.presidency.ucsb.edu

Popescu, I.-I., Altmann, G., Grzybek, P., Jayaram, B.D., Köhler, R., Krupa, V., Mačutek, J., Pustet, R., Uhlířová, L., Vidya, M.N. (2009). *Word frequency studies.* Berlin/New York: Mouton de Gruyter.

Popescu, I.-I. (2007). Text Ranking by the Weight of Highly Frequent Words. In Grzybek P, Köhler, R. (Eds.) *Exact Methods in the Study of Language and Text*: 555–565. Berlin – New York: Mouton de Gruyter,

Savoy, J. (2010). Lexical Analysis of US Political Speeches. In: *Journal of Quantitative Linguistics 17, 123–141.*

Tuzzi, A., Popescu, I.-I., Altmann, G. (2010). *Quantitative Analysis of Italian Texts.* Lüdenscheid: RAM.

Appendix

#	Year	President	Types	Tokens	MATTR	Q	STC	Thematic words
1	1789	Washington	594	1431	0.73	0.55	0.004	HAVE, GOVERNMENT, MORE
2	1793	Washington	90	135	0.70	0.65	0.022	SHALL
3	1797	Adams	794	2322	0.69	0.48	0.008	PEOPLE, GOVERNMENT, NATIONS, MORE, COUNTRY
4	1801	Jefferson	679	1732	0.70	0.46	0.003	GOVERMENT, HAVE
5	1805	Jefferson	777	2168	0.72	0.60	0.011	HAVE, PUBLIC, WHO, CITIZENS, FELLOW, STATE
6	1809	Madison	521	1177	0.70	0.53	0.001	HAVE, BEEN
7	1813	Madison	518	1211	0.71	0.62	0.017	HAVE, WAR, BEEN
8	1817	Monroe	980	3379	0.71	0.51	0.013	HAVE, BEEN, GOVERNMENT, STATES, GREAT, HAS, OTJER, PEOPLE, UNITED
9	1821	Monroe	1195	4466	0.71	0.63	0.015	BEEN, HAVE, HAS, GREAT, STATES, WERE, OTHER, WAS,

								WAR, UNITED, MADE, CITIZENS, SUCH, HAD, GOVERNMENT
10	1825	Adams	961	2917	0.67	0.56	0.016	HAVE, BEEN, HAS, UNION, GOVERNMENT, RIGHTS, OTHER, COUNTRY
11	1829	Jackson	500	1128	0.70	0.44	0.001	PUBLIC, HAVE
12	1833	Jackson	474	1177	0.70	0.57	0.010	GOVERNMENT, PEOPLE, UNION, STATES, HAVE
13	1837	Buren	1252	3846	0.74	0.61	0.015	HAS, HAVE, BEEN, PEOPLE, WAS, INSTITUTIONS, GOVERNMENT, WERE
14	1841	Harrison	1799	8469	0.68	0.58	0.008	HAVE, POWER, HAS, PEOPLE, BEEN, CONSTITUTION, GOVERNMENT, WAS, CITIZENS, OTHER, STATES, EXECUTIVE, COUNTRY, GREAT, SPIRIT, MORE, CHARACTER, SUCH, LIBERTY, STATE
15	1845	Polk	1255	4814	0.69	0.56	0.013	GOVERNMENT,STATES,HAVE, UNION, HAS, BEEN, POWERS, PEOPLE,COUNTRY, CONSTI-TUTION, INTERRESTS
16	1849	Taylor	481	1091	0.71	0.44	0.010	SHALL,GOVERNMENT, COUNTRY
17	1853	Pierce	1114	3344	0.73	0.55	0.003	HAVE, HAS, POWER, BEEN, GOVERMENT
18	1857	Buchanan	889	2836	0.72	0.50	0.015	HAS, STATES, HAVE, SHALL, CONSTITUTION,BEEN, GOVERNMENT,PEOPLE, QUESTION, GREAT
19	1861	Lincoln	1005	3635	0.71	0.53	0.010	CONSTITUTION, HAVE, PEOPLE, UNION, STATES, GOVERNMENT, SHALL, SUCH, LAW, DO
20	1865	Lincoln	335	705	0.71	0.66	0.016	WAR, GOD
21	1869	Grant	466	1132	0.73	0.50	0.002	COUNTRY
22	1873	Grant	520	1339	0.71	0.60	0.003	HAVE, BEEN, COUNTRY, WAS
23	1877	Hayes	798	2489	0.69	0.50	0.007	COUNTRY, GOVERNMENT, HAVE, STATES, PUBLIC, POLITICAL, HAS, PEOPLE, GREAT
24	1881	Garfield	966	2987	0.71	0.59	0.010	GOVERNMENT, HAVE, PEOPLE, HAS, STATES, CONSTITUTION, BEEN, UNION, GREAT, LAW
25	1885	Cleveland	643	1691	0.67	0.47	0.012	PEOPLE, GOVERNMENT, PUBLIC, WHO, SHALL, CONSTITUTION
26	1889	Harrison	1299	4398	0.71	0.52	0.010	HAVE, PEOPLE, BEEN, WHO, STTES, HAS, SHALL, LAWS, PUBLIC, WAS
27	1893	Cleveland	794	2028	0.71	0.56	0.011	PEOPLE, GOVERNMENT, HAVE
28	1897	McKinley	1186	3972	0.72	0.52	0.008	HAS, PEOPLE, GOVERNMENT, CONGRESS, BEEN, GREAT, HAVE, COUNTRY, MORE, SUCH, WAS, PUBLIC

29	1901	McKinley	809	2215	0.72	0.59	0.005	HAS, GOVERNMENT, PEOPLE, HAVE
30	1905	Roosevelt	383	991	0.70	0.49	0.012	HAVE
31	1909	Taft	1372	5438	0.69	0.51	0.006	HAS, HAVE, GOVERNMENT, BUSINESS, SUCH, PROPER, LAW, CONGRESS, BEEN, OTHER, TARIFF, RACE
32	1913	Wilson	626	1712	0.69	0.61	0.023	HAVE, GREAT, BEEN, HAS, MEN, GOVERNMENT, HAD, JUSTICE, LIFE
33	1917	Wilson	523	1531	0.68	0.49	0.023	HAVE, OWN, MORE, BEEN, SHALL
34	1921	Harding	1117	3346	0.73	0.53	0.012	WORLD, JAVE, AMERICA, WAR, HAS, NEW, CIVILIZATION, GOVERNMENT
35	1925	Coolidge	1158	4056	0.70	0.52	0.010	HAVE, HAS, COUNTRY, GREAT, WHO, BEEN, PEOPLE, GOVERNMENT, MORE, WHAT, DO
36	1929	Hoover	1022	3766	0.67	0.45	0.009	GOVERNMENT, HAVE, MORE, PEOPLE, PROGRESS, PEACE, WORLD, JEUSTICE
37	1933	Roosevelt	709	1883	0.71	0.47	0.006	HAVE, NATIONAL
38	1937	Roosevelt	684	1823	0.72	0.57	0.020	HAVE, GOVERNMENT, PEOPLE, BEEN, NATION
39	1941	Roosevelt	490	1346	0.67	0.62	0.009	NATION, HAS, DEMOCRACY, HAVE, LIFE, SPIRIT
40	1945	Roosevelt	259	559	0.68	0.63	0.011	SHALL, PEACE, HAVE
41	1949	Truman	739	2283	0.70	0.50	0.022	WORLD, HAVE, NATIONS, PEACE, FREEDOM, PEOPLE, FREE, UNITED, MORE, PEOPLES, SECURITY, DEMOCRACY
42	1953	Eisenhower	845	2461	0.70	0.46	0.012	FREE, WORLD, PEACE, SHALL, HAVE, PEOPLE, STRENGTH, FREEDOM
43	1957	Eisenhower	585	1660	0.70	0.42	0.015	WORLD, NATIONS, FREEDOM, PEOPLE, PEACE, SEEK, OWN
44	1961	Kennedy	531	1365	0.70	0.48	0.012	LET, DO, WORLD, SIDES
45	1965	Johnson	524	1493	0.67	0.63	0.020	HAVE, NATION, CHANGE, MAN, UNION, WHO, PEOPLE
46	1969	Nixon	704	2131	0.69	0.57	0.011	HAVE, PEOPLE, WORLD, PEACE, WHAT, LET, WHO
47	1973	Nixon	505	1818	0.65	0.52	0.048	LET, AMERICA, PEACE, WORLD, HAVE, NEW, DO, HAS, RESPONSIBILITY, MORE, NATION
48	1977	Carter	491	1226	0.72	0.44	0.016	NATION, NEW, HAVE, HAD
49	1981	Reagan	841	2446	0.72	0.60	0.014	HAVE, GOVERNMENT, DO, WHO, HAS, BEEN, BELIEVE, AMERICANS, WORLD, PEOPLE
50	1985	Reagan	855	2575	0.73	0.59	0.018	HAVE, GOVERNMENT, PEOPLE, WORLD, FREEDOM, WHO, HAS
51	1989	Bush	743	2335	0.70	0.56	0.013	HAVE, NEW, WHAT, WHO, NATION, WORLD, GREAT
52	1993	Clinton	596	1611	0.71	0.54	0.034	WORLD, AMERICA, HAVE, PEOPLE, TODAY, WHO

53	1997	Clinton	717	2171	0.70	0.43	0.029	NEW, CENTURY, WORLD, AMERICA, NATION, HAVE, TIME, PEOPLE, LAND, GOVERNMENT, PROMISE
54	2001	Bush	583	1593	0.71	0.53	0.010	AMERICA, NATION, STORY, COUNTRY, CITIZENS, DO
55	2005	Bush	720	2078	0.70	0.57	0.024	FREEDOM, HAVE, AMERICA, LIBERTY, NATION, OWN
56	2009	Obama	886	2407	0.73	0.59	0.010	HAVE, HAS, WHO, NATION, NEW, AMERICA
57	2013	Obama	772	2120	0.72	0.58	0.003	PEOPLE, TIME