

OTDK-dolgozat

2021

BUDAPESTI CORVINUS EGYETEM

A POLITIKAI KAMPÁNYOK LEGNAGYOBB FEGYVERE

**THE GREATEST WEAPON OF POLITICAL CAMPAIGNS
SENTIMENT AND EMOTION ANALYSIS OF US PRESIDENTIAL
CAMPAIGN SPEECHES**

Kamenár Gyöngyvér

Konzulens: Pál Gábor, Stump Árpád

Kézirat lezárásának dátuma: 2022. 04. 19.

Table of Contents

1. Introduction	1
2. Literature review	2
2.1. The role of emotions in political campaigns	2
2.2. Sentiment analysis of political content	3
3. Sentiment analysis of US Presidential speeches	8
3.1. Introduction	8
3.2. Data.....	8
3.3. Methodology.....	9
3.4. Results and discussion	10
4. Emotion analysis and thematic words	12
4.1. Emotion analysis.....	12
4.2. Comparison of candidates	15
4.3. Thematic word frequency	18
5. Discussion.....	20
6. Conclusion.....	22
Bibliography.....	24

List of Figures

Figure 1: Distribution of sentiments.....	10
Figure 2: Sentiments by candidate	11
Figure 3: Average number of emotional words per paragraph.....	13
Figure 4: Total number of emotional words.....	14
Figure 5: Boxplot of emotional words.....	15
Figure 6: Average number of emotional words by candidate	16
Figure 7: Boxplot of emotional words by candidate	17
Figure 8: Total sum of emotional words by candidate	18
Figure 9: Most frequent unique words	19

1. Introduction

Do you think that politicians affect your emotions? Surely, they do anything in order to gain our trust and more importantly, our vote. But what are the weapons they use to reach their goals? By weapons, we mean psychological and rhetorical tricks and strategies to convince their audience about the perfect program and their excellent leadership.

It is clear, that politics is much more than parties and programs, people always trust in their instincts, and consciously or unconsciously they listen to their emotions. The psychological impact of political campaigns becoming more important than ever and one of the main reasons behind this is that the significance of material conflicts is fading compared to the value and virtue-based disagreements (Enyedi, 2004). Another potential reason is that with the rise of mass and social media, decision-makers have much more opportunities to feed us with information and they do not give just dry facts but they provide a huge emotional package with it.

The theory of rational choice suggests that citizens vote rationally based on every piece of information but common knowledge and empirical research cannot confirm that (Kiss, 2013). Citizens are able to pay attention to every political information and news, but in most cases, they do not want to waste time and energy on unnecessary things so they rather trust their instincts and emotions. There is a belief (Edelman, (1988) as cited in (Enyedi, (2004)) that for an average citizen politics is purely based on emotions and myth. The symbols, decisions, and information can be interpreted in so many ways that people rather decide based on their emotions without any real political context. Campaigns can potentially reinforce such emotional dependency of citizens, however, we do not know, what kind of sentiments and emotions campaigners use to convince their target groups.

This study aims to investigate what kind of sentiments and emotions politicians use or express nowadays as their campaign tools to win the election. We hypothesize that negativity prevails the presidential campaign speeches based on Soroka's theory of media negativity. The second hypothesis is that Donald Trump used a less negative and more positive sentiment in his campaign speeches because of his incumbent position compared to Joe Biden. Besides, our third research question is what kind of emotions dominated the presidential candidates' rhetoric during the 2020 US campaign.

In detail, we analyzed all the campaign speeches of Joe Biden and Donald J. Trump leading to the 2020 US presidential election. We performed sentiment analysis with Amazon Comprehend and confirmed that negativity reigns in the campaign speeches. However, Trump

used more negative sentiment contrary to our hypotheses but he also used more positive sentiment than Biden, who prevails over Trump in neutral rhetoric. In addition, we executed a lexicon-based emotion analysis and determined that trust and anticipation are the dominating emotions in the election campaign speeches. Apart from that, we also reported the frequency of unique words of each candidate to get some impression about the language and tone they use which can help us understand the previous results.

The 2nd section reviews the relevant literature both from the aspect of emotions in political campaigns and the use of sentiment analysis for political content. The 3rd section describes the sentiment analysis of presidential speeches while the 4th section presents the emotion analysis and unique word frequencies. The 5th section discusses the results and implications of the paper and finally, section 6 concludes the research.

2. Literature review

2.1. The role of emotions in political campaigns

The role of emotions in politics might not be straightforward at first glance. However, as Marcus (2000) concludes, it has been the focus of every political thinker from Aristotle, Platon, Descartes to Smith and Hume (as cited in Marcus, 2000) and many others, all of them thought that in order to engage in politics and scrutinize human nature it is crucial to understand human emotions.

Despite that, emotion was a declining factor in politics and science for a period because of the view that democratic politics and progress demand more reason and less emotion (Marcus (2000) cites Arkes (1993)) and that extensive public education joined with the growth of scientific knowledge empower reason to repel emotions and take up the central role in political science and psychology as well (Brader & Marcus (2013) cites Marcus (2010)). But in the 1980s emotion started moving forward to the center stage in both disciplines (Brader & Marcus, 2013).

Generally, the use of emotions in politics can be categorized into two approaches (Marcus, 2000). Firstly, the particular characteristics of a person involve emotions as a feature of personality. An application of this approach is the case studies of political leaders because attention to them focuses on their characteristic emotional inclination. The other approach is how humans experience various emotional reactions to situations. Here, the target shifts from personality to the emotion attached to situations, symbols, groups, individuals, or external

events to evoke response from the audience. In this case, emotions are used to explain why individuals deviated from their nature due to some stimulus.

Our study rather fits into the first case, because we analyzed the speeches of two political leaders, which reflect their characteristics and the emotions involved or evoked by them. Besides, our research focuses on campaign speeches which is a different, much narrower, and more concrete topic than emotion in politics. However, there is limited research on examining the campaigns themselves.

Kühne et al. (2011) studied the priming effect of campaigning on opinion formatting in the case of Swiss tax reform. Priming is the phenomenon whereby the previous experience of a certain stimuli influences the response to the subsequent stimuli (American Psychological Association, n.d.). They claim the campaign-induced emotions can activate cognitive content which in turn can affect political judgment. Based on their analysis, political campaigns can influence public opinion not only by cognitive priming but also as a consequence of affective priming.

Another empirical study (Weber, 2013) examined the political effects of emotional advertisements during the 2004 presidential campaign. Weber used 4 types of emotional ads: fear, anger, hope, and sadness. The experiments showed that anger is uniquely mobilizing by leading to the belief that our behavior is politically consequential. However, fear and sadness had some lessening effects on political engagement. The study confirms that emotions are important parts of campaigning.

Brader (2005) performed an experiment during the governor's campaign to measure the impact of emotion on political behavior on nearly 300 subjects. He manipulated ads by emotional appeals, namely uplifting music and colorful, warm images of children with positive script to detect the appeal of enthusiasm, while grainy, black-and-white images of violence with discordant, tense music added to negative script to examine ads with fear. He found that cues of enthusiasm stimulated existing loyalty and induced participation while fear cues activated vigilance, encouraged persuasion, and increased confidence in existing evaluations.

These studies confirm that emotions are not just important in politics but campaigns can be intentionally designed by expressing emotions to influence political behavior and thinking. Needless to say, it is essential to further study emotional expressions in political campaigns.

2.2. Sentiment analysis of political content

After reviewing the role of emotions in politics let's study the previous literature on sentiment analysis in politics. The past of sentiment analysis is much shorter than emotions in politics

because it is a rather new discipline. Sentiment analysis evolved with the fast-paced development of information technology and artificial intelligence in the 1990s and was performed by computational linguistics (Mäntylä et al., 2018). However, the uprising of computer-based sentiment analysis was due to the availability of text on the web, therefore 99% of the studies have been published after 2004. In the last decade, the object of sentiment analysis shifted from online product reviews to Twitter, Facebook, and other social media text therefore, today's sentiment analysis techniques are aimed to capture any features of textual data like emoticons, capitalization, punctuation, and so on. Nevertheless, it is perfectly suitable for analyzing political content on any platform.

2.2.1. Sentiment analysis targeting audience

A recently popular use case of sentiment analysis is the prediction of election outcomes. Chauhan et al. (2021) reviewed 38 papers published after 2004 which all aimed to predict election results using social media data. 31 out of them used sentiment analysis on their own or as part of the prediction methodologies. 4 papers used a purely volumetric approach while 2 studies applied social network analysis and 1 paper worked with the combination of these two. Within the 31 studies that employed sentiment analysis, 16 used the lexicon-based approach, and 12 used machine learning while 3 papers performed their analysis with the combination of these two. However, the outcome is the most interesting part for us. 24 studies had positive prediction results while 5 had negative. 3 papers had negative as well as positive predictions considering different criteria and the other 6 studies did not predict anything but accomplished some other tasks. Anyhow, we can see sentiment analysis is becoming a reliable tool to predict election results based on social media data. In the future, it might become an essential part of political surveys.

González-Bailón and Paltoglou (2015) used sentiment analysis to measure public opinion of online debates. The specialty of their study is that they used several different platforms (news, discussion forum of BBC, Digg, Youtube, Twitter, and politically oriented blogs) and several different methodologies (5 different dictionary-based, human annotation, and machine learning). They concluded that reliability and validity really depend on the diversity and formality of the text, and machine-learning outperforms the dictionary-based solutions.

Another analysis of public opinion (Iacus et al., 2012) uses Hopkins & King's (2010) method of automated nonparametric content analysis for social science and the researchers showed a remarkable correlation between social media and mass survey results, as well as between public

opinion of social media and electoral results despite that internet users and not compulsorily representative of the population.

Using sentiment analysis targeting the audience to predict or explain the 2020 US election was also done by plenty of researchers lately (Xia et al. 2021, Chandra & Saini 2021, Bernábe-Loranca et al. 2022, Chaudhry et al. 2021, Akbar et al. 2021, Chandak 2022, Singh et al. 2021) mainly using social media data. However, the sentiments expressed by the campaign and the political actors themselves have not got so much attention.

2.2.2. Sentiment analysis targeting political actors

We can see, that sentiment analysis of political content usually targets the audience, citizens, or voters because the goal of most studies is to predict election results, and determine public opinion or the level of support for proposals. A limited number of researches focus on the political actors themselves to analyze their campaigning strategy from sentiment and emotional aspects. Nevertheless, sentiment analysis can be a great tool for political psychological research to inspect the strategy and methods of leading politicians and political parties. In the following part, we introduce studies that applied sentiment analysis targeting political actors.

Dilai et al. (2018) analyzed the US and Ukrainian presidential speeches, namely the speeches of Donald Trump and Petro Poroshenko with 2 programs SentiStrength¹ and UAM CorpusTool². They found that Trump's speeches are highly emotional (only 27% of the segment are neutral) but interestingly positive sentiments prevail. Contrary, in the case of Poroshenko, the speeches are mostly neutral (55%) but positivity still prevails over the negative polarity. However, Dilai's paper just aims to test the political domain and the Ukrainian language of the programs and lacks any theoretical or political psychological discussion.

Similar research to this present study was carried out by Hoffmann (2018) who analyzed the campaign speeches of Hillary Clinton and Donald Trump during the 2016 US Presidential election. Hoffmann argues that presidential speeches are much more than sentiments because both sentiment categories include many different emotions. So besides the polarized sentiment analysis, he applied emotion analysis with the *Syuzhet* R package (Jockers, 2020) which distinguished 8 distinct emotions: joy, surprise, trust, disgust, sadness, fear, anger, and anticipation. The study found that Trump used considerably more (10%) negative words than Hillary which is statistically significant however, still positive sentiments prevail in Trump's

¹ SentiStrength, <http://sentistrength.wlv.ac.uk/>

² UAM CorpusTool <http://www.corpustool.com/>

speeches just as concluded by Dilai et al. (2018). Besides, by adopting the classification of emotions, both of the candidates dominantly expressed trust (26% and 23%). However, trust is followed by anticipation and joy in the case of Hillary while fear and anticipation come next in Trump's speeches. Hoffman argues, that is why Donald Trump is described as someone who mostly relied on negative rhetoric (Bayagich et al., 2017).

A comprehensive and punctilious study was done by Liu and Lei (2018) who performed the sentiment, discourse, and thematic investigation of Clinton's and Trump's speeches during the 2016 election. The sentiment analysis concludes similarly to Hoffmann (2018) more negative sentiment in Trump's speeches. Moreover, their unique thematic analysis revealed much more about the two parties. The top 10 thematic words of Clinton were *family, pay, job, talk, hard, vote, woman, day, support* and *child*, while from Trump's side the 10 most frequent thematic words are *job, world, deal, happen, win, tax, million, trade, government* and *vote* in the listed order. Besides, both candidates had unshared, unique prominent phrases. Clinton's prominent terms are more family/education/health-related while Trump emphasizes more trade/money and law-enforcement related terms which demonstrate his business-oriented approach contrary to Clinton's focus on achieving social and economic equality for underprivileged and minorities. This methodology enabled the researchers to understand in-depth the politicians' opinions, actual attitudes, and perspectives on issues.

A more recent paper by Crabtree et al. (2019) analyzed the campaign sentiment of political parties in 8 European countries. More precisely, they analyzed the emotive language in over 400 party manifestos and stated several hypotheses regarding incumbency status, policy position, and the objective economic condition of the country. They found that incumbent parties and especially incumbent prime ministerial parties have a higher level of positive sentiment in their language compared to the opposition parties. The researchers also discovered, that ideologically moderate parties use more positive sentiment than extremist parties. Besides, they found that parties express less positive sentiment when the economy has poor performance, but this negative impact is weaker for incumbent parties.

The study of Crabtree et al. (2019) is not just important because of its findings and implications for campaigning strategies but they also provide a comprehensive theoretical model and explanation behind their hypothesis which they have confirmed empirically as well. The basis of the theory, is that voters are influenced not just by the substantive content of political campaigns but by emotive content as well. Changing the emotional nature of the language describing political content influences how it is evaluated even after controlling for the objective content and the focus of the political message (Crabtree et al. (2019) cites Utych

(2018)). Utych (2018) also concludes, that when candidates are described using negative language, they are evaluated more negatively compared to the case when they are described with neutral emotive language. Thus, individuals process information differently based on their emotional mood as concluded also by Schwarz (2000) cited in Crabtree et al., (2019). Therefore, parties have the incentive to emotionally frame their political messages based on how they want the voters perceive the world.

Crabtree et al. (2019) describe that this incentive can be tied to the models of retrospective voting. Voters reward and promote the incumbent party when they view the world as good, because the incumbent's performance in office determines the world, at least partially. On the other hand, they want to punish the incumbent when they perceive the world as poor, or bad. The good or bad state of the world is frequently understood in terms of economy, but it is not necessarily constrained to that. This model implies that incumbent parties have to frame their political content with positive emotions to influence the voters' perception of the world to be great and collect their votes. The case is the opposite for opposition parties, they have to use negative emotional framing during their campaign message to make voters see the actual status of the world to be poor and therefore, vote for the opposition's program which can change that.

Moreover, the researchers propose that citizens are better at assessing the campaign messages' overall sentiment than the detailed, substantive situation or position that are shown in these political contents. Based on this claim, we can ask the question, how important is the campaign sentiment compared to the objective political position. There is no clear answer for that, however, based on the previously described works of literature we can agree on campaign sentiment has a relevant influence on the election, therefore it should not be neglected. Not just because it can influence voters and decide elections, but also because it requires no additional cost in a campaign budget just some consciousness and a clear, well-designed strategy.

The theory has straightforward implications for political actors, but do they use it effectively? Crabtree et al. (2019) confirmed it in the case of 8 European countries and some researchers, Hoffmann, 2018, Liu & Lei, 2018) analyzed the sentiment of US presidential candidates up to the election of 2016 but the sentiment and characteristics of Donald Trump and Joe Biden's campaign before the 2020 US election have not been covered yet. Especially not extended with a theoretical framework, explanation, implications, and recommendations. This research is aimed to fill this gap.

3. Sentiment analysis of US Presidential speeches

3.1. Introduction

Clearly, sentiment analysis is a great tool to analyze large-scale political content including the political actors themselves. There are plenty of studies about election prediction and also some about political speeches, but the latest US presidential election has not been covered yet.

We have analyzed the 2020 presidential campaign speeches of Joe Biden and Donald Trump and our decision behind the 2 people is twofold. Firstly, they are the top politicians of one of the most developed countries in the world with huge attention around them. Therefore, their campaign speeches represent the most sophisticated and well-built political speeches currently, which are worth analyzing. However, in this aspect, we do not want to compare their rhetoric or make a distinction between them because their speeches are both parts of the campaign speech sample.

Secondly, the election of 2020 was a close fight between Donald Trump and Joe Biden (CNN, 2020) and they have different personalities, values, virtues, and programs and most likely personalities too. They both represent political campaigns and strategy as well, so it is interesting to compare and see the contrast between their political rhetoric. All in all, the speeches of the two candidates open the door for substantial analysis.

3.2. Data

The transcripts of presidential speeches were reached from the 2020 Election Transcript section³ of www.rev.com. This collection includes all election and campaign-related events' transcripts, including presidential debates, campaign rallies, and political speeches not just from the candidate but also from their official supporters and colleagues. We retrieved the text content of all transcripts using a data mining method called web scraping. The event date is filtered to be no later than the 3rd of November, 2020 because that was the election day itself and we wanted to avoid discussions about election results, fraud claims and everything else which are not part of the campaign so cannot influence voters.

The text content of each transcript was divided into paragraphs with a timestamp and the speaker's name. This feature enables us to filter the speeches of Donald Trump and Joe Biden. The mentioned paragraphs then served as the unit of analysis for 3 reasons.

³ 2020 Election Transcripts <https://www.rev.com/blog/transcript-category/2020-election-transcripts>

Firstly, that paragraphs usually contain one thread of thoughts which is inevitable for understanding the context. Sentences are also connected within paragraphs. Secondly, text analysis methods perform the best with the appropriate length e.g. word by word is not enough because there is no context around it but analyzing a whole chapter is also useless because we will get a mixture of all emotions which often offset themselves so it will have neutral or indecisive sentiment value. The third reason is just technically complementing the former two, the used cloud computing platform allows a limited number of requests per second so dividing the text into smaller parts would be just a waste of time and energy.

Altogether, we analyzed 189 speech transcripts which contain 12271 paragraphs of text, from which 8682 have been told by Donald Trump and the remaining 3588 by Joe Biden. The 12270 paragraphs actually mean more than a 6.7-million-character length of text, which is quite appropriate for the research. Noticeably, Trump has approximately twice as much speech as Biden but it won't harm our analysis, because even Biden has still a representative amount of text. In the case of comparison between them, we can use relative terms instead of absolute measures.

3.3. Methodology

As mentioned before, sentiment analysis can be performed by several different methods. One popular and scalable method is the lexicon/dictionary-based analysis which order scores to word and phrases supplemented by some simple algorithms like negation, end-of-sentence signs, and so on. However, a more reliable and developed method is using machine-learning-based models which is a type of artificial intelligence. Simple machine learning models can be built by anyone but there are already several services that offer us their highly developed resources.

In this research, we used a machine learning tool of Amazon Web Services (AWS) called Amazon Comprehend⁴ which offers one of the best or the best sentiment analysis tools in the world currently. The other perk of AWS is that it is a cloud computing platform so the actual computational resources do not burden our computer, we just provide the request and get the results. Besides, the service is easily implementable into most programming languages like R (R Core Team, 2021) where we have worked with the data.

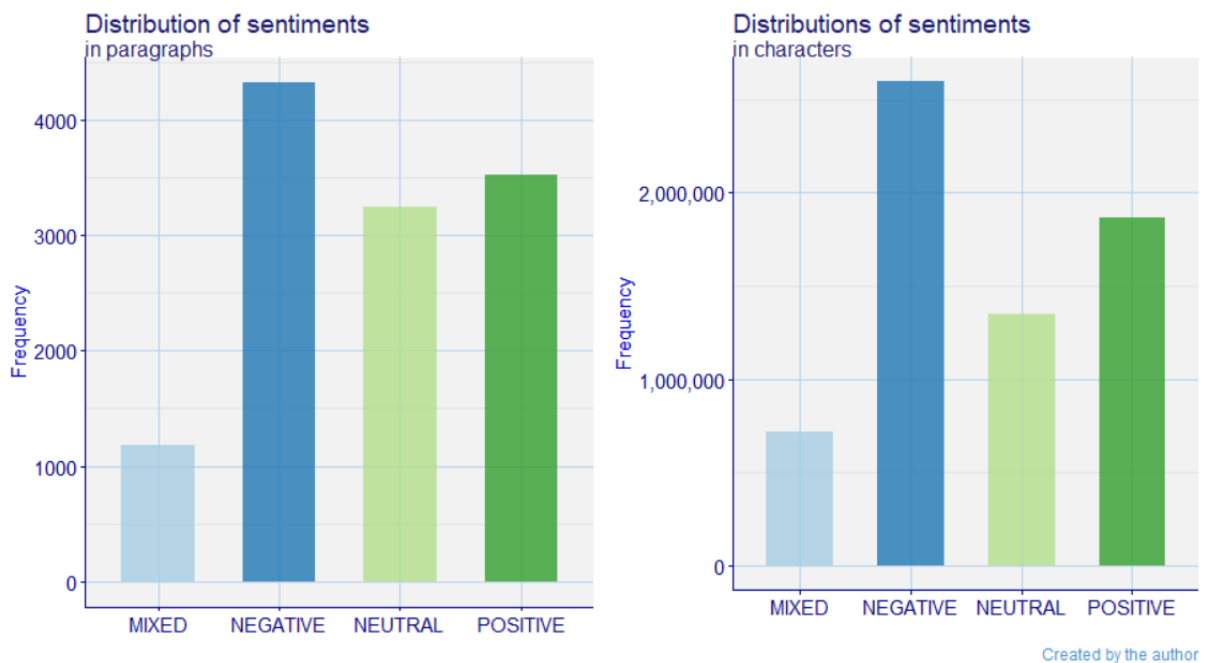
⁴ Amazon Comprehend <https://aws.amazon.com/comprehend/>

So each paragraph was analyzed with Amazon Comprehend which return a probability value for 4 sentiment categories: positive, negative, neutral, and mixed. Out of these 4 categories, the program determines the sentiment as the category with the highest probability.

3.4. Results and discussion

Figure 1 shows the distribution of sentiments in terms of paragraphs and characters as well. Obviously, the character is more appropriate here because then the text is weighted by the character and the diverse length of paragraphs won't mislead us.

Figure 1: Distribution of sentiments

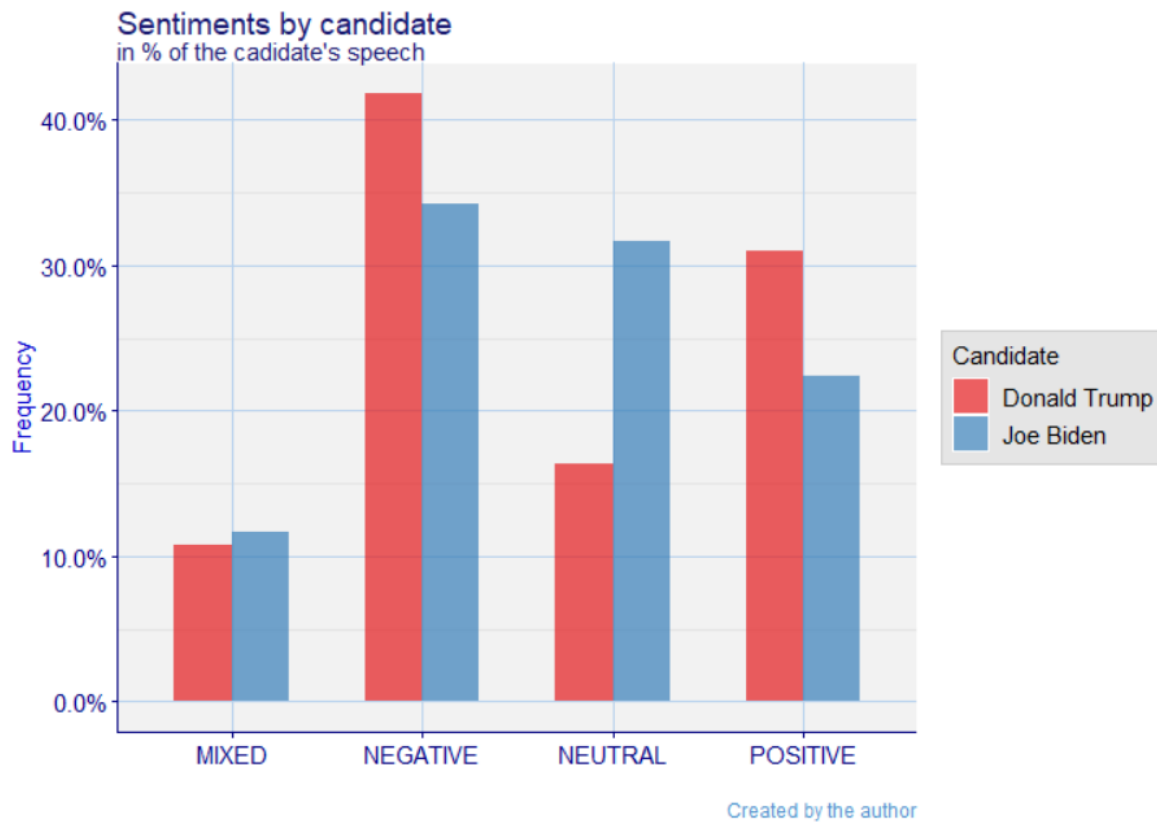


Apparently, negative sentiments dominate the speeches followed by positive, and neutral and there is a negligible amount of mixed sentiment in the text. As there are no publications yet about the 2020 US election's sentiment we can compare it only to the outcome of 2016. Our result resembles the negative tone of the 2016 campaign (Bayagich et al., 2017) and Donald Trump's rhetoric but it contradicts the result of Dilai et al. (2018) and Hoffmann, (2018) who both used lexicon-based sentiment analysis. Actually, what changed compared to 2016 is that the candidate besides Donald Trump was not Clinton but Biden. So let's see how the sentiment is distributed between the two politicians.

Figure 2 shows the distribution of sentiments by the candidate as percentage of the candidate's speech. Both nominees' rhetoric is mostly negative but Donald Trump prevails over Biden with +8% in terms of negativity. Trump also has more (~11%) positivity while Biden's neutral speech is almost as frequent as his positive (32% and 34% respectively). The results in

figure 1 and figure 2 indicate, that we can confirm our first hypothesis i.e. negativity dominated in campaign speeches of the 2020 US election. However, in the second hypothesis, we expected Trump will use more positivity and less negativity than Biden, but actually, Trump prevails in both positivity and negativity, while Biden used much more neutral sentiment.

Figure 2: Sentiments by candidate



We can conclude that negativity dominates in presidential speeches, the remaining question is why. Are they connected to negative campaigning or media negativity? We cannot claim either of them for sure, but negativity in politics has been recognized and studied by others too, like Soroka (2014). His book *Negativity in Democratic Politics: Causes and Consequences* was reviewed by Muddiman (2015) Soroka points out that humans are biased toward negativity. This negativity bias is defined as the “relative strength of negative over positive” (Muddiman (2015) cites Soroka, (2014)). Soroka explains that negativity psychologically arouses individuals, captures their attention, and reassures them to consume more of such information which inspires journalists to excessively publish negative content. He concludes from neurological and psychological studies that human beings psychologically and biologically favor negativity over positivity in various contexts and political messages are motivated by the same process.

In one of his experiments, he analyzed the long-term sales of magazines with negative covers, compared to neutral and positive covers. He found that negative cover magazines were sold at a higher rate than positive or neutral cover magazines. In another experiment, he concluded, that the psychological response to negative campaign advertisements and news was different from the response to such positive messages. He explains, that people are physiologically aroused by and drawn to negative messages.

Based on this theory, we can say with some exaggeration, that negativity is a weapon of political rhetoric and campaigns, which is indeed used by the most prominent officials of the United States. This proceeding certainly sabotages the rational decision-making of democratic systems but is it totally approved in nowadays intense fights during the US campaign period. Needless to say, that the end justifies the means in this case.

4. Emotion analysis and thematic words

4.1. Emotion analysis

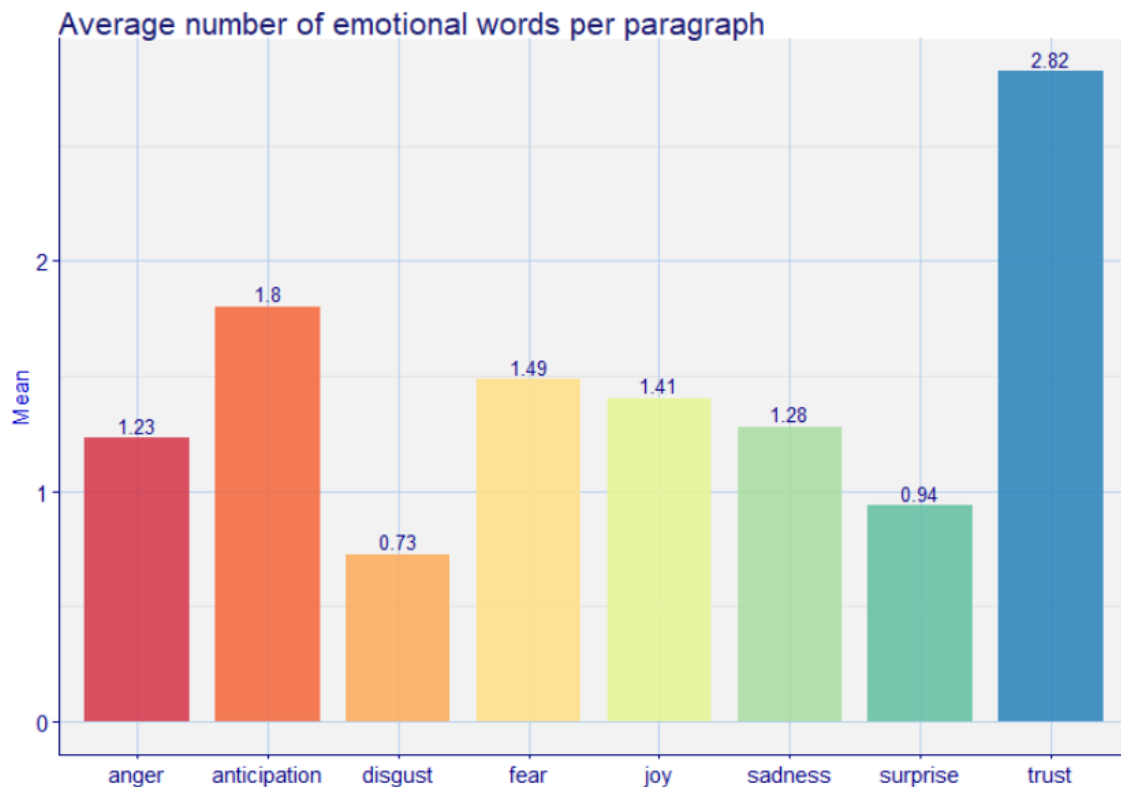
The sentiment analysis measured the polarity of sentiment from negative to positive. However, there are much more emotions that can be distinguished within presidential speeches. For this we used the same Syuzhet package (Jockers, 2020) as Hoffmann (2018) used. This package is a lexicon-based emotional emotion-word association. It uses the NRC Word-Emotion Association Lexicon (Mohammad, n.d. , S. Mohammad & Turney, 2010) which includes the relation of words with 8 elementary emotions, that were manually annotated by crowdsourcing. It means, that the association between words and emotion we annotated by humans in an efficient way including quality checks as well.

The 8 selected emotions are anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. The creators explain, that it is an extended list of the basic 6 emotions described by Ekman (1992), namely joy, sadness, anger, fear, disgust, and surprise. The NRC lexicon's creators' argument behind the 8 emotions is based on the study of Plutchnik (1980) that emotions are naturally can be paired with opposites: joy–sadness, trust–disgust, anticipation–surprise, and anger–fear. Besides, every more complex emotion can be considered as a combination of these 8 basic emotions. So let's see how the presidential campaign speeches express these 8 distinct emotions.

Figure 3 shows the average number of words per paragraph associated with the given emotion and figure 4 shows the total number of emotional words in the full dataset. It is important to view the text from both aspects, not just draw conclusions on either just the paragraph level

insight or just on the measure of the full dataset. Obviously, the two charts are consistent, but the differences between mean values on paragraph level suggest it is worth also exploring the big picture on the full dataset.

Figure 3: Average number of emotional words per paragraph



Apparently, trust is by far the most usual expression in the speeches, followed by anticipation. Fear and joy are a step behind anticipation, while anger and sadness are also at a lower level. Surprise is the one but last, while disgust has an even much lower level that makes it the least frequent expression.

Interpreting these results is not so straightforward. Disgust is not a necessary element of any campaign, but it could be used in negative campaigning to attack the opponent. Surprise, in our opinion, has no much role in political campaigns, but it can be a tool for creating the campaign mood and atmosphere of speeches. Surprise is an ambivalent feeling because it can be associated with good and bad objectives as well. Example words in the NRC lexicon associated with surprise are: *vanish, winning, unstable, terrorist, playful, optimism, emergency, heroism, judgment, greatness, and enthusiasm*.

Joy, fear, sadness, and anger are all vital elements of campaign speeches and because of the coronavirus and recession in 2020, the negative emotions are not unexpected. Finally, anticipation and trust seem trivial emotions in campaign speeches. Politicians must focus on

gaining the trust of citizens for sure. In the NRC lexicon trust is associated with the following examples: *believing, community, rational, diligence, virtue, unity, supporter, sincere, loyalty, measure, hope, improvement* and so on. Besides, anticipation is also understandably expressed by candidates, because their plans, ideas, and work should be converted to the future in election speeches. Some examples of that are: *begun, coming, continue, exciting, happen, planning, probably, ready, succeed* and *youth*.

Figure 4: Total number of emotional words

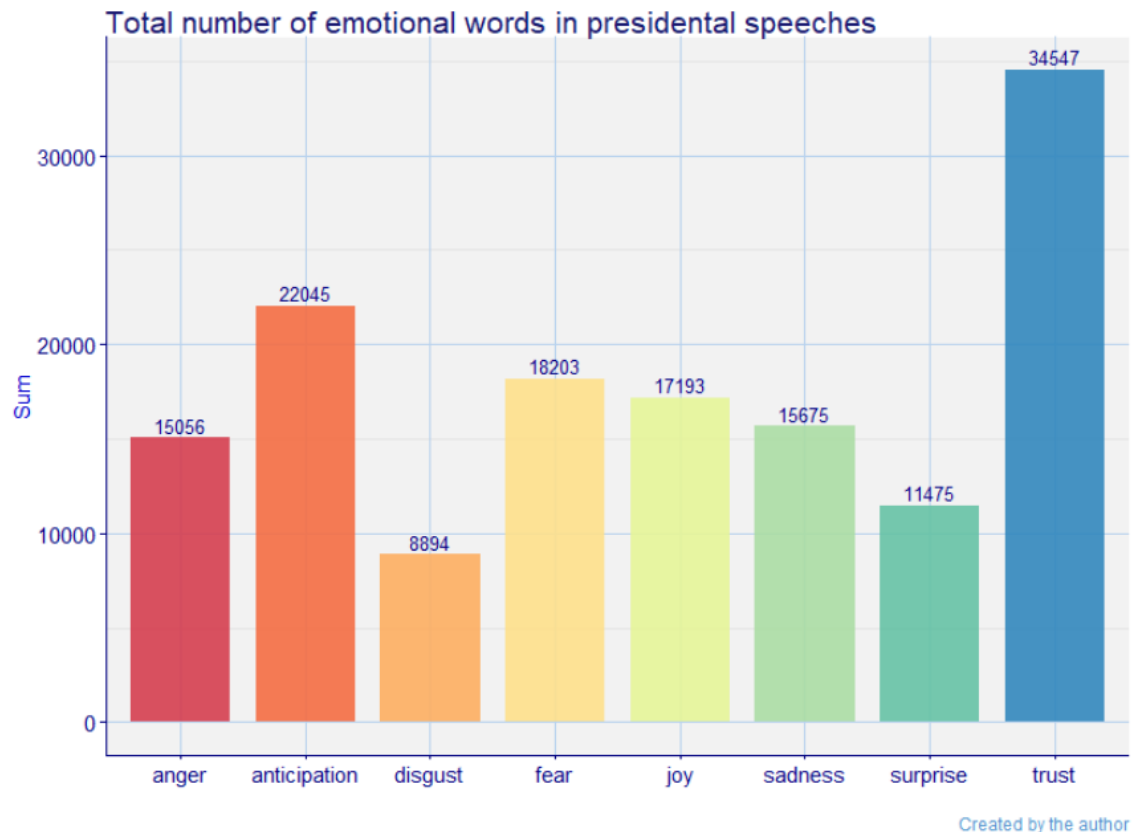
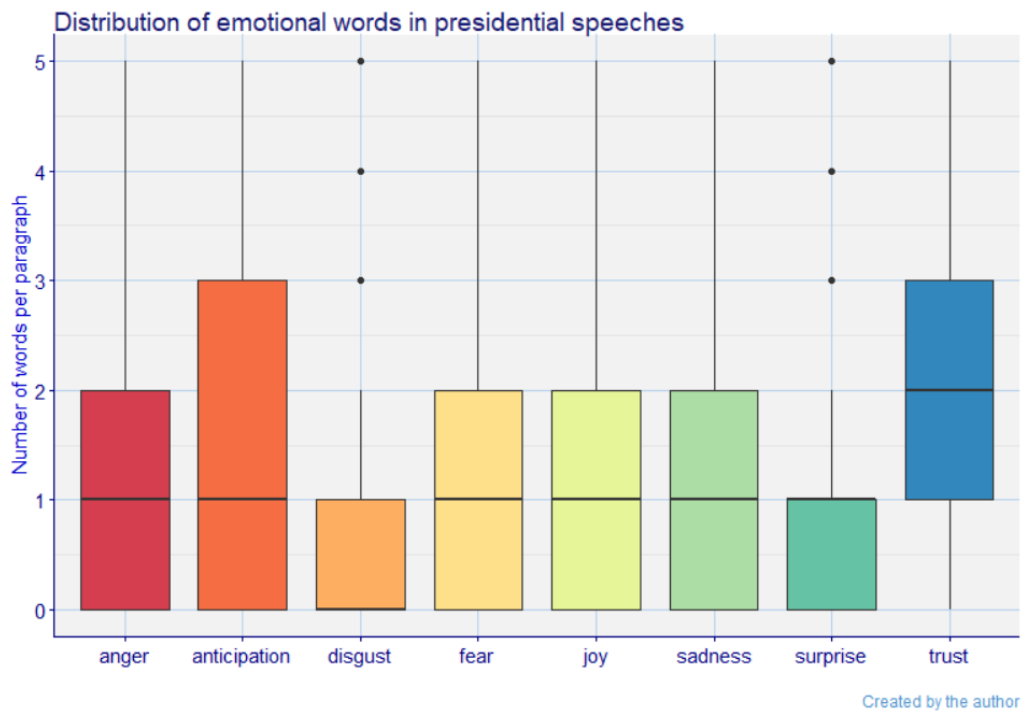


Figure 5. shows the boxplot of the number of words for each sentiment. The distribution visualized by boxplot is also a necessary check, to ensure that we do not deal with an extreme case where all the emotional words are in 1 paragraph and all other paragraphs have no such content. The median value is 2 in case of trust and 0 in case of disgust while 1 in every other case. The interquartile range (the range between the 25th and 75th percentile) is the highest for anticipation, meaning that there is greater variation in the values, and it is the lowest for disgust and surprise. For all other emotions examined, the range is from 0 to 2.

Figure 5: Boxplot of emotional words



But why trust and anticipation are the dominant emotions? Actually, we can somewhat understand it by understanding the mechanism of the campaigns themselves. Kiss (2013) explains that political campaigns are not just about the citizens but also instead of the citizens, meaning that campaigners define the voters' interests and beliefs instead of the voters themselves, which makes the politicians caring and trusted. Behind this caring attention, there is anticipation and hope. The indifference of citizens decreases when someone speaks about them and they hear what is their best interests and opinion according to someone.

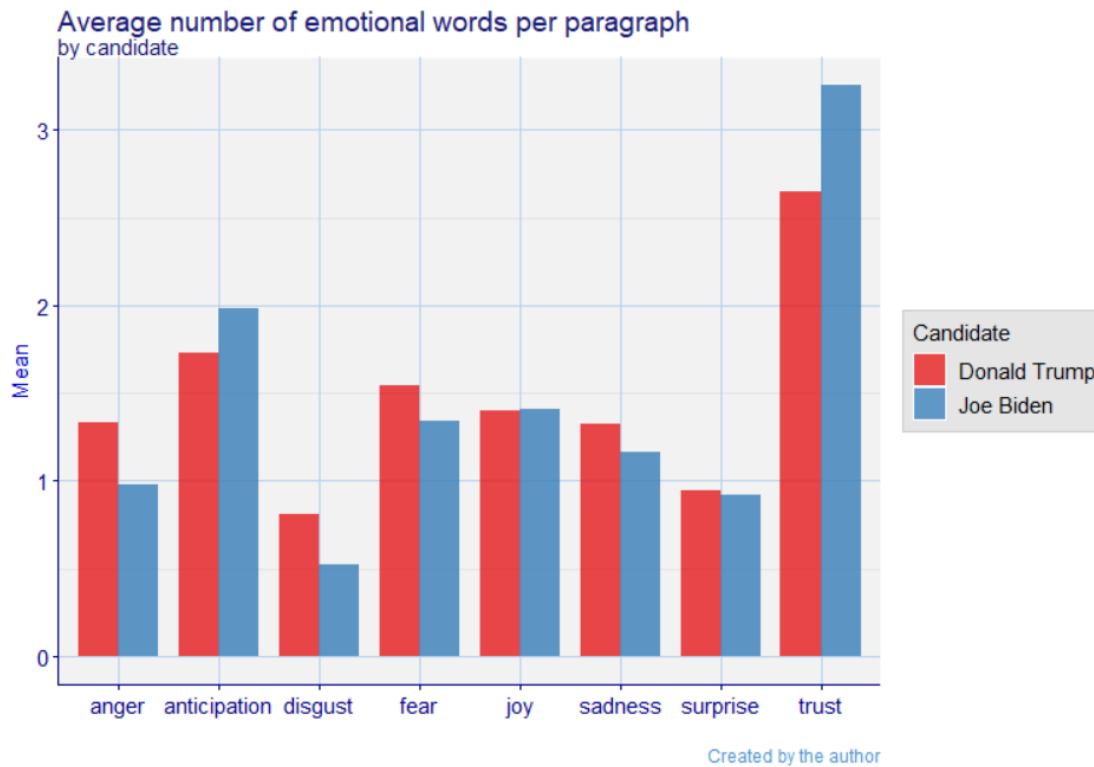
In practice, politicians tell us that we need “something”, and they care about us so if we vote on/trust them, they will give us this “something”. So this caring mechanism is built on the anticipation that they will do “something” and they can be trusted. Some citizens contradict in their values and opinions so they search for other possibilities but some citizens recognize their appropriate representative so they give their vote to this candidate (Kiss, 2013). Thus, campaigns aim for decreasing indifference by activating emotions through a caring attitude that expresses trust and positive anticipation.

4.2. Comparison of candidates

So far we have seen the emotion in the speeches of both candidates altogether, however, we can make a distinction between them too. Figure 6 shows the average number of emotional words by candidate while figure 7 displays the boxplot of the emotional words so we can see

that the median and interquartile ranges are almost the same. The reason why we use the average number of words per paragraph is that as mentioned before, Trump has more than twice as much text in the data, so it would be pointless to compare them based on the total values.

Figure 6: Average number of emotional words by candidate

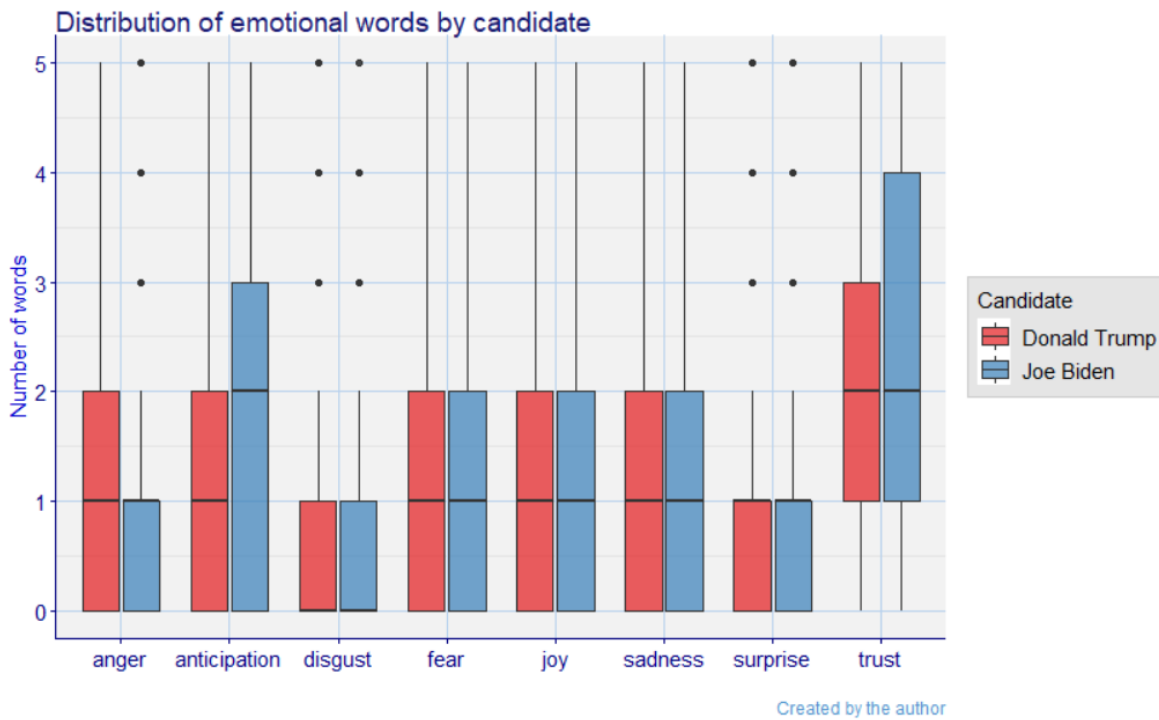


The general trend that we have seen in figure 3 and figure 4 is also visible in figure 6. Trust and anticipation are the most expressed emotions based on the average value for both candidates but Biden prevails over Trump in both of these cases. Despite that, Donald Trump uses more terms expressing fear, anger, sadness, and disgust compared to President Biden. Interestingly, joy and surprise are expressed in very similar amounts by the two politicians.

The results are somewhat similar to Hoffmann (2018) who showed that the top emotion is trust which is followed by anticipation and joy in the case of Hillary Clinton but trust is followed by fear and anticipation respectively with a minor difference between their frequency in Trump's speeches. So the pattern of Clinton is also true for Biden, yet Trump used a bit more anticipation than fear in 2020 compared to the election campaign of 2016.

Apparently, Biden is more efficient in expressing emotions behind the fundamental mechanisms of campaign, namely trust, and anticipation discussed in the previous subsection. We cannot claim that this is the reason behind his victory yet only because it was not the case for Clinton, but the differences and the patterns are remarkable.

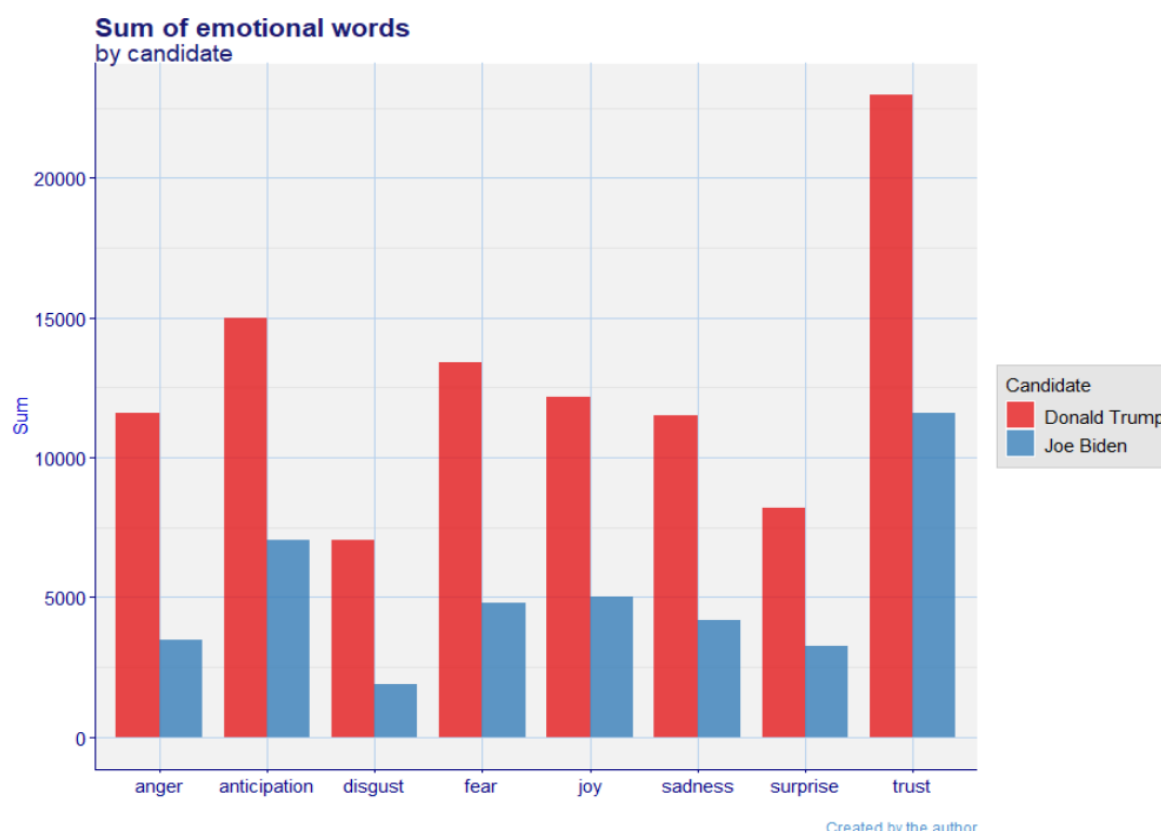
Figure 7: Boxplot of emotional words by candidate



The boxplot again just confirms the mentioned patterns, we see in the distribution too how Biden prevails in expressing trust and anticipation. However, in other categories, they are very similar, except for anger, which is more dominantly used by Biden. The number of the emotional words are obviously integer numbers, that is why there is not much variation in the median values. The median is 1 in case of anger, fear, joy, sadness, and surprise, zero in case of disgust, and 2 in case of trust for both candidates. The only emotion where the median value is different between the candidates is anticipation: 2 for Biden and 1 for Trump.

As mentioned, the average word per paragraph is the measure for comparison because we have twice as much data about Trump. However, the sum of emotive words by the candidate in the full dataset is also visualized in figure 8. It is not useful to compare the candidates but is useful to observe the previously detected and described pattern, for both candidates. The order of frequency is the same, but now we can see the differences in absolute terms.

Figure 8: Total sum of emotional words by candidate⁵



4.3. Thematic word frequency

Besides the sentiment and emotion analysis, it would be interesting to see which words and terms they use often. The frequent terms indicate that the candidate regularly spoke about that topic so it is important to him, it is his focus throughout the campaign. Our report is similar to Liu & Lei's (2018) but the methodology is simpler, yet still helps us understand the candidate's rhetoric.

To get the most frequent word of the candidates I tokenized the speeches and removed the stop words (commonly used words but without significant content or meaning). Then I selected the top 50 most frequent words for each candidate, but this list of words is not really revealing because most words are used by both of them and are probably part of any American political speech. That is why I filtered those words which are not in the top 50 words of the other candidate therefore I got the unique top words of each candidate.

The list of words with their frequency is illustrated in figure 9. Obviously, the frequency values in the case of Trump are much higher because he had more speeches in the dataset. Interestingly, both of them mention the other candidate in a top 3 place which can be a sign of

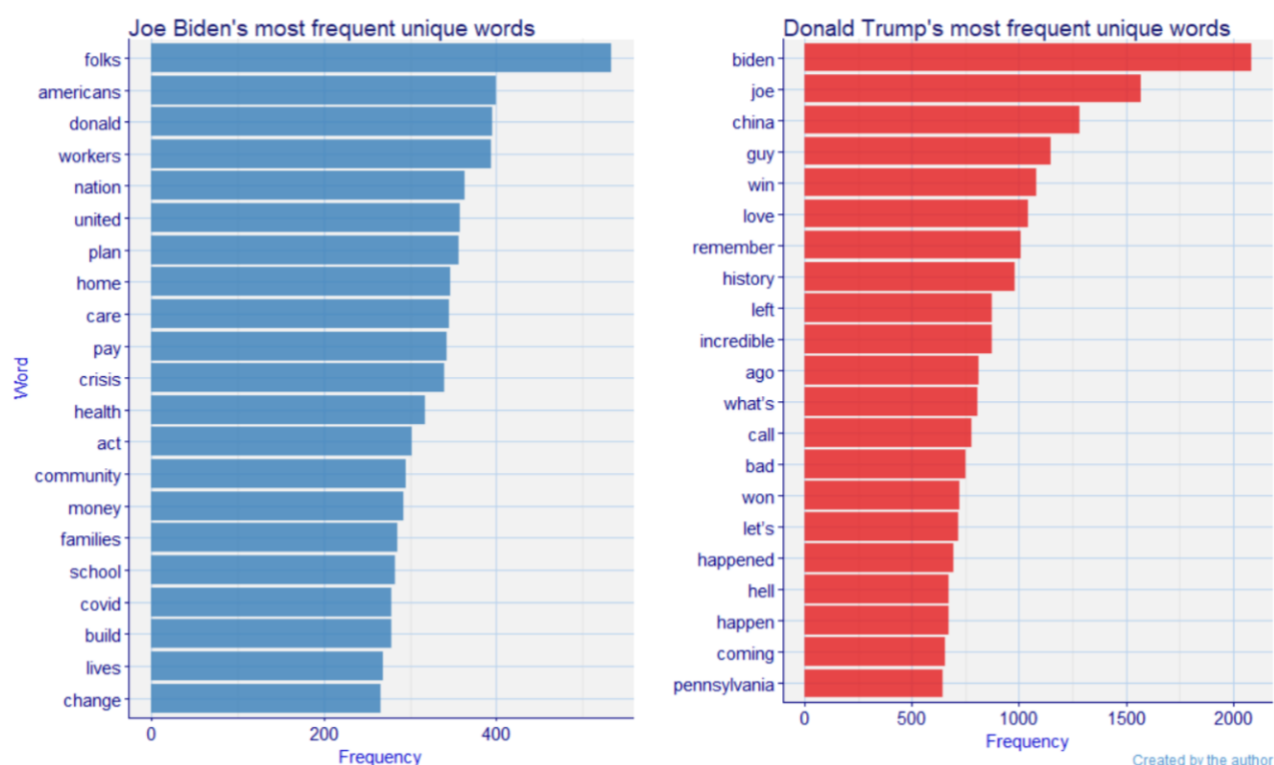
⁵ Donald Trump has more than twice as much speeches than Joe Biden

negative campaigning. Besides, there are major differences between the topics which they talk about. Joe Biden frequently mentions topics of political programs and social issues like *health*, *pay*, *workers*, *home*, *crisis*, *school*, *family*, *community*, *lives*, and *covid*. The Democratic candidate also mentions noticeably more terms referring to unity like *Americans*, *nation*, *united*, *community* and the word *change* has also relevant place in his rhetoric understandably because he wants to dismiss the Republican reign and change the trajectory of the United States.

Compared to that, Donald Trump used more verbs and emotionally more saturated words like *win*, *love*, *remember*, *history*, *incredible*, *bad*, *won*, *happened*, and *hell*. This emotional saturation helps us understand why Donald Trump prevails in both positive and negative sentiment as presented in chapter 3. We can also see, that *China* is the top third expression for Trump which properly reflects the former president's antipathy toward the country not just because of the spread of coronavirus but also because of its economic threat.

These are major differences between their rhetoric, the themes they talking about the most, could not be further from each other. As it is expressed in their rhetoric, the base of this is the difference in their political affiliation, standpoint, program, virtue, and so on. And as we know, Biden won the election at this time. However, we do not claim that his terms were the key to his success but it is a useful comparison to understand the differences and learn from them for future political events.

Figure 9: Most frequent unique words



5. Discussion

The empirical results of the study are presented in the previous 2 chapters, however, the findings can be further discussed. In this chapter, we project the results of the 3rd and 4th chapters to some theoretical models and determine the greatest weapon of campaigners regarding emotional influence to maximize their results. We detect the potential errors of the candidates for the 2020 US presidential election and make recommendations to subsequent politicians. Finally, we discuss the limitations and suggestions for further research.

Soroka's (2014) theory suggests that people are biased towards negativity, therefore, if public figures want to get attention the easiest way to get that is by emotionally negative content. Politicians also have lots of room for that, they can consistently highlight unfortunate events and situations, build a negative campaign on their opponents and consistently attack them, create or just discover new dangers or threats to the county, and so on. This strategy seems legit until reasonable amount of negativity, however, the candidate's position clearly influences the goal and methods towards that. Based on the theory explained in the paper of Crabtree et al. (2019) the incumbent party's communication should be more positive. Campaign messages communicated by positive emotive language make people adopt a positive frame when they are evaluating the state of the world. The state of the world is attributed to the incumbent's performance, so when individuals evaluate the performance above a certain threshold, they vote for the incumbent party. On the other hand, when the evaluation of the current state of the world is below that certain threshold, they will switch to the opposition. Therefore, the vote is determined by how voters perceive the current state.

Retrospective voting is only possible when individuals can attribute at least some responsibility for performance to the incumbent party. Regarding the election of 2020, we have to assume, that individuals believed, that Donald Trump had some responsibility in the state of the world at that time. This assumption seems rather real, even considering the coronavirus pandemic too, which was most probably not his fault. His activity for 4 years as president clearly had some impact which is recognized by the citizens.

Therefore, Trump, who was the candidate in office should have used more positive language, and positive framing, that would ensure citizens that he performed well, the economy is good despite the coronavirus, everything is great in the US, and so on. However, he is the one who used more negativity. We can also argue based on the media negativity theory of Soroka (2014) that the politicians should use more negativity in order to gain attention, which is also valid. But out of the 2 candidates, Trump should be the one using more positivity because of this

incumbent position. Trump also used more positivity than Biden, but more negativity too and negativity is the dominant emotion in his campaign speeches.

Although this study argues, that the incumbent party should have used more positivity than the opposition, we do not claim that it would have changed the results or this strategy is so confirmed, that it makes anyone a winner who follows it. The existing research about campaign sentiment is not so well-founded yet but it is getting more and more attention. The difficulty of any campaign recommendation is that every campaign environment and election happens only once, so no one will ever know what could have happened if something is changed.

Regarding the emotion analysis, Trump has expressed less trust and anticipation compared to Biden, which we argue to be essential based on the caring mechanism explained by Kiss (2013). However, the case was the same during the campaign of 2016 analyzed by Hoffmann (2018). At that time, Trump also used less trust and anticipation in his speeches than Clinton and he won the election. We do not claim that it had so significant impact on the results, however, we believe based on the theory of the caring mechanism, that expressing more anticipation about the future and using language expressing more trust could be beneficial for politicians. In such tight fights like the election of 2020 every parameter of the campaign communication matters.

The campaign communication can be also viewed as a limitation of the research. This study mainly analyzed public speeches and some debates, but we know that probably the majority of these were prepared by a speechwriter. It does not deduct anything from the findings and merits of the research, but we should mention that not just the candidate but the speechwriter, the campaign staff, and the whole party is reflected in these speeches. Social media data like Twitter could have been also used but since then, Trump has been banned from the site so we cannot make a comparison. Another limitation of the study is the emotion analysis, which heavily relied on the *Syuzhet* package (Jockers, 2020) and therefore used the given 8 emotions. However, it is arguable that several another, more complex emotions are included in these speeches and these might have additional impact. Others could also argue, that the used 8 emotions are too broad considering the basic emotions, and for example, anticipation is too complex. We believe that using these 8 emotions is kind of the golden mean. Our other argument besides *Syuzhet* is that there is no much better open-source program on the market, in which we could select such emotions and *Syuzhet* is perfectly suitable for political content.

To have more developed and effective results, future researchers should focus on establishing more and more accurate tools for emotion analysis and also on the effect of different sentiments and emotions on voters with the same objective content just like Brader's (2005) study. Better

understanding these influential effect is a key for political campaigns. On the other hand, researchers should also focus on analyzing the politicians themselves, to recognize their keys or flaws to success and failure.

6. Conclusion

Emotions in politics have been the target of discussion for centuries but still, there is no exact conclusion about its effect, yet most researchers agree that emotions play an important role in politics and political campaigns as well. By reviewing some prominent studies, we argue that emotions can influence political participation, behavior, and opinion, consequently, it is important to further study the emotions expressed by political entities.

Subsequently, we discuss the development of sentiment and emotion analysis and look at some remarkable use cases of sentiment analysis of political content. Noticeably, there is limited research on investigating the political actors themselves, moreover, the US presidential election of 2020 has not been covered yet at all.

In this study, we analyzed the presidential campaign speech of Joe Biden and Donald Trump to understand what emotional expression strategies they use. Our first hypothesis is that they dominantly used negative sentiment based on Soroka's (2014) theory. The second hypothesis is that Trump used more positive and less negative sentiment compared to Biden, because of his incumbent position he has to use positive framing to ensure voters that he performed well and encourage them to further support him. We also asked an open question about what kind of emotions the candidates use exactly.

Our sentiment analysis with Amazon Comprehend confirms the hypothesis that candidates prevailingly use negative sentiment in their speeches, which is matched by some existing research (Bayagich et al., 2017) (Liu & Lei, 2018). We have also found that Trump expresses more negative and more positive sentiments compared to Biden, while the latter uses significantly more neutral expressions in contrast to his opponent. Therefore, we cannot confirm our second hypothesis. The theory of Soroka (2014) explains why negativity gets more attention and importance in media and consequently, in political messages as well. However, the too much negativity from Trump might have been a fault, because as president in office, he should have framed his messages more positively.

For emotion analysis, we used a lexicon-based methodology, which showed that trust and anticipation are the ruling emotions in the candidates' speeches, furthermore Biden dominates Trump in both of these categories. The dominance of trust and anticipation can be interpreted by the mechanism of campaigns and the caring functionality of political manifestations. The

divergence between the candidates can be better understood by their commonly used terms and topics.

To sum up, we found that today's most outstanding political leaders use negativity, trust, and anticipation in their rhetoric to convince citizens about their program, leadership, and party. The general inference is that these are the tools or manipulations which are the most useful for political leaders and emerging politicians, moreover previous research also concludes the same about past election campaigns. Whether it is just a temporary pattern, the characteristics of the recent candidates, or just the political culture of the United States, remains unanswered but it might be the combination of multiple known or hidden phenomena and indeed the greatest weapon in political campaigns. This research aimed to collect empirical evidence by modern large-scale methodology and confirmed by some theoretical models, but the comprehensive theoretical background, the causes and consequences can be further studied either from political and social science aspects or from a psychological perspective.

Bibliography

- Akbar, P., Irawan, B., Taufik, M., & Nurmandi, A. (2021). Social Media in Politic: Political Campaign on United States Election 2020 Between Donald Trump and Joe Biden. *International Conference on Human-Computer Interaction*, 359–367.
- American Psychological Association. (n.d.). *APA Dictionary of Psychology*. Retrieved April 15, 2022, from <https://dictionary.apa.org/priming>
- Arkes, H. (1993). Can Emotion Supply the Place of Reason? Comments on the Chapters. In G. E. Marcus & R. L. Hanson (Eds.), *Reconsidering the democratic public* (p. 287). Pennsylvania State University Press.
- Bayagich, M., Cohen, L., Farfel, L., Krowitz, A., Kuchman, E., Lindenberg, S., Sochacki, N., Suh, H., & Soroka, S. (2017, January 19). Exploring the Tone of the 2016 Campaign. *Center for Political Studies (CPS) Blog*. <https://cpsblog.isr.umich.edu/?p=1884&p=1884>
- Bernábe-Loranca, M. B., González-Velázquez, R., Carrillo-Canán, A., & Granillo-Martínez, E. (2022). Sentiment Analysis and Multiple Means Comparison for the 2020 United States Elections. *Computación y Sistemas*, 26(1).
- Brader, T. (2005). Striking a Responsive Chord: How Political Ads Motivate and Persuade Voters by Appealing to Emotions. *American Journal of Political Science*, 49(2), 388–405. <https://doi.org/https://doi.org/10.1111/j.0092-5853.2005.00130.x>
- Brader, T., & Marcus, G. E. (2013). Emotion and political psychology. In L. Huddy, D. O. Sears, & J. S. Levy (Eds.), *The Oxford Handbook of Political Psychology* (2 ed, pp. 165–204). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199760107.013.0006>
- Chandak, A. (2022). Sentiment Analysis on Tweets in the 2020 US Presidential Election. *Journal of High School Science*, 32492.
- Chandra, R., & Saini, R. (2021). Biden vs Trump: Modeling US General Elections Using BERT Language Model. *IEEE Access*, 9, 128494–128505. <https://doi.org/10.1109/ACCESS.2021.3111035>
- Chaudhry, H. N., Javed, Y., Kulsoom, F., Mehmood, Z., Khan, Z. I., Shoaib, U., & Janjua, S. H. (2021). Sentiment analysis of before and after elections: Twitter data of US election 2020. *Electronics*, 10(17), 2082.
- Chauhan, P., Sharma, N., & Sikka, G. (2021). The emergence of social media data and sentiment analysis in election prediction. *Journal of Ambient Intelligence and Humanized*

- Computing*, 12(2), 2601–2627.
- CNN. (2020). *Presidential Election Results and Electoral Map 2020*. CNN Politics. <https://edition.cnn.com/election/2020/results/president>
- Crabtree, C., Golder, M., Gschwend, T., & Indridason, I. H. (2019). It Is Not Only What You Say, It Is Also How You Say It: The Strategic Use of Campaign Sentiment. *The Journal of Politics*, 82(3), 1044–1060. <https://doi.org/10.1086/707613>
- Dilai, M., Onukevych, Y., & Dilay, I. (2018). Sentiment analysis of the US and Ukrainian presidential speeches. *Computational Linguistics and Intelligent Systems* (2), 2018, 60–70.
- Edelman, M. (1988). *Constructing the political spectacle*. University of Chicago Press, Chicago.
- Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6(3–4), 169–200.
- Enyedi, Z. (2004). Érték, értelem, érzelem. Politikai pszichológia Bibó nyomán. *Politikatudományi Szemle*, 4, 5–18.
- González-Bailón, S., & Paltoglou, G. (2015). Signals of Public Opinion in Online Communication: A Comparison of Methods and Data Sources. *The ANNALS of the American Academy of Political and Social Science*, 659(1), 95–107. <https://doi.org/10.1177/0002716215569192>
- Hoffmann, T. (2018). “Too many Americans are trapped in fear, violence and poverty”: a psychology-informed sentiment analysis of campaign speeches from the 2016 US Presidential Election. *Linguistics Vanguard*, 4(1), 20170008.
- Hopkins, D. J., & King, G. (2010). A Method of Automated Nonparametric Content Analysis for Social Science. *American Journal of Political Science*, 54(1), 229–247. <https://doi.org/https://doi.org/10.1111/j.1540-5907.2009.00428.x>
- Iacus, S. M., Porro, G., Curini, L., & Ceron, A. (2014). Every tweet counts? How sentiment analysis of social networks can improve our knowledge of citizens’ policy preferences. An application to Italy and France. *News Media & Society*, 16(2), 340–358.
- Jockers, M. (2020). *Introduction to the Syuzhet package*. The Comprehensive R Archive Network. <https://cran.r-project.org/web/packages/syuzhet/vignettes/syuzhet-vignette.html>
- Kiss, B. (2013). Érzelmek és politikatudomány. *Politikatudományi Szemle*, 22(3), 7–28.
- Kühne, R., Schemer, C., Matthes, J., & Wirth, W. (2011). Affective Priming in Political Campaigns: How Campaign-Induced Emotions Prime Political Opinions. *International*

- Journal of Public Opinion Research*, 23(4), 485–507. <https://doi.org/10.1093/ijpor/edr004>
- Liu, D., & Lei, L. (2018). The appeal to political sentiment: An analysis of Donald Trump's and Hillary Clinton's speech themes and discourse strategies in the 2016 US presidential election. *Discourse, Context & Media*, 25, 143–152. <https://doi.org/https://doi.org/10.1016/j.dcm.2018.05.001>
- Mäntylä, M. V, Graziotin, D., & Kuuttila, M. (2018). The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. *Computer Science Review*, 27, 16–32. <https://doi.org/https://doi.org/10.1016/j.cosrev.2017.10.002>
- Marcus, G. E. (2000). Emotions in politics. *Annual Review of Political Science*, 3(1), 221–250.
- Marcus, G. E. (2010). *Sentimental citizen: Emotion in democratic politics*. Penn State Press.
- Mohammad, S. (n.d.). *NRC Emotion Lexicon*. Retrieved November 29, 2021, from <http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>
- Mohammad, S., & Turney, P. (2010). Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, 26–34.
- Muddiman, A. (2015). Negativity in Democratic Politics: Causes and Consequences, by Stuart N. Soroka. *Political Communication*, 32(2), 339–341. <https://doi.org/10.1080/10584609.2015.1018110>
- Plutchnik, R. (1980). *Chapter 1 - A GENERAL PSYCHOEVOLUTIONARY THEORY OF EMOTION* (R. Plutchik & H. B. T.-T. of E. Kellerman (eds.); pp. 3–33). Academic Press. <https://doi.org/https://doi.org/10.1016/B978-0-12-558701-3.50007-7>
- R Core Team. (2021). *R: A language and environment for statistical computing* (4.1.2). R Foundation for Statistical Computing. <https://www.r-project.org/>
- Schwarz, N. (2000). Emotion, cognition, and decision making. *Cognition & Emotion*, 14(4), 433–440.
- Singh, A., Dua, N., Mishra, V. K., Singh, D., & Agrawal, A. (2021). Predicting Elections Results using Social Media Activity A Case Study: USA Presidential Election 2020. *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 1, 314–319.
- Soroka, S. N. (2014). *Negativity in democratic politics: Causes and consequences*. Cambridge University Press, Cambridge.

- Utych, S. M. (2018). Negative affective language in politics. *American Politics Research*, 46(1), 77–102.
- Weber, C. (2013). Emotions, campaigns, and political participation. *Political Research Quarterly*, 66(2), 414–428.
- Xia, E., Yue, H., & Liu, H. (2021). Tweet sentiment analysis of the 2020 US presidential election. *Companion Proceedings of the Web Conference 2021*, 367–371.