New Year’s Day Speeches of Czech Presidents: Phonetic Analysis and Text Analysis
Milan Jičínský, Jaroslav Marek

To cite this version:
Milan Jičínský, Jaroslav Marek. New Year’s Day Speeches of Czech Presidents: Phonetic Analysis and Text Analysis. 16th IFIP International Conference on Computer Information Systems and Industrial Management (CISIM), Jun 2017, Bialystok, Poland. pp.110-121, 10.1007/978-3-319-59105-6_10. hal-01656221

HAL Id: hal-01656221
https://hal.inria.fr/hal-01656221
Submitted on 5 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
New Year’s Day speeches of Czech presidents: phonetic analysis and text analysis

Milan Jičínský, Jaroslav Marek
Univerzita Pardubice, Fakulta elektrotechniky a informatiky
Studentská 95, 530 02 Pardubice I, Czech Republic
milan.jicinsky@student.upce.cz, Jaroslav.Marek@upce.cz

Abstract. The aim of our study is verification of programmed algorithms of phonetic analysis using concrete data, and reassurance that it works as also sought after. For our testing, the appropriate recordings of New Year’s Day speeches of Czech and Czechoslovak presidents are available. The very first available recording of presidential speech comes from 1935. All transcripts and recordings of the last 87 speeches are located on the web page www.rozhlas.cz. The primary goal of this paper is to analyze voice characteristics of the speaker (log energy, speech velocity and Zero crossing rate). Especially words "with greatest energy" will be found. There will be a list of words having the highest energy values. The most interesting results will be presented by graphical tools. Using a software, capable of text analysis, transcript characteristics such as most frequent words, length of words, total number of words and different words will be computed. The most frequent words will be presented. Political speeches often become the subject of various analyses. Our calculation allows a new perspective on speeches. It is interesting to compare the most frequent semantic words and words with the greatest energy. The results can be historically important. It allows an extraction of new information from available data and scientifically different approach.

Keywords: New Year’s Day speeches of Czech presidents, intensity of speech, vocabulary richness, speech velocity, Zero crossing rate, length of words

1 Introduction

Nowadays analysis of speech is very popular. It started in the second half of 20th century when basic signal characteristics were discovered. Fundamentals are well described in [1-3]. This field of study evolved very quickly so there are many applications from key word detection [4] across transcription of fluent speech [5] to recognition of speaker [6]. This article is reserved for those who want to study phonetic analysis and its fundamentals. Even historians whose field of study is 20th century in the Czech Republic can appreciate the most frequented words and words with the highest energy. Linguists can be satisfied with changes in individual speeches whether it's written or spoken.

Many authors aim their work on linguistic analysis of political speeches. For example, articles End-of-year speeches of Italian presidents or inaugural speeches of US
presidents were researched in [7, 8]. Relationship between ideology and language and thematic concentration of Czechoslovak New Year’s Day speeches is analyzed in article [9]. Of course, it is very interesting to study influences of ideology, originality of author and his abilities to differ from uniformity. The most frequent words can provide an information about recent years because they react on the most important events. Some of those words will be listed.

Main goal of this publication is to present words that were said with the greatest energy. Words with the greatest energy allows to track what president emphasized on during a reading. Emphasis of the speaker will be probably on positive words. The only exception could be the time of war. Ideological words may be emphasized in some speeches. There will be one more characteristic calculated for each speech – speech velocity.

2 Voice characteristics

Analysis of recordings of New Year’s Day speeches will be introduced in this chapter. The intensity of voice (energy) and speech velocity will represent voice characteristic of speaker. Energy tells something about how much emphasis speaker uses and speech velocity shows how fast speaker speaks. These variables can be influenced e.g. by age or by sickness. Then the words having the greatest energy can be found. It could be interesting to compare these words with most frequent thematic words. President didn’t have to be an author of written text. But he could highlight any words he wanted to. It depended on what he considered to be important. This is the example of individuality. Then ZCR (zero crossing rate) characteristics will be shown.

Figure 1: “Dear fellow citizens” – Václav Havel (1998). Source: own.

2.1 Obtaining data and its processing

Source data have been obtained from website www.rozhlas.cz. Speeches are recorded with useful software Audacity. Sampling rate of each speech is 8 kHz. Each recording is modified because the original ones contain a music before the speech starts. Calculations of voice parameters are realized in MATLAB. Scheme of processing can be simplified as on Figure 3.
Segmentation means that the record is divided into frames of the same length (typically 20 ms long). Frames are overlapping each other right in the half (in this case). Overlapping is recommended to the fact that parameters can be changed in jumps. So this enhancement improves the dramatic changes and it can describe even changes near an edge of frame without loss of useful data. After the segmentation follows a parameterization. Feature vector values of each frame is computed during the parameterization. Features can be divided into: basic, spectral, cepstral and dynamic. ZCR and energy ranks among the basic features. Feature extraction and segmentation is also discussed in [2].

![Audio processing diagram](image)

Figure 3: Audio processing. Source: own.

2.2 Intensity of voice

The intensity of voice is characterized by energy. So the energy is a key parameter which defines the intensity of voice. Energy is defined as the sum of squared values of samples within one frame. Logarithm function is used for better range of energy values. In this case Log energy of ordinary noise is around 5. Whenever speech is contained in recording, values of energy are greater for those frames. Typically, the energy of speaking person can reach even value of 15. It depends on how loud speaker speaks. Log energy is defined as

\[
E = \log \left( \sum_{n=0}^{L-1} x^2(n) \right),
\]

where L is the frame length, concretely the number of samples contained in the frame. \( x(n) \) is the designation for the current sample value.

In comparison of all presidents it’s evident to see that president Hácha spoke not as loud as others. He had no emphasis. This could be caused by political situation. Hácha used to be a president during the hardest time of the Czechoslovak history. He was helpless president of protectorate state. The only thing he could do was to make peo-
ple feel calm and safe, even if it wasn't possible. As for president Husák, very significant decrease of energy was observed between years 1978 and 1979.

Figure 4: Log energy: mean value. Source: own.

2.3 ZCR

Zero crossing rate is a parameter that characterizes changing of sign from negative to positive or back. Zero crossing rate is related to the frequency. There is one value of ZCR for each frame (the same as for the energy). The principle of ZCR can be easily explained with Figure 5. ZCR value is equal to the count of all dots. The dots are placed to the points where signal intersects x axis and changes the sign.

Figure 5: Explanation of ZCR - 20ms frame containing of phoneme "a". Source: own.
It’s often used for voice activity detection [11] – to find out if human speech in record is present or not. As for voiced signal ZCR values are typically low. Noises and unvoiced signals have higher values. This method is sensitive to noises and direct component shifts. It even allows us to find out if concrete phoneme is voiced \((b,d,g,z,v,h,\ldots)\) or unvoiced \((p,t,k,s,f,ch,c,\ldots)\). Especially the sibilance \((s,c,š,č,\ldots)\) have higher ZCR values.

\[
ZCR = \frac{1}{2} \sum_{n=1}^{L-1} \left| \text{sgn}(x(n)) - \text{sgn}(x(n-1)) \right|.
\]  

\(2\)

Data variability is relatively high. So, the mean value of ZCR is not that good to represent individual speaker. Better results can be obtained using ZCR dynamically. That means ZCR of each frame is used. Then search for dynamic changes instead of treating it as one static value. It’s preferable to use it for each frame.

### 2.4 Speech velocity

For the purposes of the article there is a created parameter that can be used to link results of text and voice interpretation into one value that characterizes the speaker. It’s called speech velocity. This mean value represents how many words the speaker pronounces during the time of one second. The speech of president Husák from 1989 is significantly the slowest. President Hacha is speaking relatively slowly too. On the
other hand, the fastest tempo of speaking can be recognized in speeches from 1938, 1943 (Beneš), 1959 (Novotný) and 1996 (Havel).

Figure 7: Speech velocity: mean value. Source: own.

3 Characteristics of written text

All studies are realized for 87 speeches of Czechoslovak presidents, Czech presidents or Czechoslovak prime ministers. The unique situation happened due to the World War II. The Czechoslovak Republic had two presidents. President Beneš left his country and exiled to the Great Britain. But he was still very politically active. Then Hácha was chosen to be a president of protectorate. So, both groups of speeches were analyzed between 1940 and 1945. On the other hand, president in exile is considered to be more important subject of our analysis.

We can expect changes in using of different length words during a long history of New Year’s Day speeches. Therefore, the first aim of our calculation is to determine average of word length. Calculations of text parameters are made using software based on Java [10] called „Statistika v lexikální analýze“. This GUI (Graphical User Interface) has been created during diploma thesis. It makes easier the whole text processing. The software allows to analyze frequent letters and words, length of words, aggregation and alliteration and some other features. The original purpose of existence of software is analysis of poems and its translations as in [12].

3.1 Mean of word length

On the Figure 8 can be seen mean values of length of words of analyzed texts. Length of words of communistic presidents (Zápotocký, Novotný, Husák) is much
greater than nowadays (Havel, Klaus, Zeman). President Beneš used the shortest words. The greatest variability can be seen at Svoboda and Hácha. Estimate of expected value is given by

$$
\bar{x} = \frac{1}{n} \sum_{i=1}^{k} if_i.
$$

(3)

where $i = 1, ..., k$ is length of word, $k$ is length of the longest word.

3.2 The most frequent words

Conjunctions and prepositions of course ranks among the most frequent words. Conjunction “a” is the most frequent in all speeches except Novotný (1964) – “v”, Svoboda (1973) – “v” and Hácha (1944) – “se”. Figure 10 is the list of sorted conjunctions and prepositions used from first to fourth position. The common words can be seen. The most frequent words with meaning will be presented in Figure 10 too. These words differ much more than prepositions. Presidents react on current political events such as crisis, protectorate, war or return of democracy. Meaning words can provide a quick preview of content. Comparison with inaugural speeches of US presidents can be interesting. As for words with meaning, for example Roosevelt (1933) said words: HAVE, NATIONAL and Truman (1949) used the words: WORD, HAVE, NATIONS, PEACE, FREEDOM, PEOPLE, FREE, UNITED, MORE, SECURITY, DEMOCRACY [7]. See Figure 9 or 10.
Figure 9: The most frequent words (Czech version). Source: own.
Figure 10: The most frequent words and words with the highest energy. Source: own.
Figures are divided into some subsections. As mentioned before, the Czechoslovak Republic had two presidents between 1940 and 1945. In 1993 the second anomaly appeared. The Czechoslovak republic ceased to exist. Since 1993 the country was divided into two smaller autonomous countries: The Czech Republic and the Slovak Republic. So, the president Václav Havel had no speech in 1993. Prime ministers were speaking to their nations instead of president.

Rows are sorted by years. Each row has its color depending on president or prime minister. Colors were chosen according to all figures. The first column contains first four most frequent conjunctions and prepositions. Then there are three columns containing the most frequent words sorted by order. The last column shows the word with the highest energy. Those words are written by uppercase.

### 3.3 Number of words

Scatter chart will be used to demonstrate a vocabulary richness. Coordinates on axis x means total number of words in speech and coordinates on axis y means total number of different words. Functional dependency can be modeled by Gompertz curve. Presidents with values above the curve have greater ratio of words than other presidents. Language richness of speeches under the curve can be considered lesser. In article [9] author mentioned that thematic concentration of president Havel is surprisingly low. But this claim doesn't seem to correspond with language richness. According to the Figure 10, ratio of number of different words and total number of words is greater as for Havel. This could be caused by choosing different methods of evaluating language richness.

![Figure 11: Number of different words. Source: own.](image-url)
4 Conclusions

This article’s goal is to present results of our research and show that data we already had can be processed in different way. The extraction of information is much discussed nowadays. Main purpose of research is finding the words with the greatest energy. Because they have historical importance, they can be used as keywords and they even characterize the speaker.

Scale of publication doesn’t allow to detail comment and the description of used algorithms. Many hours of machine time have been needed during the calculations of phonetic parameters of speeches. Archive [13] contains 74 speeches. So, this is more than 19 hours of recordings to be analyzed. Before calculation of mean values of ZCR and Log energy, there was an extensive table for each speech. Presented parameters were created by reducing the table containing millions of values (each frame parameter values) into one mean value. Unreduced data may be used for further analysis.

Comparing the table of most frequent thematic words with table containing the words pronounced with the greatest energy brings almost no match. The speaker didn’t emphasize the most frequent words. But he chose to highlight other words. For example, Masaryk talked about economy. Beneš emphasized the war and human kind. Novotný insisted on hard work and improving the communistic country. Havel emphasized the very first words: “Dear fellow citizens.” It can provide some information without listening to the whole speech. It even characterizes the president himself and an era of each president (the most important events, standard of living, relationship between president and citizens).

5 References

1. LIANG, Bai, et al. Feature analysis and extraction for audio automatic classification. In: Systems, Man and Cybernetics, 2005 IEEE International Conference on. IEEE, 2005. pp. 767-772.
2. LIU, Zhu; WANG, Yao; CHEN, Tsuhan. Audio feature extraction and analysis for scene segmentation and classification. The Journal of VLSI Signal Processing, 1998, 20.1: 61-79.
3. Short Term Time Domain Processing of Speech (Theory): Speech Signal Processing Laboratory. IIT GUWAHATI Virtual Lab [online]. Guwahati: Bhattacharjee, 2016. URL: http://iitg.vlab.co.in/?sub=59&brch=164&sim=857&cnt=1. [accessed 2016-12-05]
4. KANDA, Naoyuki, et al. Open-vocabulary keyword detection from super-large scale speech database. In: Multimedia Signal Processing, 2008 IEEE 10th Workshop on. IEEE, 2008. pp. 939-944.
5. NOUZA, Jan, et al. Very large vocabulary speech recognition system for automatic transcription of Czech broadcast programs. In: INTERSPEECH, 2004.
6. WASSON, D.; DONALDSON, R. Speech amplitude and zero crossings for automated identification of human speakers. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1975, 23.4: 390-392.
7. KUBÁT, M., ČECH, R. Quantitative Analysis of US Presidential Inaugural Addresses. Glottometrics (34) 2016, pp. 14-27.
8. TUZZI, Arjuna; POPESCU, Ioan-Iovitz; ALTMANN, Gabriel. *Quantitative analysis of Italian texts*. Lüdenscheid: RAM-Verlag, 2010.

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

10. ŠLAHORA, Jan. *Matematická lingvistika a překlady básně E.A. Poea Havran* [online]. Pardubice, 2015. Master thesis. University Pardubice, Faculty of electrical engineering and informatics. URL: https://dk.upce.cz/bitstream/handle/10195/60403/SlahoraJ_MatematickaLingvistika_JM_2015.zip?sequence=1&isAllowed=y. [accessed. 2016-12-05].

11. R. Bachu, S. Kopparthi, B. Adapa, and B. Barkana, “Separation of voiced and unvoiced using zero crossing rate and energy of the speech signal” in *American Society for Engineering Education (ASEE) Zone Conference Proceedings*, 2008, pp. 1–7.

12. MAREK, J. - ŠLAHORA, J. Měření podobnosti překladů básně Havran. *Forum Statisticum Slovacia*, 2013, vol. X (2013), no. 5, pp. 90-95. ISSN: 1336-7420

13. URL: http://www.rozhlas.cz/zpravy/data/_zprava/od-ign-k-zemanovi-poslechnete-sivanoce-a-novoroce-projevy-vsech-presidentu--1436738. [accessed. 2016-12-17].