The influence of words on political marketing

ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

Master Thesis Marketing

Name: Mark Scholten
Student number: 349166
Study: Economics and Business Economics

Supervisor: Prof. Dr. Willem Verbeke

Abstract

This research focuses on the way politicians in the United States of America market themselves to the public. Specifically, the influence of linguistic markers in their use of language is studied. Using software, this research explores the differences in the vocabulary used by politicians running for president. Categories are formed that determine how politicians differ from each other for different variables. The probability for significant differences is modelled using independent t-test analyses. Findings illustrate statistically significant results in almost all analyses.

Index

Introduction

Politics and elections are a challenging playing field for both politicians and their respective campaign teams. The means of communicating one's message and gaining support are constantly evolving and changing, as technology plays a more vital role in campaigning. A recent phenomenon is the rise of the Internet as a medium for political communication and persuasion. Politicians now have more ways than ever of influencing potential voters as information becomes more widely accessible through the Internet. Whereas traditional media has been standing still, new media like Facebook, Twitter, Youtube, etc., gain more followers each and every day and thus a larger viewing audience (Duggan et al, 2015). These new sources of media outlets can change the nature of competition in politics and thus how politicians market themselves. With this new found well of information sources come new and extensive ways to study how politicians market themselves. By researching and analysing this, politicians can adapt their campaign strategy and use different linguistic markers in their vocabulary to position themselves better for the general public (Wattal et al., 2010).

Originally, electoral systems were designed to give candidates the opportunity to show the public who they are and what they stand for. However, there has been an interesting twist in how candidates nowadays campaign during the presidential race. Marketing as it seems now plays a pivotal role during the course of elections. Politicians now for instance can use scientific polling and other marketing techniques to get a better understanding of who the public is, and what they wish to hear. In return, political candidates feed back the ideas to the voters that they know will sell in the marketplace. In combination with politics, marketing techniques are now used and manipulated to influence the voters (Newman, 1999).

Politics, and in particular elections, have become big business in the United States of America. In 2008, the total spending on the presidential race was in excess of one billion dollars, marking it as the most expensive presidential race in history. Since then, the number in presidential contributions and expenditure has only increased. Logically, there is a lot financially riding on who becomes the next president. One could thus argue that a presidential candidate nowadays faces more difficulty with outside political influence than before. It is therefore vital that he or she retains his or her political voice and agenda (Mosk, 2008).

One of the prime goals of a politician is to increase his or her influence with the purpose of making society benefit as a whole. In achieving that, he or she will need to gain more support. Be it support in the form of more votes from the public, or the outspoken support from your political colleagues and peers. Next to that is also the financial support, which often acts as a tool to help achieve the other forms of support. The financial support usually comes in the form of Super PACs and lobbyists, or be it individual donations from the public. By allocating their financial resources to different marketing strategies, presidential candidates have more means to make their voice heard by a broader audience. In doing so, this will allow for more advertising, events, PR and other marketing activities that will enable presidential candidates to voice their beliefs more strongly to a larger audience and persuade them (Akande, 2012).

When addressing how politicians can use marketing to gain more support, it is vital that they voice a strong message. By communicating and advocating their beliefs and policies, candidates try to market themselves as the best possible candidate to vote for. They aim to display character traits and statements that resonate in the potential voters' mind. It is therefore essential that a politician picks his or her words carefully. The differences in the vocabulary used by politicians could have the potential to drastically alter the way the public perceives them. Studies have shown that voters tend to construct images of political leaders that reinforce and support their own ideological views. This is supported by a positive correlation that shows that those who admire a candidate, tend to see that candidate promoting policies they themselves favour. Alternatively, those who dislike a candidate tend to distance themselves from the policies that candidate promotes (Kinder, 1978).

Adding to that, research in the past has shown that politicians are more effective in using language to shape beliefs than they are with the actions they display in dealing with problems they are supposed to ameliorate. This fuels the argument that words have more meaning than actions do. The use of language can form a symbol for a politician in a sense that the ideological appeal can be stronger than the observable condition of that symbol. It can therefore also trigger different psychological responses in how we use language and respond to it (Edelman, 2013).

Moreover, when looking at the use of language, a recent study conducted by the Boston Globe for the 2016 presidential race found that there is a significant difference between the political candidates and their respective word choice and sentence structure. On the basis of the algorithm of the Flesch-Kincaid readability test, the candidates were scored on their use of language and ranked accordingly to the corresponding school grade level. The results depict vastly different levels on which the candidates communicate. Higher grades show more complexity and comprehension in terms of speech whereas lower grades indicate the use of more simple language. However, a politician's job isn't necessarily to educate. It is to inspire the public and persuade them. This often entails speaking in words that are easily accessible to the broadest possible audience. Therein it seems as if the use of language is the deciding factor that can define and differentiate presidential candidates from each other (Viser, 2015).

Consequently, the main question of this research will be: to what extent do politicians market themselves by their use of language? More to the point, this research is focused around the 2016 US elections and in particular the major candidates from both parties. Furthermore, it expands on previous research by not solely focusing on what a politician is saying, but by going more in depth into how a politician is communicating with their choice of words. Marketing hereby entails how the choice of words can differentiate political candidates from each other. The use of language refers to the word identification and association that is measured in different dictionary categories. These categories will act as the explanatory variables in determining how the choice of words can influence the marketing of politicians.

The results of this study could provide useful information for politicians, wishing to run for office, and marketing campaigns as marketing strategies are developed more and more to intensify the support for a presidential candidate. Studying and analysing the data, which provides information on speech patterns and linguistic markers that are associated with different psychologically-related categories, will enable to determine the optimal scenario in how a politician can use his or her words to position themselves differently from others. Moreover, it will allow for politicians and their campaign teams to form new strategies that are in coherence with their use of language and the audience they wish to reach. This could ultimately lead to totally restyled political marketing campaigns and the ability to distinguish yourself better in order to satisfy the public with your political agenda more efficiently (Tausczik & Pennebaker, 2010).

History has already shown that vocabulary can play a significant role in winning over the public. There can be quite a substantial difference in outcome when the public only hears a candidate speak versus seeing a candidate speak. This is of course illustrated by the infamous debate between then presidential candidates John F. Kennedy and Richard Nixon. People who watched the debate on television thought Kennedy won by a landslide, however the people who listened to the debate on the radio favoured Nixon. To this day, based on the evidence researches cannot come to a definitive agreement on what caused these different outcomes. Instead, it calls for further research to analyse the power of words (Kraus, 1996).

From a social standpoint, this study is highly relevant as many politicians and marketers seek out new and different ways to influence and persuade the public. Furthermore, it is aimed at the current US elections and how the presidential candidates thus far have voiced their opinions and beliefs. There has already been wide controversy about the remarks several candidates have made in order to attract more followers and distinguish themselves from each other. However, there is a lack of current information and research available that addresses these remarks and their effects (Zernicke, 1990). Many studies in the past have already focused on the sociological relevance of today's politics. Moreover, whilst marketing research is mostly done with the intent of influencing the public in order to generate more sales, in politics it is often used to identify what the public wants to hear. However, it doesn't necessarily address how the public wishes to hear it. There has been little to no research done on how these different linguistic markers and word choices can be used for a politician's marketing (Villar & Krosnick, 2011).

Scientifically spoken this study is very relevant. It expands on previous research that merely tried to identify sentence structure and the length of words politicians use. Whereas different uses of vocabulary can provide important psychological cues to people's thought processes and intentions, there is still a lot of unexploited territory to be studied (Tetlock, 1981). Adding to that, whilst conventional marketing research often entails the use of empirical data and evidence to support claims, neuromarketing offers a new wide range of tools and different data sources that can be used to get a clearer understanding of the use of language. More specifically, this research addresses the theoretical concepts behind categorising words in response to triggering different psychological reactions. With this, this research aims to combine the new insights neuromarketing has to offer and apply those with the help of the

Linguistic Inquiry and Word Count (LIWC) software to analyse and determine differences and patterns in the use of language of politicians (Tausczik & Pennebaker, 2010).

To get a clear overview of this research, the structuring of this paper will look as follows. The subsequent chapters focus firstly on developing a theoretical framework. The theoretical framework will add to the introduction and the literature to further discuss what is already known on the research topic and more importantly, what is lacking. Furthermore, on the basis of the literature the research question is further explored and elaborated on. The chapter will conclude with several hypotheses and a conceptual model that will support in answering the main research question. Following the theoretical framework, the methodology chapter will describe how the data is collected. More specifically, the characteristics of the data will be listed but more importantly, how the data is used to answer the main and sub-questions. This entails describing the different research methods and instruments like LIWC2015 and SPSS, that allow for statistical analysis to be done. Consequently, the results will depict what the research has revealed alongside a clear overview of how the data used. The most striking results that guide us to answering the main research question will be shown along with the corresponding tables and figures. Finally, after interpretation of the results, this research formulates an answer to the main research question. Furthermore, having answered the main research question, recommendations will be given on further research in cohesion with the limitations. With this clear structure, any uncontemplated problems that can occur will be solved.

Theoretical Framework

Before describing the methodology, it is vital to completely understand the research question and all the terms used. The main research question ‘To what extent do politicians market themselves by their use of language’ is made op from several components that need extra explaining. Firstly, politicians commonly refers to the group of people that either campaigns for or holds a position in government (Conge, 1988). In order to make this research more feasible, this is narrowed down to only campaigning for a position in government. Moreover, considering the vast amount of politicians running campaigns for different positions, this needs further reduction. The chosen position in government is that of the president of the United States of America. Furthermore, the politicians that will be used for analysis are

Hillary Clinton, Bernie Sanders, Ted Cruz and Donald Trump. As at the time of writing, they are the frontrunners in the polls to secure their respective party's nomination to run for president.

Secondly, market themselves revolves around political marketing. Although there is no single unambiguous answer that defines political marketing. In the academic field it is widely viewed and accepted as a subgenre in marketing. Its focus lies on the marketing activities a politician pursues in order to increase interest from the public and stand out from other competing politicians successfully (Scammel, 1999). More to the point, this research is focused on the linguistic differences between presidential candidates. Furthermore, some of the promotional activities politicians exploit include press advertisements, TV ads, party political broadcasts and election addresses. For research purposes, these marketing activities and events are narrowed down to debates, party conventions and campaign rallies as they provide a solid platform for candidates to verbally express themselves. Moreover, it is therefore essential that the data gathered comes from spoken language in order to get accurate results from the LIWC software for further statistical analysis (Pennebaker, 2015).

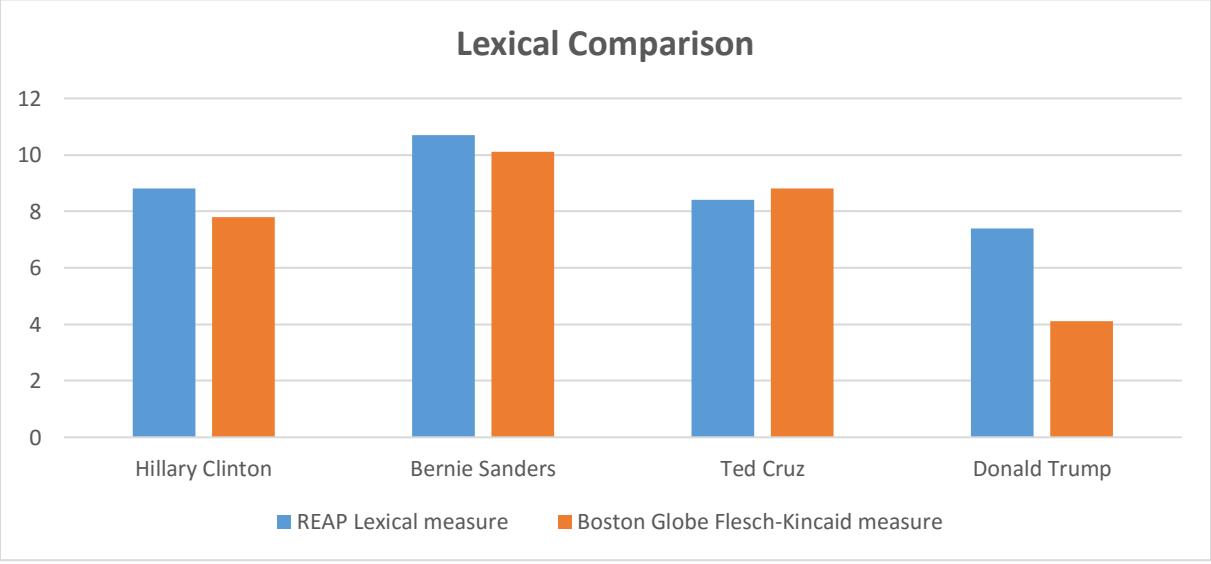
Lastly, use of language is at the centre of this research. This study is aimed to highlight differences in use of language between presidential candidates. Use of language hereby entails the politicians' respective word choice during the current US elections. Furthermore, to statistically identify and determine the differences in use of language, using the LIWC software the words are sampled together and categorised accordingly. The main categories are; linguistic processes, other grammar, psychological processes and punctuation. These categories will act as the key elements in translating spoken words to numbers for analysis purposes. Adding to that, each category has a different set of subcategories belonging to them. For example, linguistic processes contains elements of the Flesch-Kincaid formula like words per sentence and words containing more than six letters. Other grammar for example refers to common verbs, adjectives and comparisons. Furthermore, psychological processes has the largest set of subcategories and is presumed to be the most influential in highlighting differences in use of language. Some examples of these subcategories include affective process, social processes, cognitive processes, drivers and informal language. Lastly, punctuation logically lists the different conventional signs and typographical devices like periods, commas and question marks that are used to space out aloud speaking. The abovementioned categories are used when referring to use of language and the comparison

between the candidates. Additionally, exhibit A in the appendix gives a full overview of an example containing all the categories, except for punctuation, belonging to the LIWC software output (Pennebaker et al., 2015).

Current knowledge on the research topic stems from the REAP (REAder-specific Practice) project conducted in 2004. Its main purpose is to provide reader-specific practice for improved reading comprehension. A search model was developed that lists documents satisfying a set of diverse and possibly complex lexical constraints corresponding to a particular grade level (Collings-Thompson & Callan, 2004). This model was further developed to measure the lexical reading difficulty based on the smoothed individual probabilities of words occurring at each grade level and the corresponding syntax. Unlike the Flesch-Kincaid readability test, this model measures the frequency in which words are used whereas the Flesch-Kincaid formula depends on the number of words in each sentence and the number of syllables in each word. However, both studies made assumptions about what a difficult text was. More importantly, written speech is very different from spoken speech. When communicating verbally, we usually use less structured language and shorter sentences. So although these studies showcase a spectrum on which a politician can be scored, they aren't really reflective when considering spoken language (Schumacher & Eskenazi, 2016).

With this in mind however, figure 1 already shows some differences between the candidates' use of language when scoring them on grade level.

Figure 1



Results compiled from the Schumacher and Eskenazi study showing the grade levels of the politicians using different measuring techniques

Overall, there are more similarities than differences among the candidates for the different measuring techniques. However, Trump's vocabulary already shows hints of contradictory results. This could be because too few variables are taken into consideration. Furthermore, scoring them on grade level is quite subjective. However, these results form the basis of breaking the main research question down into a set of sub-questions that will enable to solve the problem statement and answer the main research question.

Hypotheses

Based on the recent studies and literature, the expectations are to see a significant difference in the use of language between Hillary Clinton, Bernie Sanders, Ted Cruz and Donald Trump. These expectations result from the analysis done with REAP lexical measure and the Flesch-Kincaid formula, but will be further explored using a wide range of different dictionary categories derived from the LIWC software. The research is focussed on determining the differences separately for each candidate, but also combined when grouping the candidates together in Democrats and Republicans. Furthermore, it will also analyse the differences across time for each candidate. It could be the case that some categories show no statistical difference, however overall it is expected to see more differences than similarities between the candidates and aforementioned political parties. Therefore, the first hypothesis is:

(1) Candidates market themselves. This hypothesis is tested for Hillary Clinton and Donald Trump as they are the frontrunners in the polls for presidency. The main focus is on marketing language as politics revolves around selling yourself. Slogans such as "I'm With Her" and "Make America Great Again" will be analysed in coherence with how Hillary and Trump voice their political agenda in linguistic terms. Previous literature may suggest Trump tries to market himself as a more intelligent person as he often states he comes from an excellent educational background (Bennett, 2015). Furthermore, stereotypes in male and female spoken language would propose Hillary will use more emotional and expressive language (Haas, 1979). Adding to that, Hillary will most likely use more social language directed at people's needs whilst Trump will talk more about business and economic needs (Haas, 1979). Lastly, the results from the

Boston Globe study would suggest Trump is a better salesperson as simpler language resonates more in the voters' mind.

The first hypothesis is represented by the statistical hypothesis:

(1^s) $H_0: \hat{p}_{x,i} = \hat{p}_{y,i}$

$H_a: \hat{p}_{x,i} \neq \hat{p}_{y,i}$

- Here, $\hat{p}_{x,i}$ and $\hat{p}_{y,i}$ are the average proportion that candidates x and y use dictionary subcategory i. The hypotheses will be tested across candidates and dictionary categories.

The second hypothesis states:

(2) Democratic candidates market themselves differently from Republicans by using different dictionary categories in their speeches. It is expected that Democrats focus more on the middle class and the poor. Early results from previous debates support this expectation by listing the number of times Democrats refer to these groups in comparison with Republicans (Luhby, 2015). Adding to that, the 2016 Democratic Party Platform often mentions the importance of how "Our country depends on a thriving middle class" along with "We are stronger together". Furthermore, it is expected that Republicans focus more on family values, tradition, the nation, reward and achievement. This is supported by the various statements the GOP Platform makes regarding these issues. For instance, "families and communities should be strong and free from government intrusion". The "constitution should be honoured, valued and upheld" which is followed by "we believe America is exceptional because of our historic role" and that "Americans have earned and deserve a strong and healthy economy". Regarding the differences in political standpoints and how these parties frame and address these issues, it is expected to see significant differences in how they voice their beliefs. For this analysis, Hillary Clinton and Bernie Sanders form the Democrat group and Donald Trump and Ted Cruz form the Republican group.

The second hypothesis is represented by the following statistical hypothesis:

(2^s) $H_0: \hat{p}_{d,i} = \hat{p}_{r,i}$

$H_a: \hat{p}_{d,i} \neq \hat{p}_{r,i}$

- Here, $\hat{p}_{d,i}$ and $\hat{p}_{r,i}$ are the average proportion that Democrat and Republican candidates use dictionary subcategory i. The hypotheses will be tested across dictionary categories.

The third hypothesis suggests that a candidate's use of language changes as their political campaign progresses. In business and branding, the first one focusses on is consistency (Schultz & Schultz, 2000). One could therefore say that Trump will be more consistent in his use of language than Clinton with regards to his background in business. However, Clinton has years of experience in politics and campaigning so a logical counter hypothesis would be that she is more consistent in her use of language. Next to this, previous research has indicated that presidents become more complex in their manner of thinking over time (Tetlock, 1981). One could therefore argue that this is also expressed in linguistic terms by presidential candidates over a period of time.

The third hypothesis will therefore be:

(3) A candidate's use of language changes throughout their political campaign.

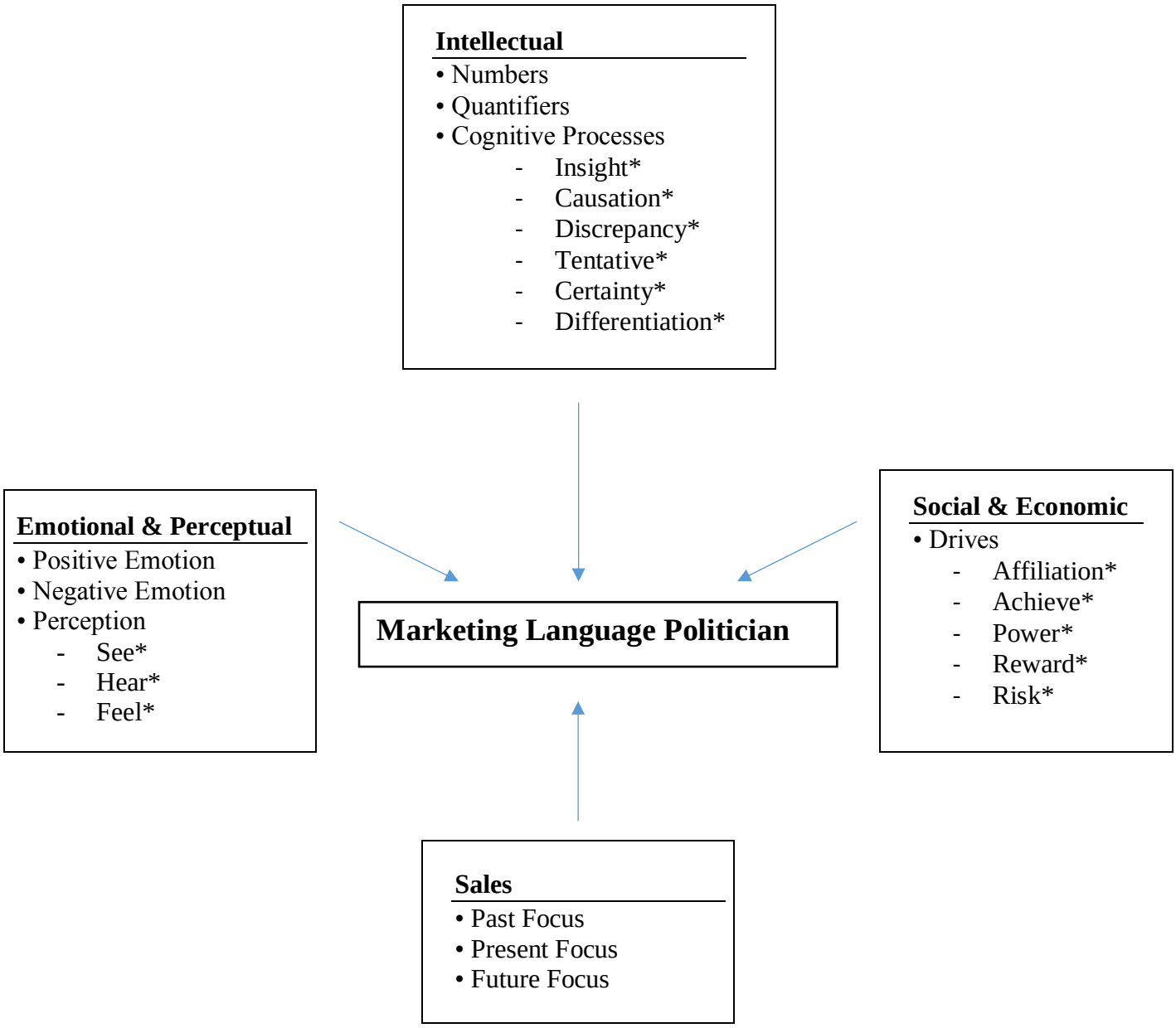
This is represented by the statistical hypothesis:

(3^s) $H_0: \hat{p}_{x,i,t} = \hat{p}_{x,i,t+n}$

$H_a: \hat{p}_{x,i,t} \neq \hat{p}_{x,i,t+n}$

- Here, $\hat{p}_{x,i,t}$ and $\hat{p}_{x,i,t+n}$ represent the average proportion that a candidate uses dictionary subcategory i at time t and at time t + n. The hypotheses will be tested across candidates and dictionary categories.

In conjunction with the hypotheses, a conceptual model is developed to illustrate how the different variables influence and form political marketing language.



The model showcases different forms of language for which the hypotheses will be tested. Namely, intellectual language, emotional, perceptual, social, economic and sales language. Furthermore, the different types of language are formed by allocating the variables to them that correspond to them.

With these hypotheses, the problem statement will be solved and the main research question will be answered.

Methodology

The research performed in this study is of descriptive nature. More specifically, via quantitative analysis differences in the use of language of several politicians will be studied. The purpose of this is to identify certain relationships between different word use and a politician's marketing. In line with the type of research, observations will be collected in a large data set showcasing the type of language candidates use during speaking events. Accordingly, the data collection is done with a cross-section observation. The cross-section measures differences across individuals or groups. This is combination with the observations results in witnessing differences in use of language. The form of the repeated cross-section, which is applicable to this research, means analysing differences over time. This will enable to make statistical statements about groups and individuals over time (Wooldridge, 2010).

The actual research conducted will consist of collecting speeches made by Hillary Clinton, Bernie Sanders, Ted Cruz and Donald Trump during the current presidential elections. The selection for these four candidates is based on how well they did in the polls and their different political views and backgrounds. In total, thirty-six speeches will be collected and transcribed from different speaking events throughout the presidential race. These events range from debates to party conventions to campaign rallies. Moreover, the transcriptions also include speeches they made when they either won a state or lost it when attempting to secure their party's nomination. This is extended with providing the transcriptions a place and date that marks the nature and time of the speaking event. Adding to that, the starting point for the data collection is each individual's announcement speech for candidacy. After this, the data is collected with predominantly one or two-month interval periods.

Furthermore, the speeches collected show similarities that stem from either the nature of the event from which they are transcribed or the topics that are discussed within these speeches. This is done to keep the internal reliability and consistency of the data as high as possible and reduce the number of outliers for statistical purposes. For example, the transcription of the New Hampshire Democratic party state convention on the 19th of September shows both Bernie Sanders and Hillary Clinton speaking at that particular event. Likewise, the transcription from the very first GOP 2016 presidential debate held on August 7th shows both Ted Cruz and Donald Trump speak at that event as part of the top ten Republican candidates. Consequently, by covering different events with different audiences and different topics that

each of them discuss, other factors that can contribute to their use of language are reduced to a minimum. The actual sources for these speeches are various media outlets and YouTube channels from which these speeches are transcribed. These transcriptions are then sampled together and used as input for the LIWC software. With the software, the output is then categorised according to different dictionary categories. The number of words collected will allow one to statistically determine if trends are present. Furthermore, the interpretation of the data will aid in either confirming or rejecting the statistic hypotheses.

As the Linguistic Inquiry and Word Count (LIWC) software forms the basis for the statistical analysis to be done with SPSS, it deems further clarification in how it operates. The LIWC software was developed with the intent of providing an effective and efficient method to study various emotional, cognitive and structural components present in an individual's written and verbal speech. The first application was developed as part of an exploratory study in language and disclosure. Through the years it has been further developed with the most recent version being LIWC2015, which has significantly altered both the dictionary and software options. The latest version incorporates new and more dictionaries that allows for the user to explore the use of language in multiple ways. Words that are read and analysed by the software are referred to as target words. Subsequently, words in the software's dictionary file are named dictionary words. Groups of dictionary words that tap a particular domain, for example negative emotion, are referred to as dictionaries subcategories (Pennebaker et al., 2015). Furthermore, the LIWC2015 dictionary forms the basis for text analysis. The dictionary is composed of nearly 6,400 words, word stems and select emoticons. It is noteworthy to state that a single word can be part of multiple categories. For instance, if the target word is "cried" then it will be scored in five categories (e.g. sadness, negative emotion, overall affect, verbs and past focus). Moreover, the output generated by the software lists 93 categories and sub categories. Except for the total word count, words per sentences and the four summary variables analytic, clout, authentic and tone, all means are expressed as percentages of the total word count in the given sample (Pennebaker et al., 2015). With this in mind, the analysis done with SPSS will consist of running several t-tests based on the different categories that are listed in the output.

The reliability and validity of this research along with the data collection should be a primary concern. Firstly, assessing the reliability and validity of text analysis is quite complex. Whereas sales data and questionnaires often provide consistent coefficients, the measurement

of use of language isn't so straightforward. Basic assumptions suggest that the more coefficients correlate, the more reliable and internally consistent the study is. However, when measuring use of language, it should be questioned how the observations are produced. Repeating the same words over and over again would aid in justifying statistical reliable numbers, however it is a phenomenon that seldom occurs. It should therefore be accepted that language reliability coefficients are typically lower than other forms of studies and data variables. Secondly, the validity proves to be complex when classifying words into categories with a corresponding scale. In order to validate the LIWC's dictionary scales and output, studies were conducted that measured the LIWC output with objective judges that evaluate the emotional content of a large number of papers. With these results, the Pearson correlation analysis was performed to determine the external validity. The findings suggest that the LIWC software successfully measures emotions, cognitive elements, thematic content and various language compositions. The level of agreement between the judges and the vast amount of studies that have found the LIWC categories to be valid for dozens of psychological domains support the external validity (Pennebaker et al., 2015).

Furthermore, in order not to get sidetracked, the validity of the entire research needs to be confirmed. By narrowing the research down to just four politicians and their use of language, the opportunity will be given to see certain relationships between these politicians and their use of language. Moreover, it will make it easier to validate that the correct use of language is measured by limiting the focus to the dictionary categories provided by the LIWC software. Next to that, it should be noted that there has to be a validation that the transcriptions correlate with the data output. As use of language is at the centre of this research, the transcriptions need to match the politician's verbal expressions. Hereby human error needs to be taken into account for transcribing the speeches. However, with the vast amount of speeches collected, as much human error and confounding effects as possible are cancelled out. Following this, the measurement and interpretation of the data needs to be done so correctly. Performing analysis with the aid of SPSS, the results produced should be considered valid and reliable. Thus the selected type of research, the manner in which the data is collected, the output of the LIWC software and the analysis with SPSS, ensures that the main research question can be answered.

T-test Analysis

SPSS will assist in testing the hypotheses but it will also rule out variables that show no significant differences between the politicians' respective use of language. Most commonly this will be done with t-tests that test the significance of the mean difference between proportions of dictionary category use. More specifically, the exact nature of the t-test depends on what is tested in each hypothesis. Choosing the right t-test will determine if the average proportion of words used in different categories significantly differs from each other for said politicians. The hypotheses that are stated are all two-sided. This will be taken into account when computing P-values and critical values. Furthermore, as hypothesis (1) and (2) compare the means between different groups and individuals, an independent samples t-test will be used. We are comparing different sample sizes and do not assume the variances to be equal. Hence, the Welch's t-test will be used. The t-statistic on the coefficient will be computed as:

(1)

$$t_{x,y,i} = \frac{\hat{p}_{x,i} - \hat{p}_{y,i}}{\sqrt{\frac{S_{X,i}^2}{N_X} + \frac{S_{Y,i}^2}{N_Y}}}$$

In this equation $t_{x,y,i}$ is the test-statistic for the difference in proportion of use in language of dictionary subcategory i between candidates x and y. $\hat{p}_{x,i}$ Is the proportion of use of subcategory i for candidate x. $s_{x,i}^2$ is the sample variance of the proportion of use of subcategory i for candidate x. N_x is the word count for candidate x. In equation (1), the denominator is based on an unpooled variance estimate. The test-statistic is distributed t (v). The degrees of freedom corresponding with the variance is estimated using the following equation:

(2)

$$v \approx \frac{\left(\frac{S_{X,i}^2}{N_X} + \frac{S_{Y,i}^2}{N_Y}\right)^2}{\frac{S_{X,i}^4}{N_X^2(N_x - 1)} + \frac{S_{Y,i}^4}{N_y^2(N_y - 1)}}$$

For hypothesis 2, where the differences in the use of language between respectively democrats and republicans is measured, the same equation is applicable. Hereby, x represents the democrats and y the republicans. However, testing the third hypothesis requires a slightly

different test-statistic as it compares the means for a single politician at two different intervals. Namely, their use of language at the beginning of their campaign compared to the end. One could therefore expect that a paired samples t-test should be used. However, in coherence to the differences in sample size, the assumed differences in variance and the origin from which the sample stems, a Welch t-test is used. In this case the test-statistic is given by the following equation:

(3)

$$t_{x,i,t} = \frac{\hat{p}_{x,i,t} - \hat{p}_{x,i,t+n}}{\sqrt{\frac{S^2_{X,i,t}}{N_{X,t}} + \frac{S^2_{x,i,t+n}}{N_{x,t+n}}}}$$

In equation (3) $t_{x,i,t}$ is the test-statistic for the difference in proportion of use in language of dictionary subcategory i for candidate x at time t and t+n. $\hat{p}_{x,i,t}$ Is the proportion of use of subcategory i for candidate x at time t. $\hat{p}_{x,i,t+n}$ Is the proportion of use of subcategory i for candidate x at time t+n. Hereby, t indicates the data retrieved from each politician's announcement speech and t+n refers to the data retrieved from their last speech collected. $s^2_{x,i,t}$ is the sample variance of the proportion of use of subcategory i for candidate x. $N_{x,t}$ is the word count for candidate x from their first speech. Similar to equation (1), the denominator is based on an unpooled variance estimate and the test-statistic is distributed t (v). The degrees of freedom corresponding with the variance is estimated using the following equation:

(4)

$$v \approx \frac{\left(\frac{S^2_{X,i,t}}{N_{X,t}} + \frac{S^2_{x,i,t+n}}{N_{x,t+n}}\right)^2}{\frac{S^4_{X,i,t}}{N^2_{X,t}(N_{x,t}-1)} + \frac{S^4_{x,i,t+n}}{N^2_{x,t+n}(N_{x,t+n}-1)}}$$

These equations will statistically determine whether to confirm or reject the hypotheses for different dictionary subcategories.

Data

The data collected shows a total of thirty-six transcriptions, collected from four different politicians over an approximate nine-month time span. The sample linked to the research exists of Hillary Clinton, Bernie Sanders, Ted Cruz and Donald Trump. At the time of writing, they are fiercely campaigning to secure their party's nomination for presidential candidate. During their political campaign, it is of vital importance that they distinguish and differentiate themselves from each other in order to gain as much support as possible. Their use of language therein seems like the ideal tool to do so. Furthermore, the transcriptions provide us with a total of 143167 words to be further analysed with the LIWC software and SPSS. As the software recommends a minimum of fifty words to provide any reliable and significant results, the sample in the dataset has well covered this. This in conjunction with the aforementioned politicians shows that the sample in the dataset matches the ideal group to retrieve information from in order to analyse the differences in use of language and form new marketing strategies.

Furthermore, the dataset from the LIWC output depicts several categories along with corresponding subcategories by which the data is categorised and ordered. On top of that, there are the four main categories Linguistic Processes, Other Grammar, Psychological Processes and Punctuation. These categories are further defined with their set of corresponding subcategories. The categories belonging to linguistic processes are:
Word Count; Analytic; Clout; Authentic; Tone; Words/sentence; Words>6 letters; Dictionary words; Total function words; Total pronouns; Personal Pronouns; 1st person singular; 1st person plural; 2nd person; 3rd person singular; 3rd person plural; Impersonal pronouns; Articles; Propositions; Auxiliary verbs; Common adverbs; Conjunctions and Negations.

Other grammar is made up from the following categories:
Common verbs, Common adjectives, Comparisons, Interrogatives, Numbers and Quantifiers.

Psychological processes consists of the following categories:
Affective processes, Positive emotion, Negative emotion, Anxiety, Anger, Sadness, Social processes, Family, Friends, Female references, Male references, Cognitive processes, Insight, Causation, Discrepancy, Tentative, Certainty, Differentiation, Perceptual processes,

See, Hear, Feel, Biological processes, Body, Health, Sexual. Ingestion, Drives, Affiliation, Achievement, Power, Reward, Risk, Past focus, Present focus, Future focus, Relativity, Motion, Space, Time, Work, Leisure, Home, Money, Religion, Death, Informal language, Swear words, Netspeak, Assent, Nonfluencies and Fillers.

Lastly, punctuation contains the following categories:
Total punctuation, Periods, Commas, Colons, Semicolons, Question marks, Exclamation marks, Dashes, Quotation marks, Apostrophes, Parentheses and Other punctuation.

The aforementioned categories vary slightly in how they are measured. Overall most LIWC output variables are measured in percentages of the total words within a text. However, the word count, words per sentence and the summary variables differ in how they are measured. The word count is measured in absolute numbers, the words per sentence is the mean number of words within each sentence and the summary variables are algorithms based on various LIWC variables from previous language research. The numbers corresponding to them are standardized scores converted to percentiles ranging from zero to hundred based on the area under a normal curve. With this said, most of the other categories are quite straightforward. However, the following categories are exemplified as they deem further clarification and are expected to be used most frequently in the analysis. (Pennebaker et al., 2015).

Numbers: refers to actual numbers and words that indicate a specific amount.

Quantifiers: is a type of determiner that indicates a certain quantity without revealing a specific amount.

Cognitive processes: Activities that affect the mental contents of a person such as thinking or remembering. This category is further diversified with *Insight, Causation, Discrepancy, Tentative, Certainty* and *Differentiation.*
Hereby, the increase in usage of variables pertaining to cognitive processes hints to the politician showcasing a better understanding of cognitive thinking and thus being more intelligent. Thus these variables indicate a certain level of IQ. They are expected to show a certain contrast wherein using numbers is more factual based and therefore shows more hints of knowledge whereas using quantifiers is more ambiguous and leaves room open for suggestion. Commonly during elections, they are part of factual or statistical statements so it

is expected to see variances between the candidates that either favour to include more data in their speeches or those that don't.

The category psychological processes with its set of subcategories will predominantly be used to measure the differences in the use of language. Within this category, the following variables are used and are therefore further exemplified.

Positive emotion; this variable lists the words that show the experience of feeling or an emotion in a positive way.

Negative emotion: lists the words that show the experience of feeling or an emotion but in a negative way.
With these variables it can be assumed that using more positive emotion hints to predominantly selling yourself in a positive way and spreading a more positive message. However, it is noteworthy to state that a higher usage of negative emotion could mean that a candidate tries to sell him or herself by making the others look bad. Furthermore, a negative emotion usually has a more powerful effect than a positive one.

Perceptual processes: Words that indicate the way people take in information from different stimuli. Logically, this category is expanded with *See, Hear* and *Feel*.
It is assumed women use more perceptual language which often correlates to more emotional language.

Drives: This category includes words related to different motivations that drives people. The drives are further diversified into *Affiliation, Achievement, Power, Reward* and *Risk.* Each has their own set of words corresponding to them where it is expected to see significant differences between the politicians and the different drives they predominantly communicate. Hereby a distinction is made between affiliation, achievement and power with reward and risk. Affiliation, achievement and power imply a need for achievement, power and a sense of belonging. With reference to the golden circle this is aimed at the most inner circle, "why?", that questions why a consumer has a particular need (Sinek, 2010). Reward and risk are more economic needs that often showcase a conscience form of bias (Farmer & Foley, 2009).

Time orientation: This category is split up in three different variables. *Past focus, Present focus* and *Future* focus. It measures how often politicians use time related words and in what sense.
Moreover, this category relates to the sales capacity of a politician. For instance, a politian can remark how the situation in the past wasn't good. Currently it hasn't improved but with his or her help the future will look much better. Using more time related words indicates how a politician can sell himself that he or she will improve the future.

Considering the LIWC software gives a wide range of 93 variables that can be used for further analysis, this sheer amount will most likely produce significant results for several categories. For research and statistical purposes, the lens through which these results are viewed needs to be narrowed down. This will allow for the research to more accurately and precisely determine the differences in use of language for the different politicians. The variables are chosen on the basis of several assumptions and factors corresponding to them. Most importantly, they can be grouped together to define certain characteristics in the use of language. Numbers, quantifiers and the subcategories belonging to cognitive processes define a certain level of IQ. Positive and negative emotion clearly defines the emotional context in which words are used. See, hear and feel define the use of perceptual language. The subcategories belonging to drives are split up into two categories; social needs and economic needs. Social needs pertain to affiliation, achievement and power whilst risk and reward belong to economic needs. Finally, the categories referring to time orientation will be used to define sales language.

The following table is an overview of the data collected from the LIWC output that shows the variables deemed most important to put into perspective the differences in use of language. Variables that are marked with * represent subcategories. The statistics in the table represent the means and standard deviations in percentages of several explanatory variables and could already hindsight expected results for different categories.

Overview Data

Variable	Mean	Standard deviation
Numbers	2,07	,60
Quantifiers	2,45	,55
Positive Emotion	3,71	,73
Negative Emotion	1,80	,70
Cognitive Processes	10,41	1,76
Insight*	1,83	,58
Causation*	1,56	,36
Discrepancy*	1,63	,41
Tentative*	1,76	,41
Certainty*	1,88	,45
Differentiation*	2,92	,69
Perception	1,74	,48
See*	,63	,25
Hear*	,81	,40
Feel*	,22	,13
Drives	11,93	1,97
Affiliation*	4,42	1,43
Achieve*	1,94	,64
Power*	4,54	1,02
Reward*	1,56	,58
Risk*	,73	,31
Past Focus	3,00	,83
Present Focus	12,18	2,95
Future Focus	1,72	,66

Data compiled from 36 transcriptions that have been analysed with the LIWC software. Descriptive statistics that come from figure 1 from the appendix describing the data rounded off.

Results

T-test Analysis

With a clear set of data, the hypotheses can now be tested. The first hypothesis "Candidates market themselves" is analysed with an independent t-test on the selected dictionary categories. The candidates in this hypothesis are represented by Hillary Clinton and Donald Trump. The independent t-test is used to determine if the average use of language per variable significantly differs between the candidates. These variables and subsequent sub variables are used to either confirm or reject the hypothesis, that states candidates market themselves, when looking at their use of language that correlates with IQ, emotions, perception, drives and sales. Furthermore, by comparing the two, we can determine on the

basis of the means per category which candidate prefers what type of language over the other. Table 1a shows the results from an independent t-test analysis for the categories related to IQ.

Table 1a

Variable	Mean (Hillary Clinton)	Mean (Donald Trump)	P-value (2-tailed)
Numbers	,013	,022	,000
Quantifiers	,027	,021	,000
Cognitive Processes	,115	,122	,001
Insight*	,020	,024	,000
Causation*	,018	,017	,270
Discrepancy*	,018	,018	,931
Tentative*	,019	,020	,441
Certainty*	,020	,019	,358
Differentiation*	,031	,034	,005

Data compiled from figures 2 and 3 from the appendix showing an independent t-test analysis with the selected variables for Hillary Clinton and Donald Trump.

The P-value for numbers and quantifiers shows that the average use in language for these categories significantly differs between Hillary Clinton and Donald Trump. Moreover, the mean values show a contrast where Trump favours to incorporate more numbers in his speeches whilst Hillary does so more with quantifiers. This insinuates that Trump has more accurate knowledge compared to Hillary on certain topics. However, that insinuation is on the basis that it is assumed that the numbers and statements that he makes are factually correct. The media has already pointed out this isn't always the case (Sharockman, 2016).

Furthermore, the P-value for cognitive processes shows there is a significant difference between Hillary Clinton and Donald Trump when it comes to cognition. The mean value suggests Trump shows more cognitive thinking when he speaks than Hillary. But, if we look more closely at the sub variables for cognitive processes, we see that in most cases the P-value indicates there to be no significant difference between Hillary and Trump. Next to that, when considering the small differences in mean values between the candidates, it points out that there is hardly any difference at all between them for cognitive processes. Thus, overall the null hypothesis is rejected that there is no significant difference between Clinton and Trump based on the variables pertaining to IQ. However, considering the mean values, there is hardly any difference between them and it is safe to assume they hardly vary in IQ compared to each other. This is in agreement with previous news articles that have already pointed out both Hillary and Trump were well educated and went to prestigious schools

(Strauss, 2015). Most notably is the difference in usage of numbers and quantifiers that favours Trump. However, for the analysis, the numbers Trump mentions needn't be factually correct to suggest he has more accurate knowledge than Hillary. It does however show that he likes to incorporate more numbers in his speech to sell himself as a clever and successful man. On a side note, although the linguistic variables do not necessarily point out major differences between Hillary and Trump for IQ, Trump does however repeatedly states "I'm, like, really a smart person." (Carroll, 2015). This goes to show that Trump is actively marketing himself as an intellectual candidate, especially in relation to his socioeconomic status and academic achievements (Sirin, 2005).

Table 1b shows the results from an independent t-test for the variables pertaining to emotion and perceptual language.

Table 1b

Variable	Mean (Hillary Clinton)	Mean (Donald Trump)	P-value (2-tailed)
Positive Emotion	,039	,039	,671
Negative Emotion	,018	,017	,337
Perception	,017	,022	,000
See*	,007	,007	,390
Hear*	,006	,013	,000
Feel*	,003	,002	,000

Data compiled from figures 4 and 5 from the appendix showing an independent t-test analysis with the selected variables for Hillary Clinton and Donald Trump.

When looking at the P-values for both positive and negative emotion, it can be determined that there is no significant difference between Trump and Hillary for emotional language. In this case, the null hypothesis is confirmed. The mean values show that both have a tendency to use more positive emotion than negative, but on average there is no significant difference between. So neither Trump or Hillary differentiates him or herself from the other on an emotional level. However, when examining the results for perceptual language, the P-value depicts a clear significant difference between them. In this case, the null hypothesis is rejected. Whilst we expected Hillary to use more perceptual language, it is in fact Trump who does so. Based on the mean values, Trump clearly uses more perceptual language related to hearing whilst Hillary slightly uses more related to feeling. A linguistic interpretation of this could be that Trump tries to sell himself by often stating that he listens and hears what the public wants whilst Hillary focusses more on what the public feels.

Table 1c depicts the results for the category drives.

Table 1c

Variable	Mean (Hillary Clinton)	Mean (Donald Trump)	P-value (2-tailed)
Drives	,134	,102	,000
Affiliation*	,050	,041	,000
Achieve*	,028	,013	,000
Power*	,045	,032	,000
Reward*	,018	,022	,000
Risk*	,009	,006	,000

Data compiled from figures 6 and 7 from the appendix showing an independent t-test analysis with the selected variables for Hillary Clinton and Donald Trump.

The results clearly indicate significant differences between Trump and Hillary for not only drives, but all sub variables. Thus the null hypothesis that states there to be no difference between Trump and Hillary for this category is completely rejected. Looking closely at the results, the mean values for almost every category indicate Hillary focusses more on drives. For affiliation, achieve and power Hillary certainly addresses more people's needs and desires. Although she has changed her slogan multiple times, her current one "Stronger Together" adds to this type of language. The need to belong to a group, that together you can achieve more and the results of this can be more power support this message. Hillary therefore seems to be more focussed on selling herself as being part of the solution, but not being the solution in itself.

The variable reward however clearly shows Trump using more language related to benefits. This is also in response to his slogan "Make America Great Again". It suggests that one of his major selling points is that he is the one who can make this country great again which in return will benefit everyone. Be it from an economic standpoint, or immigration, or security or any other topic, he will make sure everyone benefits (Bennett, 2015). Perhaps in response to that, Hillary incorporates more risk related language in her manner of speaking. Playing into the public's fear is not so uncommon in politics (Browser, 2011). However, where this is commonly done so from an economic, social or security standpoint, it seems as if the risk card is now used more to discredit the other candidate personally and not just their standpoints. Thus an economic interpretation would be that Trump focusses more on the economic gains for this country, whilst Hillary focusses more on the economic risks this

country faces. This shows that they have very conflicting views on economic issues (Boak, 2016).

Table 1d lists the results for sales related language.

Table 1d

Variable	Mean (Hillary Clinton)	Mean (Donald Trump)	P-value (2-tailed)
Past Focus	,031	,039	,000
Present Focus	,121	,164	,000
Future Focus	,014	,023	,000

Data compiled from figures 8 and 9 from the appendix showing an independent t-test analysis with the selected variables for Hillary Clinton and Donald Trump.

The P-values for all variables show there to be a significant difference between Hillary Clinton and Donald Trump. In this case, the null hypothesis is rejected that states there is no difference between Hillary and Trump for sales language. The mean values show that in all cases Trump incorporates more language related to the past, present and future. A marketing approach of this in response to his slogan "Make America Great Again" could suggest that the past, possibly when there was a Republican president, was better than the present situation. Moreover, that with him as a president, the future will be prosperous again. In other terms, he creates a need that is focussed on the current climate, how this need was addressed in the past and how he will fix it in the future. It needs to be stated that the results don't show Hillary not using sales language, just that Trump does so more. Perhaps this is explained due to his business background, which could make him more equipped and experienced in doing so. Next to that, studies have already found Trump's language isn't grammatically correct, but he manipulates his sentences in such a way as to end with a strong, punchy word (Shontell, 2016). So, whilst Hillary is focussed more on the public's needs, Trump likes to sell the public better times a manages to do so by even altering his sentence structures.

The second hypothesis "Democratic candidates market themselves differently from Republicans by using different dictionary categories in their speech" questions how Democrats and Republicans differ from each other in political marketing terms. Whilst the first hypothesis only had data from Trump and Hillary, the second hypothesis increases the data set by incorporating speeches from Bernie Sanders and Ted Cruz. Most notably, the results will be used to accurately determine the differences between Democrats and

Republicans, but it will also allow us to see if the outcomes vary from the first hypothesis by increasing the population and data set.

Table 2a shows the results from an independent t-test analysis for Democrats and Republicans with the variables related to IQ.

Table 2a

Variable	Mean (Democrats)	Mean (Republicans)	P-value (2-tailed)
Numbers	,019	,023	,000
Quantifiers	,027	,022	,000
Cognitive Processes	,110	,103	,000
Insight*	,019	,020	,068
Causation*	,017	,016	,401
Discrepancy*	,018	,016	,038
Tentative*	,018	,018	,523
Certainty*	,019	,017	,044
Differentiation*	,032	,027	,000

Data compiled from figures 10 and 11 from the appendix showing an independent t-test analysis with the selected variables for the Democrats and Republicans.

The P-values for both numbers and quantifiers shows there is a significant difference between Democrats and Republicans in their use of language for this category. Moreover, the mean values show that Republicans like to use more numbers in their speeches whereas Democrats show a higher use of quantifiers. Furthermore, the results show that the differences in use of language between Democrats and Republicans is smaller in comparison to merely analysing Hillary Clinton and Donald Trump. Although the difference between quantifiers has stayed relatively the same, the mean difference for using numbers has narrowed. A linguistic impression of this is that within the Democratic party there are some differences between the party members in their preferences to incorporate numbers in their speeches. In contrast to that, the Republican party stays quite consistent in the amount of numbers and quantifiers they incorporate in their speeches.

Furthermore, the P-value for cognitive processes shows there is a significant difference between Democrats and Republicans. However, in contrast to just comparing Hillary and Trump, the mean values in this case illustrate Democrats overall show more cognitive thinking in their use of language than Republicans. Yet, almost all sub variables for cognitive processes show no significant difference except for differentiation. In linguistic terms, this

means that Democrats show a higher level in flexibility for word usage compared to Republicans. Recent research that has tracked and analysed Facebook messages politicians and their supporters from both parties have sent confirm that Democrats have a larger vocabulary and better grammar than Republicans (Brekhus, 2015). Thus, when it comes to the variables contributing in determining the differences in IQ between Democrats and Republicans, overall the null hypothesis is rejected that suggests there is no significant difference between them. However, considering the insignificant differences for most sub variables and the small mean differences, it is fair to suggest that on an intellectual level there is hardly any difference between Democrats and Republicans. The major differences between them stem from their use of numbers and quantifiers. Whereas it is presumed that incorporating more numbers in your use of language shows a more accurate understanding and knowledge of the current political climate, for the analysis and results these numbers needn't be factually correct. They do prove however that Republicans like to use more numbers in their speeches which translates to selling yourself to the public as a person that knows exactly what is going on when discussing important issues. Next to that, Democrats often target the middle and lower class whereas the Republican's primary target group consists of the upper class. Thus, the higher frequency of numbers could also originate from the numbers Republicans have accumulated in the form of economic wealth which they use to sell themselves. A marketing approach to this could be that Republicans advertise they have more wealth, so if you vote for a Republican, surely you will get richer. Although the logics behind that marketing approach should be questioned, it can however sell the idea Republicans are better for your wallet than Democrats.

Table 2b provides the results from an independent t-test for the variables corresponding to emotion and perceptual language.

Table 2b

Variable	Mean (Democrats)	Mean (Republicans)	P-value (2-tailed)
Positive Emotion	,036	,036	,876
Negative Emotion	,019	,017	,149
Perception	,016	,020	,000
See*	,006	,006	,820
Hear*	,006	,010	,000
Feel*	,003	,002	,000

Data compiled from figures 12 and 13 from the appendix showing an independent t-test analysis with the selected variables for the Democrats and Republicans.

From these results it can be determined that there is no significant difference between Democrats and Republicans for the amount of emotion that is used in their use of language. Both the variables positive and negative emotion with their corresponding P-values indicate no significant difference between Democrats and Republicans. The mean values do highlight that both parties use more positive emotion than negative emotion, but the differences between the parties is insignificant. Various news outlets pointed out that Democrats often speak from the head and Republicans try to inspire from the heart using more emotion (Shapiro & Newby, 2015). However, when looking at the linguistic characteristics for emotion, the null hypothesis is confirmed that suggests there is no significant difference between Republicans and Democrats for this category. However, when it comes to incorporating perception in your use of language, there is a significant difference between the democratic and republican party. Overall, the mean values show that Republicans use more language that is related to a person's perception than Democrats do. But the sub variables show that this is mostly due to the variable hearing. Furthermore, the results are quite similar to the case of only comparing Hillary and Trump. Therefore, a marketing approach to this could be that Republicans sell the idea of better hearing and listening to the public's needs whilst Democrats have a better understanding of what the public feels.

Table 2c displays the results for the variables corresponding to needs and economic factors.

Table 2c

Variable	Mean (Democrats)	Mean (Republicans)	P-value (2-tailed)
Drives	,127	,105	,000
Affiliation*	,046	,037	,000
Achieve*	,024	,015	,000
Power*	,047	,042	,000
Reward*	,015	,016	,096
Risk*	,008	,007	,016

Data compiled from figures 14 and 15 from the appendix showing an independent t-test analysis with the selected variables for the Democrats and Republicans.

From table 2c it can be determined that there is a significant difference between Democrats and Republicans for the variables corresponding to needs. The null hypothesis that resembles no significant difference between Democrats and Republicans for this category is therefore clearly rejected. This conclusion is supported by the corresponding P-values for affiliation, achieve and power. However, when looking at more economical related language in the form

of the reward and risk variables, the null hypothesis is only partially rejected. The P-value of ,096 is higher than the significance level of 5% which in return confirms the hypothesis that there is no significant difference for economic factors in the use of language for Democrats and Republicans. When it comes to risk, there is however a significant difference. But due to the small mean difference between the parties, it should be questioned to what extent the null hypothesis can either be rejected or confirmed.

Whilst it was expected that the Republicans would focus more on needs and economic factors, it is in fact the Democrats who do so. The results are in fact quite similar to the case of comparing Hillary and Trump. Democrats tend to generate a sense of belonging, that together it is possible to achieve more which leads to a better outcome for everyone and more power. It needs to be noted that use of language related to power can also be used to discredit the other party. For instance, suggesting that Republicans feel superior or act with the intent of only making themselves better. Thus the expectation that Republicans would focus more on family values, the nation, achievement and reward isn't supported by the results. It can however be determined that the differences between Democrats and Republicans is smaller than the differences between Hillary and Trump. This is reflected by the mean values. But overall, Democrats clearly prefer to address needs in their use of language whilst Republicans lack to differentiate themselves with economic factors.

Table 2d shows the results for Democrats and Republicans in the case of comparing sales related language.

Table 2d

Variable	Mean (Democrats)	Mean (Republicans)	P-value (2-tailed)
Past Focus	,027	,034	,000
Present Focus	,117	,131	,000
Future Focus	,013	,019	,000

Data compiled from figures 16 and 17 from the appendix showing an independent t-test analysis with the selected variables for the Democrats and Republicans.

The results clearly illustrate there is a significant difference between Democrats and Republicans in their use of sales language. The null hypothesis is therefore completely rejected. Moreover, the higher mean values for all three variables highlight that Republicans overall incorporate more sales language when speaking. As was expected, Republicans focus

more on the past. In linguistic terms, this corresponds to having a higher value for tradition and most likely referring more to better times. Next to that, they also address the present more. Most likely they aim to sell the idea that, under the current democratic president, things are worse compared to the past when there was a republican president. Following this, a logical marketing campaign would be that America will be better off in the future again with a republican president that knows the countries traditions. Furthermore, the results are quite similar to the case of comparing Hillary and Trump. The mean differences have narrowed but there is still a clear distinction between Democrats and Republicans when it comes to sales language. Overall, this leads to stating Republicans focus more on selling themselves by reflecting on better times whilst Democrats focus more on needs and how they should be addressed.

For the third and final hypothesis "a candidate's use of language changes throughout their political campaign" multiple speeches from a single candidate are analysed. The mean differences are calculated along with corresponding P-values to determine how much a candidate's use of language has significantly changed starting at the launch of their campaign and the progression towards the end of campaigning. The results are analysed and compared for Hillary Clinton and Donald Trump. One could say Trump's language would stay more consistent in response to his background in business whilst another might counter that with stating Hillary will remain more consistent in her use of language due to her years of experience in politics and political marketing. The results will show if a candidate changed his or how manner of speaking, how much so, and who has changed the most compared to the other.

Table 3a shows the results of Hillary Clinton from an independent t-test for all variables from the beginning and end of her campaign.

Table 3a

Variable	Mean Difference (Hillary Clinton first-last speech)	P-value (2-tailed)
Numbers	-,002	,519
Quantifiers	-,004	,463
Positive Emotion	-,002	,769
Negative Emotion	,005	,101
Cognitive Processes	,015	,091
Insight*	,003	,488

Causation*	,004	,203
Discrepancy*	,011	,000
Tentative*	,005	,162
Certainty*	-,008	,066
Differentiation*	,005	,334
Perception	-,007	,061
See*	-,001	,776
Hear*	-,006	,022
Feel*	-,001	,504
Drives	,002	,821
Affiliation*	-,008	,218
Achieve*	,015	,001
Power*	,009	,149
Reward*	,002	,549
Risk*	-,003	,226
Past Focus	-,008	,142
Present Focus	-,004	,639
Future Focus	,000	,908

Data compiled from figures 18 to 21 from the appendix showing an independent t-test analysis with the selected variables for Hillary Clinton.

From table 3a it can be determined that Hillary Clinton stays quite consistent in her use of language throughout her political campaign. Almost every P-value indicates there is no significant difference from how she talked at the beginning of her campaign and how she communicates towards the end. When reviewing the variables corresponding with IQ, it is clear that there is no significant difference between them. The mean differences show that Hillary did use more numbers and quantifiers towards to end of her campaign, however the difference remains insignificant. Furthermore, only the sub variable discrepancy shows any significant difference from how Hillary spoke at the beginning. Considering how cognitive processes and all the other sub variables show no significant difference, the null hypothesis is confirmed that leads to stating Hillary's use of language hasn't changed in how she expresses her IQ.

As for emotion and perception, the results show that a change in her use of language corresponding to these variables has remained to a minimum. Both the P-values for positive and negative emotion show no significant difference. The mean differences show that Hillary did use a little more positive emotion and less negative from how she spoke in the beginning, but the difference remains insignificant. Furthermore, looking at how she incorporates perception in her use of language, the differences remain fairly insignificant. Only the sub variable hear shows a significant difference from how Hillary incorporated words

corresponding to hearing and listening in the beginning compared to the end. Although most P-values show no significant difference, it is noteworthy to state that the mean differences illustrate that Hillary used more perception in her use of language at the end compared to the beginning. Considering there is only one sub variable that shows a significant difference, overall the null hypothesis is confirmed which shows that Hillary's use of language hasn't changed that much throughout her campaign for how she expresses emotion and perception. Similar to emotion and perception, Hillary's use of language for needs and economic factors hardly shifts. This is supported by the P-values which show no significant difference except for the sub variable achieve. The mean difference for achieve shows Hillary focussed more on the need for achievement in the beginning of her campaign than she does so in the end. Although there is no significant difference, the mean differences show this is the same case for the sub variables power and reward. Furthermore, she increased her use of affiliation and risk. A marketing impression of this could be that she started campaigning more with the message that the public can achieve more together. This is also encouraged by her final slogan "Stronger Together", which was changed and adjusted multiple times throughout campaign course. Next to that, although the difference for risk is insignificant, al small shift in adding a little extra risk to your campaign helps in marketing a worse outcome if Trump would become president. Thus, overall the null hypothesis for a change in her use of language for needs is partially rejected whereas the null hypothesis for economic factors is completely confirmed. This once again shows that Hillary's use of language remains quite consistent throughout her political campaign.

In conjunction with the other variables, the results for her sales language show Hillary remains quite consistent. From the P-values it can be determined that there is no significant difference from how Hillary incorporated sales language at the beginning of her campaign and how she does so in the end. Alongside this, the mean differences show Hillary did in fact focussed a bit more on the past and present compared to the beginning, but there is no shift at all in her focus on the future. From a marketing perspective, this could mean that she remains consistent in her message for how the future should look and for this, she starts using more examples from the past and present to support that message. So when it comes to sales language, the null hypothesis is confirmed that leads to stating Hillary's language hasn't changed when it comes to using sales language. Finally, when reviewing all the variables pertaining to IQ, emotion, perception, needs, economic factors and sales, the general null hypothesis that states Hillary's use of language shows no significant difference from how she

speaks at the beginning of her campaign and at the end is confirmed. Although there are minor significant differences, a general consensus would be that she is very consistent in how she sells herself during the presidential race.

Table 3b shows the results for Donald Trump's use of language from the beginning of his campaign and the end.

Table 3b

Variable	Mean Difference (Donald Trump first-last speech)	P-value (2-tailed)
Numbers	,010	,000
Quantifiers	,000	,961
Positive Emotion	-,007	,054
Negative Emotion	-,015	,000
Cognitive Processes	,008	,181
Insight*	,004	,157
Causation*	,002	,361
Discrepancy*	-,001	,622
Tentative*	,004	,077
Certainty*	-,014	,000
Differentiation*	,004	,223
Perception	,005	,034
See*	-,003	,023
Hear*	,011	,000
Feel*	-,003	,003
Drives	-,055	,000
Affiliation*	-,029	,000
Achieve*	,010	,000
Power*	-,024	,000
Reward*	,004	,199
Risk*	-,004	,048
Past Focus	,007	,047
Present Focus	,049	,000
Future Focus	,000	,895

Data compiled from figures 22 to 25 from the appendix showing an independent t-test analysis with the selected variables for Donald Trump.

From table 3b it can be determined that Trump's language changes throughout his campaign significantly for various variables. However, when looking specifically at the P-values corresponding to Trump's use of IQ in his use of language, there are only minor differences. Only the variable numbers and sub variable certainty shows a significant difference from how Trump incorporated numbers and has shown certainty in his speeches in the beginning of his campaign and in the end. More specifically, the mean differences show Trump started to use

less numbers in his speeches and more certainty as the campaign progresses. Yet, the main variable cognitive processes showed no significant difference. Therefore, the difference in certainty is deemed less instrumental in determining a change in his use of language showcasing his IQ. Thus, considering the P-values and mean differences for the variables pertaining to Trump's use of IQ in his use of language, the null hypothesis is only partially rejected. This leads to stating that there are some changes in how Trump expressed his IQ in the beginning of his campaign and how he does so in the end. However, the main reason for this change is the amount of numbers he uses in his speeches. Considering that amount has diminished, it shows Trump is starting to refrain from using exact figures and most likely focusses on other areas to sell himself as the ideal candidate. Thus when it comes to IQ, it is fair to state that overall Trump stays relatively consistent in how he presents himself as a knowledgeable man. The shift in using numbers could be explained by Trump learning and adjusting his language and focus as the campaign trail continues. This would support Tetlock's research that a politician starts to adjust his thinking as time progresses. Moreover, it is noteworthy to state that the mean differences show Trump making more changes in his use of language related to IQ than Hillary did so.

Following this, the results for emotion and perception show Trump making some significant changes in his use of language for these categories. Firstly, although the P-value of positive emotion is slightly higher than a significance level of 5%, the mean difference shows Trump starting to use more positive emotion in his speeches. Next to that, the P-value for negative emotion clearly demonstrates a significant difference in Trump's use of emotional language. The mean difference shows Trump using more negative emotion in his speeches at the end compared to the beginning. Overall, Trump used more emotional language in the end than he did so in the beginning. Furthermore, all the variables for perception show a significant difference. The mean differences show that overall Trump used less perception in his use of language as the campaign progresses. Yet, for the sub variables see and feel the opposite is in fact the case. The sub variable hear illustrates Trump using less language related to hearing and listening in the end of his campaign. The null hypothesis that states there is no significant difference in Trump's use of language for emotion and perception when comparing them at the start of his campaign and the end, is partially rejected for emotion and completely for perception. Moreover, whereas Hillary has only shown a minor difference for the sub variable hear, Trump clearly displays vast differences in how he incorporates emotion and perception in his speeches throughout his political campaign. Whereas Hillary started to use

more perceptual language related to hearing, Trump does the exact opposite. A marketing response to this would be that Trump adjusts the tone and feeling of his message throughout his campaign (Healy & Burns, 2016).

Alongside the results for emotion and perception, the results for the category drives demonstrate Trump also significantly changes his choice of words for this category. Almost every P-value indicates a significant difference from how Trump focussed on social and economic needs in the beginning and how he does so near the end of his campaign. The mean differences for affiliation and power show a greater focus on a sense of belonging and a need for power towards the end of his campaign. In contrast to that, Trump used less language that focussed on the need for achievement. Whilst Hillary only showed a significant difference for achieve, both her and Trump's use of language diminished for this category. Furthermore, although Hillary's change in use of language for risk was insignificant, the P-value for Trump shows a clear significant difference. Both her and his mean difference for this variable show they started to incorporate more risk related words. In marketing terms this could imply that Trump is campaigning more fiercely with a message that Hillary is a poorer choice for president than Hillary does so the other way around. For instance, Trump refers to Hillary as 'the devil' (Diamond, 2016). Overall, the null hypothesis that states Trump's use of language for needs remains the same throughout his campaign is completely rejected. In the case of economical language, the null hypothesis is only partially rejected as the P-value for reward shows an insignificant difference. Thus unlike Hillary, Trump changes his language considerably throughout his campaign for this category.

As for sales language, Trump again displays that his use of language changes considerably. The P-values for past and present focus illustrate a significant difference when comparing the beginning of his campaign and the end. However, similar to Hillary, there is no significant difference for future focus. Whilst the mean differences for Hillary showed that she started to focus more on the past and present, the opposite is in fact the case for Trump. Taking into account his slogan "Make America Great Again", it would have been fair to assume that his focus on better times and referring to those would increase as his campaign progresses. However, the mean differences signal that he did so more often in the beginning of his campaign than towards the end. Taking into account the P-values for past, present and future focus, the null hypothesis that states Trump's sales language remains the same throughout his political campaign is only partially rejected. It does however show his flexibility to change

his language more so than Hillary. When making a fair assessment of all the variables that make up IQ, emotion, perception, needs, economic factors and sales, the main null hypothesis that states Trump's use of language shows no significant difference from how he communicates at the beginning of his campaign and near the end is rejected. Where it was expected that Trump would stay consistent in his use of language due to his business background, the opposite is in fact the case. This demonstrates Trump's ability to adapt and learn as his political career progresses. The counter hypothesis that suggests Hillary would be more consistent in her use of language throughout her campaign is clearly proven to be correct. Most likely this is explained due to her years of experience in not only politics altogether, but also specifically elections and political campaigning. It is in fact remarkable how much Trump's language has changed throughout his campaign and how little this is the case for Hillary. Whilst Hillary has shown a strong and consistent focus on drives, Trump manages to shift and diversify his focus more for different categories throughout his campaign. Thus this implies Trump has learned more during his campaign and manages to put his new political knowledge to good use by changing his choice of words.

Having captured all these results, a clear image is portrayed that shows to what extent politicians market themselves in their use of language. The interpretation and understanding of the results will lead up to solving the problem statement.

Conclusion

When evaluating the problem statement, the data collected and the results from the analyses, several conclusions can be drawn that answer the main research question. The purpose of this research was to determine to what extent politicians market themselves in their use of language. In order to make it feasible to solve the problem statement, 36 speeches were collected and transcribed from four politicians. These transcriptions were then used as input for the LIWC software. The output of the LIWC software formed the basis of the dataset to determine quantitative differences between politicians regarding their use of language.

with respect to the three hypotheses, several conclusions can be drawn. The first hypothesis was used to analyse the differences between Hillary Clinton and Donald Trump regarding their use of language throughout their campaign. The linguistic variables made up several

categories on which the differences could be scored and judged to determine the differences in their choice of words. These categories were IQ, emotion, perception, social needs, economic needs and sales language. The analysis shows that Hillary and Trump hardly vary in how their IQ is shown through their choice of words. The results show there is no significant difference for language corresponding with cognitive thinking. With that being said, Trump does show a clear preference to incorporate more numbers in his speeches whilst Hillary uses more quantifiers. This leaves us with concluding that Hillary and Trump don't differ significantly in their IQ. This conclusion is supported by both their educational background. Hillary was accepted to Yale, which reportedly shows a very high IQ, well in excess of the 98th percentile and Trump was an Ivy League school graduate (Strauss, 2015). However, the research question was aimed to what extent politicians market themselves in their use of language. Although there is no difference in how their IQ is expressed, there are strong differences in how they inform and persuade the public with numbers and quantifiers. Trump has a tendency to do so more with numbers which speaks in favour of his socioeconomic status and academic achievement (Sirin, 2005).

As far as emotion and perception concerns, the results only partially show a significant difference. Both candidates use the same amount of positive and negative emotion in their use of language and therefore contradicts previous literature that suggested Hillary would do more so. However, they do differ significantly when it comes to perception. Whilst it was expected that Hillary would use more perception in her language, it is in fact Trump who does so. This is mostly explained by his use of words relating to hearing. This makes us conclude that Trump tries to market himself as a person that better knows and understands what the public concerns and demands are. This conclusion is supported by his business-like approach to politics that attempts to recognize and focus on people's needs (Bennett, 2015). However, when we review the results for needs, all variables show a significant difference that may point out something else. From the mean values it can be concluded that Hillary speaks more strongly about social needs. Furthermore, Trump focusses more on rewards whilst Hillary has a tendency to focus more on risks. Thus it is clear that they market themselves differently regarding the public needs. Hillary strongly addresses a sense for togetherness, achievement and power whilst Trump is more concerned with economic prosperity. This leads to very conflicting views on economic growth (Boak, 2016).

Additionally, in politics it is of vital importance that a candidate knows how to sell himself. The results for sales language clearly show Trump is a better salesman. All variables show a significant difference between him and Hillary. The mean values show Trump to be a strong favourite of who's a better salesperson. This is supported by previous studies that have analysed and concluded that Trump's speech patterns indicate he's a good salesperson. He keeps it simple and often directly talks to the audience and gives them instructions (Shontell, 2016). This in combination with the results leads to the conclusion that Trump is a better salesperson than Hillary. Following this, considering the vast differences between Hillary and Trump with regards to their use of language, overall it can be concluded that they market themselves extensively and differently from one and another.

The second hypothesis reveals the differences in word choice between Democrats and Republics. More to the point, from the results it can be concluded if Democrats and Republicans differ from each other in their use of language and how much so. The dataset was expanded by incorporating speeches from both Bernie Sanders and Ted Cruz. The results provide us with a general sense of how Democrats and Republicans choose their words carefully and differ from each other for each category. From the results it can be concluded that the differences in IQ remain to a minimum. Whereas the results for the first hypothesis illustrate Trump showing a little more cognitive thinking in his language, the mean value now is in favour of the Democrats. This supports previous literature that has shown that Democrats have a larger vocabulary and better grammar and can therefore express themselves better (Brekhus, 2015). Next to that, similar to only comparing Hillary and Trump, the results show Democrats using more quantifiers and Republicans more numbers in their speeches. However, the mean differences have grown smaller. Yet, from this it can be concluded that when Democrats inform the public, they have a higher flexibility in using language that leaves more room open to suggestion than Republicans do.

Next to that, the results for emotion and perception show Democrats and Republicans don't significantly differ from each other for influencing the public with either more positive or negative emotional language. This contradicts previous literature that suggested Republicans would incorporate more emotion in their language. In return, the results do display differences for perception. Republicans use more perception in their choice of words than Democrats do so. The major difference is explained by the variable hear. Therefore, it is concluded that Republicans like to market themselves more with having heard what the

public wants compared to Democrats. But, the results for needs show may portray a different image. All variables corresponding with social language show a significant difference between Democrats and Republicans. More specifically, Democrats strongly incorporate more social needs in their speeches than Republicans do so. From this it can be concluded that Democrats market themselves with a message that they have a better understanding of the public's social needs. This message is reinforced by their party's platform that calls for "Change that matters" and to "Stand with Democrats". Furthermore, this confirms previous expectations that suggested Democrats would target more the middle and lower class in their manner of speaking. On the other hand, the results for the variables that make up economical language hardly show any difference. From this it can be concluded that Democrats and Republicans speak in equal amounts about economic rewards and risks. This conclusion contradicts our expectations that suggested Republicans would do so more. Aside from this, it is noteworthy to state that both party's standpoints for economic needs can be completely different, but the results show they do not vary linguistically how they talk about these concerns.

Following this, the results for sales language tries to answer the question which party would be better in selling their ideas. From the results it can be concluded that Democrats and Republicans significantly differ from each other regarding sales language. Furthermore, it can be said that Republicans are theoretically better at selling themselves and their ideas than Democrats. However, this does not lead to the conclusion that they are better candidates or have better ideas. The results merely illustrate that linguistically they use more sales language to market themselves. Previous studies already pointed out Trump would be a good salesperson due to his vocabulary and grammar, but it seems that this is the case for Republicans in general. In retrospect to the results from the first hypothesis, the results for the second hypothesis show similar findings for almost every category. Therefore, it is concluded that Democrats and Republicans market themselves differently from each other in their use of language.

From the result of the final hypothesis the following can be concluded. Hillary stays very consistent in her use of language throughout the course of her campaign. Although some variables show a minor significant difference, in almost every case the null hypothesis was confirmed. From a marketing perspective it therefore seems that she keeps selling herself linguistically in the same manner. Along with this she has over thirty years of experience in

politics and campaigning, so the counter hypothesis that Hillary would stay consistent in her use of language seems correct. Moreover, this also supports Schultz's research that consistency is vital when branding yourself. However, when reflecting upon the results for Trump, a different conclusion can be drawn. Whereas Hillary's language has stayed quite consistent, Trump's language has changed considerably throughout the course of his campaign. The major changes in his language are related to perception, drives and sales language. Almost every variable shows a significant difference. Moreover, it can be concluded that Trump starts speaking more about needs as his campaign progresses. From this it is concluded that Trump's has learned the most during his political campaign. Where some might have thought that his run for presidency was a stunt, his language and the changes he made prove the opposite. This is in line with Tetlock's research that pointed out that when politicians face complex issues and huge responsibilities, their manner of thinking changes over time. As the linguistic differences with other politicians grow smaller, it can be said that not only does Trump market himself as a businessman, he can now be considered an experienced politician by his use of language. Therefore, it is concluded that a politician's language may or may not change during a political campaign. However, they have the ability to do so to position and market themselves differently if need be.

To relate all these findings to the main research question and answer it, the following can be concluded. Politicians, and in particular Hillary Clinton and Donald Trump, vary considerably in their use of language to market themselves. The disparities differ per category. In general, politicians don't market themselves differently in relation to their IQ. However, they clearly choose to promote themselves differently by either incorporating more numbers or quantifiers in their speeches. Furthermore, although research has pointed out that the public may feel that some politicians spread more fear in their marketing, in linguistic terms they do so equally. Politicians do market themselves differently when it comes to perception. One may choose to repeat that he hears what the public wants whilst the other chooses his words differently that he feels what the public wants. What the stronger message is from a marketing perspective is open to debate. Next to that, characterizing why the public has certain needs highlights major differences in their use of language. This can be used to market a politician differently from his competitors by addressing certain needs more prominently than others. This allows for politicians to target groups more specifically that have certain needs. Next to that, politicians need to sell themselves. Based on the results, it is concluded that Trump and Republicans in general are better salespeople. This doesn't mean

Hillary and Democrats are bad salespeople, just that Trump and Republicans linguistically sell their ideas more prominently. Finally, politicians can stay consistent in their use of language or change remarkably during the elections. This does prove to a certain degree that they choose their words carefully in order to market themselves differently from each other. So overall politicians market themselves substantially by the differences in their use of language. They can choose their words in a distinctive manner to position themselves differently from other candidates and sell themselves more strongly. Therefore, as a bit of consumer advice; next time you go to the voting booths, don't just listen to what a politician has said, but also how he or she has said it.

Limitations and Recommendations

With this extensive research also come certain limitation that need to be addressed. The research method which is of descriptive nature looks at the frequency certain words are used and then categories them, with the help of the LIWC software, into different dictionary categories. In order to make this research feasible there needs to be a limit on the linguistic differences that determine a politician's marketing given their use of language. Accordingly, the explanatory variables from the LIWC software are narrowed down to the ones corresponding to IQ, emotion, perception, needs and sales language. For future research however, one can extend this research by analysing other variables that the LIWC software provides. Furthermore, this research used the variables from the LIWC to determine to what extent they influence a politician's marketing. However, these variables were not weighed and ranked accordingly to determine if some are more important than others when analysing the influence they have on a politician's marketing. Each variable was considered equally important. Moreover, actually witnessing the influence language has on a politician's marketing is more powerful than collecting data that supports statistical outcomes. Adding to that, this research was one sided in the sense that it only looked at use of language to determine differences between politicians and to what extent from a political marketing perspective. The effects of this on the public leave room for future research.

Tough the manner of data collection is suitable for this research, there is a form of human error that needs to be taken into consideration when transcribing speeches. Considering the importance of the validity and reliability of the data collection and results, one can try and

limit the effect human error has by increasing the amount of speeches that are collected. Also, the models used in the form of the independent t-test produce useful results. Yet as an extension, more models could be used that allow for more variables to be tested to determine the differences in use of language for politicians. One can for instance look at a factor and cluster analysis to group the variables differently together (Harman, 1960).

This paper can also appeal to future research on several different aspects. Firstly, the results from the analyses are valid, yet more variables can be analysed. Even though this research is done from several different aspects regarding the influence use of language has on a politician's marketing, extending this research by analysing word choice specifically for different topics may provide new insights. Next to that, voting outcomes should be analysed by cross referencing the use of language for a specific state and the voting results for that particular state. This may provide new insights as to what type of language scores best in which state. A linear regression could determine the effects different word categories have on votes (Seber & Lee, 2012). Accordingly, new political marketing strategies can be developed that focus on marketing yourself differently in your use of language for different states. Alongside this, surveys should be conducted that determine if and how the public's image of a politician has changed throughout their campaign. These results should then also be cross referenced with the type of language used at that particular moment.

Finally, future research should be done to find an optimal scenario for different types of language and election outcomes. With this research, the influence different word choices have on politician's marketing have been clearly indicated. However, the goal of a politician is to market themselves in such a way that they win the most number of votes. Since the dataset didn't show the results for number of votes gathered, future research in this field can determine the most optimal ratio for different uses of language to secure the most number of votes. First, research needs to be done to determine if and to what extent differences in vocabulary have an effect on votes. Secondly, a model can be developed that accurately determines the effects different ratios in use of language have on voter outcome. Finally, this model will than show the probability scenarios of how the outcome of elections is influenced by a different choice in words.

References

Akande, A. (2012, July 30). Quality of a good politician in a democratic society. Retrieved from http://www.vanguardngr.com/2012/07/quality-of-a-good-politician-in-a-democratic-society/

Bennett, A. (2015, December 11). Donald Trump: a bigot, or simply a man who knows what his public wants? *The Telegraph.* Retrieved from http://www.telegraph.co.uk/news/worldnews/donald-trump/12045284/Donald-Trump-a-bigot-or-simply-a-man-who-knows-what-his-public-wants.html

Boak, J. (2016, October 24). Policy Prescriptions: Trump and Clinton on economic growth. *The Washington Post.* Retrieved from https://www.washingtonpost.com/politics/policy-prescriptions-trump-and-clinton-on-economic-growth/2016/10/24/17ae2cb0-99b8-11e6-b552-b1f85e484086_story.html

Brekhus, K. (2015, October 6). Evidence that the left is smarter: Democrats have bigger vocabularies, better grammar than GOP. *PoliticsUSA.* Retrieved from http://www.politicususa.com/2015/10/06/study-finds-democrats-bigger-vocabularies-grammar-republicans.html

Brower, S. (2011). Culture of Fear. In *Journeys in Social Education* (pp. 93-110). SensePublishers.

Carroll, R. (2015, July 12). 'I'm, like, really a smart person': Donald Trump exults in outsider status. *The Guardian.* Retrieved from https://www.theguardian.com/us-news/2015/jul/12/im-like-a-really-smart-person-donald-trump-exults-in-outsider-status

Collins-Thompson, K., & Callan, J. (2004, July). Information retrieval for language tutoring: An overview of the REAP project. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 544-545). ACM.

Conge, P. (1988). The Concept of Political Participation: Toward a Definition. *Comparative Politics*, *20*(2), 241-249

Diamond, J. (2016, August 3). Donald Trump calls Hillary Clinton 'the devil'. *CNN politics.* Retrieved from http://edition.cnn.com/2016/08/01/politics/donald-trump-hillary-clinton-devil-election-2016/

Duggan, M., Ellison, N. B., Lampe, C., Lenhart, A., & Madden, M. (2015). Social media update 2014. *Pew Research Center*, *9*.

Edelman, M. (2013). *Political language: Words that succeed and policies that fail*. Elsevier.

Farmer, J. D., & Foley, D. (2009). The economy needs agent-based modelling. *Nature*, *460*(7256), 685-686.

Haas, A. (1979). Male and female spoken language differences: Stereotypes and evidence. *Psychological Bulletin*, *86*(3), 616.

Harman, H. H. (1960). Modern factor analysis.

Healy, P., & Burns, A. (2016, August 24). A Conflicted Trump Tries a New Tactic: Sticking to the Script. *The New York Times.* Retrieved from http://www.nytimes.com/2016/08/25/us/politics/donald-trump-presidential-race.html

Kinder, D. R. (1978). Political person perception: The asymmetrical influence of sentiment and choice on perceptions of presidential candidates. *Journal of Personality and Social Psychology*, *36*(8), 859.

Kraus, S. (1996). Winners of the first 1960 televised presidential debate between Kennedy and Nixon. *Journal of Communication*, *46*(4), 78-96.

Luhby, T (2015, October 14). Democrats said 'middle class' 11 times; the Republicans just three. *CNNMoney.* Retrieved from http://money.cnn.com/2015/10/14/news/economy/republican-democrats-middle-class/

Mosk, M. (2008). "A $2 Billion Presidential Race," *Washington Post* (online), Politics Section, October 22. Retrieved from http://voices.washingtonpost.com/44/2008/10/22/a_2_billion_presidential_race.html

Newman, B. I. (1999). *The mass marketing of politics: Democracy in an age of manufactured images*. Sage Publications.

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). The development and psychometric properties of LIWC2015. *UT Faculty/Researcher Works*.

Scammell, M. (1999). Political marketing: Lessons for political science. *Political studies*, *47*(4), 718-739.

Scharockman, A. (2016, June 29). The truth (so far) behind the 2016 campaign. *Politifact.* Retrieved from http://www.politifact.com/truth-o-meter/article/2016/jun/29/fact-checking-2016-clinton-trump/

Schultz, D. E., & Schultz, H. F. (2000). How to build a billion dollar business-to-business brand. *Marketing Management*, *9*(2), 22.

Schumacher, E., & Eskenazi, M. (2016). A Readability Analysis of Campaign Speeches from the 2016 US Presidential Campaign. *arXiv preprint arXiv:1603.05739*.

Seber, G. A., & Lee, A. J. (2012). *Linear regression analysis* (Vol. 936). John Wiley & Sons.

Shapiro, J., & Newby, A. (2015, October 29). Republicans are from the heart; Democrats are from the head. *Brookings.* Retrieved from https://www.brookings.edu/blog/order-from-chaos/2015/10/29/republicans-are-from-the-heart-democrats-are-from-the-head/

Shontell, A. (2016, January 2). Analyzing the way Donald Trump speaks like a salesperson. *Business Insider.* Retrieved from http://www.businessinsider.com/analyzing-the-way-donald-trump-speaks-like-a-salesman-2016-1?international=true&r=US&IR=T

Sinek, S. (2010). Simon Sinek: How great leaders inspire action.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of educational research*, *75*(3), 417-453.

Strauss, V. (2015, July 17). Yes. Donald Trump really went to an Ivy League School. *The Washington Post.* Retrieved from https://www.washingtonpost.com/news/answer-sheet/wp/2015/07/17/yes-donald-trump-really-went-to-an-ivy-league-school/

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology*, *29*(1), 24-54.

Tetlock, P. E. (1981). Pre-to postelection shifts in presidential rhetoric: Impression management or cognitive adjustment. *Journal of Personality and Social Psychology*, *41*(2), 207.

Villar, A., & Krosnick, J. A. (2011). Global warming vs. climate change, taxes vs. prices: Does word choice matter?. *Climatic change*, *105*(1-2), 1-12.

Viser, M. (2015, October 20). For Presidential Hopefuls, Simpler Language Resonates. *Boston Globe*.

Wattal, S., Schuff, D., Mandviwalla, M., & Williams, C. B. (2010). Web 2.0 and politics: the 2008 US presidential election and an e-politics research agenda. *Mis Quarterly*, 669-688.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Zernicke, P. H. (1990). Presidential roles and rhetoric. *Political Communication*, *7*(4), 231-245.

2016 Democratic Party Platform. Retrieved from https://www.democrats.org/party-platform

2016 Republican Platform. Retrieved from https://prod-static-ngop-pbl.s3.amazonaws.com/static/home/data/platform.pdf

Appendix

Exhibit A

Category	Abbrev	Examples	Words in category
Word count	WC	-	-
Summary Language Variables			
Analytical thinking	Analytic	-	-
Clout	Clout	-	-
Authentic	Authentic	-	-
Emotional tone	Tone	-	-
Words/sentence	WPS	-	-
Dictionary words	Dic	-	-
Linguistic Dimension			
Total function words	Funct	It, to, no, very	491
Total pronouns	Pronoun	I, them, itself	153
Personal pronouns	Ppron	I, them, her	93
1st pers singular	I	I, me, mine	24
1st pers plurar	We	We, us our	12
2nd person	You	You, your, thou	30
3rd pers singular	Shehe	She, her, him	17
3rd pers plural	They	They, their, they'd	11
Impersonal pronouns	Ipron	It, it's, those	59
Articles	Article	A, an, the	3
Prepositions	Prep	To, with, above	74
Auxiliary verbs	Auxverb	Am, will, have	141
Common Adverbs	Adverb	Very, really	140
Conjunctions	Conj	And, but, whereas	43
Negations	Negate	No, not, never	62
Other Grammar			
Common verbs	Verb	Eat, come, carry	1000
Common adjectives	Adj	Free, happy, long	764
Comparisons	Compare	Greater, best, after	317
Interrogative	Interrog	How, when, what	48
Numbers	Number	Second, thousand	36
Quantifiers	Quant	Few, many, much	77
Psychological Processes			
Affective processes	Affect	Happy, cried	1393
Positive emotion	Posemo	Love, nice, sweet	620
Negative emotion	Negemo	Hurt, ugly, nasty	744
Anxiety	Anx	Worried, fearful	116
Anger	Anger	Hate, kill, annoyed	230
Sadness	Sad	Crying, grief, sad	136
Social processes	Social	Mate, talk, they	756
Family	Family	Daughter, dad, aunt	118
Friends	Friend	Buddy, neighbor	95

Female references	Female	Girl, her, mom	124
Male references	Male	Boy, his, dad	116
Cognitive processes	Cogproc	Cause, know, ought	797
Insight	Insight	Think, know	259
Causation	Cause	Because, effect	135
Discrepancy	Discrep	Should would	83
Tentative	Tentat	Maybe, perhaps	178
Certainty	Certain	Always, never	113
Differentiation	Differ	Hasn't, but, else	81
Perceptual processes	Percept	Look, heard, feeling	436
See	See	View, saw, seen	126
Hear	Hear	Listen hearing	93
Feel	Feel	Feels, touch	128
Biological processes	Bio	Eat, blood, pain	748
Body	Body	Cheeks, hands, spit	215
Health	Health	Clinic, flu, pill	294
Sexual	Sexual	Horny, love, incest	131
Ingestion	Ingest	Dish, eat, pizza	184
Drives	Drives		1103
Affiliation	Affiliation	Ally, friend, social	248
Achievement	Achieve	Win, success, better	213
Power	Power	Superior, bully	518
Reward	Reward	Take, price, benefit	120
Risk	Risk	Danger, doubt	103
Time orientations	TimeOrient		
Past focus	Focuspast	Ago, did, talked	341
Present focus	Focuspresent	Today, is, now	424
Future focus	Focusfuture	May, will, soon	97
Relativity	Relative	Area, bend, exit	974
Motion	Motion	Arrive, car, go	325
Space	Space	Down, in, thin	360
Time	Time	End, until, season	310
Personal concerns			
Work	Work	Job, majors, xerox	444
Leisure	Leisure	Cook, chat, movie	296
Home	Home	Kitchen, landlord	100
Money	Money	Audit, cash, owe	226
Religion	Relig	Altar, church	174
Death	Death	Bury, coffin, kill	74
Informal language	Informal		380
Swear words	Swear	Fuck, damn, shit	131
Netspeak	Netspeak	Btw, lol, thx	209
Assent	Assent	Agree, OK, yes	36
Nonfluencies	Nonflu	Er, hm, umm	19
Fillers	Filler	Imean, youknow	14

Figure 1.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
number	36	,8900	3,5800	2,074444	,6000688
quant	36	1,2100	3,8700	2,454167	,5529266
posemo	36	2,3900	5,6300	3,711389	,7273912
negemo	36	,6500	3,8200	1,799167	,6981296
cogproc	36	7,3700	13,8300	10,406667	1,7636796
insight	36	,9000	2,8800	1,832222	,5835561
cause	36	,6400	2,3800	1,558889	,3579448
discrep	36	,5000	2,4800	1,630278	,4141531
tentat	36	,90	3,07	1,7622	,41200
certain	36	1,2700	3,1800	1,884167	,4526770
differ	36	1,4200	4,2600	2,915833	,6852460
percept	36	,8000	2,7100	1,736667	,4812840
see	36	,2100	1,5200	,631667	,2471841
hear	36	,0900	1,6200	,807500	,3966458
feel	36	,0000	,5200	,217500	,1304799
drives	36	8,3200	16,7500	11,929722	1,9667776
affiliation	36	1,8100	7,6900	4,423056	1,4310971
achieve	36	,8500	3,8100	1,943333	,6391244
power	36	2,3900	6,4000	4,542500	1,0156360
reward	36	,6000	3,0200	1,555278	,5797560
risk	36	,2000	1,7500	,734722	,3134097
focuspast	36	1,4200	4,8600	3,003611	,8293962
focuspresent	36	7,1800	18,1700	12,184444	2,9485361
focusfuture	36	,6300	3,7200	1,716111	,6579547
Valid N (listwise)	36				

Hypothesis 1.

Figure 2.

Group Statistics					
	Segment	N	Mean	Std. Deviation	Std. Error Mean
number	1,0	34712	,013	,1135	,0006
	2,0	50859	,022	,1464	,0006
quant	1,0	34712	,027	,1623	,0009
	2,0	50859	,021	,1445	,0006
cogproc	1,0	34712	,115	,3184	,0017
	2,0	50859	,122	,3270	,0015
insight	1,0	34712	,020	,1411	,0008
	2,0	50859	,024	,1537	,0007
cause	1,0	34712	,018	,1321	,0007
	2,0	50859	,017	,1284	,0006
discrep	1,0	34712	,018	,1339	,0007
	2,0	50859	,018	,1342	,0006
tentat	1,0	34712	,019	,1373	,0007
	2,0	50859	,020	,1399	,0006
certain	1,0	34712	,020	,1403	,0008
	2,0	50859	,019	,1372	,0006
differ	1,0	34712	,031	,1720	,0009
	2,0	50859	,034	,1812	,0008

Figure 3.

Independent Samples Test

		t	df	Sig. (2-tailed)	Mean Difference
number	Equal variances assumed	-9,508	85569	,000	-,0089
	Equal variances not assumed	-9,966	84206,414	,000	-,0089
quant	Equal variances assumed	5,431	85569	,000	,0057
	Equal variances not assumed	5,314	68696,631	,000	,0057
cogproc	Equal variances assumed	-3,229	85569	,001	-,0073
	Equal variances not assumed	-3,245	75848,391	,001	-,0073
insight	Equal variances assumed	-3,780	85569	,000	-,0039
	Equal variances not assumed	-3,842	78565,709	,000	-,0039
cause	Equal variances assumed	1,109	85569	,268	,0010
	Equal variances not assumed	1,103	73143,273	,270	,0010
discrep	Equal variances assumed	-,086	85569	,931	-,0001
	Equal variances not assumed	-,086	74662,848	,931	-,0001
tentat	Equal variances assumed	-,768	85569	,443	-,0007
	Equal variances not assumed	-,770	75459,817	,441	-,0007
certain	Equal variances assumed	,922	85569	,356	,0009
	Equal variances not assumed	,919	73459,723	,358	,0009
differ	Equal variances assumed	-2,806	85569	,005	-,0035
	Equal variances not assumed	-2,834	77039,490	,005	-,0035

Figure 4.

Group Statistics

	Segment	N	Mean	Std. Deviation	Std. Error Mean
posemo	1,0	34712	,039	,1924	,0010
	2,0	50859	,039	,1938	,0009
negemo	1,0	34712	,018	,1337	,0007
	2,0	50859	,017	,1305	,0006
percept	1,0	34712	,017	,1310	,0007
	2,0	50859	,022	,1475	,0007
see	1,0	34712	,007	,0834	,0004
	2,0	50859	,007	,0804	,0004
hear	1,0	34712	,006	,0785	,0004
	2,0	50859	,013	,1135	,0005
feel	1,0	34712	,003	,0587	,0003
	2,0	50859	,002	,0423	,0002

Figure 5.

Independent Samples Test

		t	df	Sig. (2-tailed)	Mean Difference
posemo	Equal variances assumed	-,425	85569	,671	-,0006
	Equal variances not assumed	-,425	74902,906	,671	-,0006
negemo	Equal variances assumed	,964	85569	,335	,0009
	Equal variances not assumed	,960	73347,108	,337	,0009
percept	Equal variances assumed	-4,888	85569	,000	-,0048
	Equal variances not assumed	-4,998	79940,811	,000	-,0048
see	Equal variances assumed	,866	85569	,386	,0005
	Equal variances not assumed	,860	72758,282	,390	,0005
hear	Equal variances assumed	-9,780	85569	,000	-,0069
	Equal variances not assumed	-10,456	85555,821	,000	-,0069
feel	Equal variances assumed	4,831	85569	,000	,0017
	Equal variances not assumed	4,550	58612,588	,000	,0017
	Equal variances not assumed	-9,670	83802,842	,000	-,0090

Figure 6.

Group Statistics

	Segment	N	Mean	Std. Deviation	Std. Error Mean
drives	1,0	34712	,134	,3408	,0018
	2,0	50859	,102	,3030	,0013
affiliation	1,0	34712	,050	,2186	,0012
	2,0	50859	,041	,1983	,0009
achieve	1,0	34712	,028	,1646	,0009
	2,0	50859	,013	,1126	,0005
power	1,0	34712	,045	,2072	,0011
	2,0	50859	,032	,1749	,0008
reward	1,0	34712	,018	,1313	,0007
	2,0	50859	,022	,1466	,0007
risk	1,0	34712	,009	,0930	,0005
	2,0	50859	,006	,0768	,0003

Figure 7.

Independent Samples Test

		t	df	Sig. (2-tailed)	Mean Difference
drives	Equal variances assumed	14,377	85569	,000	,0319
	Equal variances not assumed	14,063	68624,180	,000	,0319
affiliation	Equal variances assumed	6,449	85569	,000	,0093
	Equal variances not assumed	6,332	69669,507	,000	,0093
achieve	Equal variances assumed	15,852	85569	,000	,0150
	Equal variances not assumed	14,802	56493,458	,000	,0150
power	Equal variances assumed	10,175	85569	,000	,0134
	Equal variances not assumed	9,858	66021,532	,000	,0134
reward	Equal variances assumed	-4,534	85569	,000	-,0044
	Equal variances not assumed	-4,629	79595,451	,000	-,0044
risk	Equal variances assumed	4,785	85569	,000	,0028
	Equal variances not assumed	4,617	64948,578	,000	,0028

Figure 8.

Group Statistics

	Segment	N	Mean	Std. Deviation	Std. Error Mean
focuspast	1,0	34712	,031	,1727	,0009
	2,0	50859	,039	,1943	,0009
focuspresent	1,0	34712	,121	,3264	,0018
	2,0	50859	,164	,3705	,0016
focusfuture	1,0	34712	,014	,1192	,0006
	2,0	50859	,023	,1510	,0007

Figure 9.

Independent Samples Test

		t	df	Sig. (2-tailed)	Mean Difference
focuspast	Equal variances assumed	-6,584	85569	,000	-,0085
	Equal variances not assumed	-6,731	79891,905	,000	-,0085
focuspresent	Equal variances assumed	-17,508	85569	,000	-,0431
	Equal variances not assumed	-17,930	80254,321	,000	-,0431
focusfuture	Equal variances assumed	-9,254	85569	,000	-,0090
	Equal variances not assumed	-9,670	83802,842	,000	-,0090

Hypothesis 2.

Figure 10.

Group Statistics					
	Segment	N	Mean	Std. Deviation	Std. Error Mean
number	1,0	75380	,023	,1515	,0006
	2,0	67787	,019	,1357	,0005
quant	1,0	75380	,022	,1454	,0005
	2,0	67787	,027	,1611	,0006
cogproc	1,0	75380	,103	,3039	,0011
	2,0	67787	,110	,3127	,0012
insight	1,0	75380	,020	,1397	,0005
	2,0	67787	,019	,1351	,0005
cause	1,0	75380	,016	,1253	,0005
	2,0	67787	,017	,1274	,0005
discrep	1,0	75380	,016	,1264	,0005
	2,0	67787	,018	,1317	,0005
tentat	1,0	75380	,018	,1330	,0005
	2,0	67787	,018	,1346	,0005
certain	1,0	75380	,017	,1311	,0005
	2,0	67787	,019	,1362	,0005
differ	1,0	75380	,027	,1633	,0006
	2,0	67787	,032	,1760	,0007

Figure 11.

Independent Samples Test

		t	df	Sig. (2-tailed)	Mean Difference
number	Equal variances assumed	6,196	143165	,000	,0047
	Equal variances not assumed	6,232	143162,937	,000	,0047
quant	Equal variances assumed	-6,230	143165	,000	-,0050
	Equal variances not assumed	-6,197	137253,860	,000	-,0050
cogproc	Equal variances assumed	-4,232	143165	,000	-,0069
	Equal variances not assumed	-4,225	140613,384	,000	-,0069
insight	Equal variances assumed	1,820	143165	,069	,0013
	Equal variances not assumed	1,823	142416,956	,068	,0013
cause	Equal variances assumed	-,840	143165	,401	-,0006
	Equal variances not assumed	-,840	141024,272	,401	-,0006
discrep	Equal variances assumed	-2,081	143165	,037	-,0014
	Equal variances not assumed	-2,077	140126,194	,038	-,0014
tentat	Equal variances assumed	-,640	143165	,522	-,0005
	Equal variances not assumed	-,639	141186,243	,523	-,0005
certain	Equal variances assumed	-2,020	143165	,043	-,0014
	Equal variances not assumed	-2,015	140233,517	,044	-,0014
differ	Equal variances assumed	-5,071	143165	,000	-,0045
	Equal variances not assumed	-5,051	138669,861	,000	-,0045

Figure 12.

Group Statistics

	Segment	N	Mean	Std. Deviation	Std. Error Mean
posemo	1,0	75380	,036	,1856	,0007
	2,0	67787	,036	,1860	,0007
negemo	1,0	75380	,017	,1311	,0005
	2,0	67787	,019	,1348	,0005
percept	1,0	75380	,020	,1383	,0005
	2,0	67787	,016	,1259	,0005
see	1,0	75380	,006	,0797	,0003
	2,0	67787	,006	,0803	,0003
hear	1,0	75380	,010	,0995	,0004
	2,0	67787	,006	,0787	,0003
feel	1,0	75380	,002	,0423	,0002
	2,0	67787	,003	,0522	,0002

Figure 13.

Independent Samples Test

		t	df	Sig. (2-tailed)	Mean Difference
posemo	Equal variances assumed	-,155	143165	,876	-,0002
	Equal variances not assumed	-,155	141506,049	,876	-,0002
negemo	Equal variances assumed	-1,444	143165	,149	-,0010
	Equal variances not assumed	-1,442	140645,755	,149	-,0010
percept	Equal variances assumed	4,834	143165	,000	,0034
	Equal variances not assumed	4,858	143143,168	,000	,0034
see	Equal variances assumed	-,228	143165	,819	-,0001
	Equal variances not assumed	-,228	141339,041	,820	-,0001
hear	Equal variances assumed	7,908	143165	,000	,0038
	Equal variances not assumed	8,005	140882,937	,000	,0038
feel	Equal variances assumed	-3,753	143165	,000	-,0009
	Equal variances not assumed	-3,712	130553,920	,000	-,0009

Figure 14.

Group Statistics

	Segment	N	Mean	Std. Deviation	Std. Error Mean
drives	1,0	75380	,105	,3066	,0011
	2,0	67787	,127	,3328	,0013
affiliation	1,0	75380	,037	,1883	,0007
	2,0	67787	,046	,2091	,0008
achieve	1,0	75380	,015	,1212	,0004
	2,0	67787	,024	,1518	,0006
power	1,0	75380	,042	,2014	,0007
	2,0	67787	,047	,2125	,0008
reward	1,0	75380	,016	,1254	,0005
	2,0	67787	,015	,1212	,0005
risk	1,0	75380	,007	,0822	,0003
	2,0	67787	,008	,0885	,0003

Figure 15.

Independent Samples Test

		t	df	Sig. (2-tailed)	Mean Difference
drives	Equal variances assumed	-12,914	143165	,000	-,0218
	Equal variances not assumed	-12,859	138296,347	,000	-,0218
affiliation	Equal variances assumed	-8,549	143165	,000	-,0090
	Equal variances not assumed	-8,502	137143,817	,000	-,0090
achieve	Equal variances assumed	-12,026	143165	,000	-,0087
	Equal variances not assumed	-11,885	129501,478	,000	-,0087
power	Equal variances assumed	-4,606	143165	,000	-,0050
	Equal variances not assumed	-4,593	139612,566	,000	-,0050
reward	Equal variances assumed	1,662	143165	,096	,0011
	Equal variances not assumed	1,666	142434,909	,096	,0011
risk	Equal variances assumed	-2,409	143165	,016	-,0011
	Equal variances not assumed	-2,399	138708,858	,016	-,0011

Figure 16.

Group Statistics

	Segment	N	Mean	Std. Deviation	Std. Error Mean
focuspast	1,0	75380	,034	,1821	,0007
	2,0	67787	,027	,1619	,0006
focuspresent	1,0	75380	,131	,3373	,0012
	2,0	67787	,117	,3218	,0012
focusfuture	1,0	75380	,019	,1372	,0005
	2,0	67787	,013	,1143	,0004

Figure 17.

Independent Samples Test

		t	df	Sig. (2-tailed)	Mean Difference
focuspast	Equal variances assumed	8,084	143165	,000	,0074
	Equal variances not assumed	8,135	143146,779	,000	,0074
focuspresent	Equal variances assumed	7,825	143165	,000	,0137
	Equal variances not assumed	7,845	142670,361	,000	,0137
focusfuture	Equal variances assumed	8,856	143165	,000	,0059
	Equal variances not assumed	8,941	142349,345	,000	,0059

Hypothesis 3.

Figure 18.

Independent Samples Test

		t	df	Sig. (2-tailed)	Mean Difference
number	Equal variances assumed	-,645	6457	,519	-,002
	Equal variances not assumed	-,624	2906,366	,533	-,002
quant	Equal variances assumed	-,752	6457	,452	-,004
	Equal variances not assumed	-,733	2948,177	,463	-,004
cogproc	Equal variances assumed	1,644	6457	,100	,015
	Equal variances not assumed	1,690	3258,174	,091	,015
insight	Equal variances assumed	,672	6457	,501	,003
	Equal variances not assumed	,694	3291,203	,488	,003
cause	Equal variances assumed	1,199	6457	,231	,004
	Equal variances not assumed	1,274	3494,286	,203	,004
discrep	Equal variances assumed	2,979	6457	,003	,011
	Equal variances not assumed	3,551	4575,919	,000	,011
tentat	Equal variances assumed	1,309	6457	,191	,005
	Equal variances not assumed	1,400	3544,992	,162	,005
certain	Equal variances assumed	-1,987	6457	,047	-,008
	Equal variances not assumed	-1,836	2687,504	,066	-,008
differ	Equal variances assumed	,934	6457	,350	,005
	Equal variances not assumed	,965	3298,698	,334	,005

Figure 19.

Independent Samples Test

		t	df	Sig. (2-tailed)	Mean Difference
posemo	Equal variances assumed	-,296	6457	,767	-,002
	Equal variances not assumed	-,294	3037,984	,769	-,002
negemo	Equal variances assumed	1,507	6457	,132	,005
	Equal variances not assumed	1,639	3679,012	,101	,005
percept	Equal variances assumed	-2,065	6457	,039	-,007
	Equal variances not assumed	-1,874	2613,160	,061	-,007
see	Equal variances assumed	-,291	6457	,771	-,001
	Equal variances not assumed	-,284	2953,777	,776	-,001
hear	Equal variances assumed	-2,811	6457	,005	-,006
	Equal variances not assumed	-2,290	2252,503	,022	-,006
feel	Equal variances assumed	-,717	6457	,474	-,001
	Equal variances not assumed	-,668	2723,914	,504	-,001

Figure 20.

Independent Samples Test

		t	df	Sig. (2-tailed)	Mean Difference
drives	Equal variances assumed	,226	6457	,821	,002
	Equal variances not assumed	,226	3105,678	,821	,002
affiliation	Equal variances assumed	-1,273	6457	,203	-,008
	Equal variances not assumed	-1,233	2913,289	,218	-,008
achieve	Equal variances assumed	2,957	6457	,003	,015
	Equal variances not assumed	3,304	3912,163	,001	,015
power	Equal variances assumed	1,388	6457	,165	,009
	Equal variances not assumed	1,442	3334,656	,149	,009
reward	Equal variances assumed	,584	6457	,559	,002
	Equal variances not assumed	,599	3254,273	,549	,002
risk	Equal variances assumed	-1,315	6457	,188	-,003
	Equal variances not assumed	-1,212	2675,316	,226	-,003

Figure 21.

Independent Samples Test

		t	df	Sig. (2-tailed)	Mean Difference
focuspast	Equal variances assumed	-1,544	6457	,123	-,008
	Equal variances not assumed	-1,468	2817,232	,142	-,008
focuspresent	Equal variances assumed	-,473	6457	,636	-,004
	Equal variances not assumed	-,469	3043,480	,639	-,004
focusfuture	Equal variances assumed	-,116	6457	,907	,000
	Equal variances not assumed	-,116	3054,731	,908	,000

Figure 22.

Independent Samples Test

		t	df	Sig. (2-tailed)	Mean Difference
number	Equal variances assumed	3,530	11034	,000	,010
	Equal variances not assumed	3,660	10989,011	,000	,010
quant	Equal variances assumed	-,049	11034	,961	,000
	Equal variances not assumed	-,049	10062,731	,961	,000
cogproc	Equal variances assumed	1,332	11034	,183	,008
	Equal variances not assumed	1,338	10245,725	,181	,008
insight	Equal variances assumed	1,396	11034	,163	,004
	Equal variances not assumed	1,415	10537,176	,157	,004
cause	Equal variances assumed	,905	11034	,366	,002
	Equal variances not assumed	,913	10409,723	,361	,002
discrep	Equal variances assumed	-,495	11034	,620	-,001
	Equal variances not assumed	-,493	9889,849	,622	-,001
tentat	Equal variances assumed	1,736	11034	,083	,004
	Equal variances not assumed	1,769	10678,021	,077	,004
certain	Equal variances assumed	-4,799	11034	,000	-,014
	Equal variances not assumed	-4,601	8329,416	,000	-,014
differ	Equal variances assumed	1,206	11034	,228	,004
	Equal variances not assumed	1,217	10413,179	,223	,004

Figure 23.

Independent Samples Test

		t	df	Sig. (2-tailed)	Mean Difference
posemo	Equal variances assumed	-1,954	11034	,051	-,007
	Equal variances not assumed	-1,928	9566,734	,054	-,007
negemo	Equal variances assumed	-5,310	11034	,000	-,015
	Equal variances not assumed	-5,065	8115,433	,000	-,015
percept	Equal variances assumed	2,067	11034	,039	,005
	Equal variances not assumed	2,115	10776,220	,034	,005
see	Equal variances assumed	-2,373	11034	,018	-,003
	Equal variances not assumed	-2,267	8171,496	,023	-,003
hear	Equal variances assumed	6,155	11034	,000	,011
	Equal variances not assumed	6,816	9322,449	,000	,011
feel	Equal variances assumed	-3,337	11034	,001	-,003
	Equal variances not assumed	-2,987	5774,658	,003	-,003

Figure 24.

Independent Samples Test

		t	df	Sig. (2-tailed)	Mean Difference
drives	Equal variances assumed	-8,338	11034	,000	-,055
	Equal variances not assumed	-8,132	9062,402	,000	-,055
affiliation	Equal variances assumed	-6,494	11034	,000	-,029
	Equal variances not assumed	-6,268	8615,809	,000	-,029
achieve	Equal variances assumed	-4,212	11034	,000	-,010
	Equal variances not assumed	-4,031	8247,706	,000	-,010
power	Equal variances assumed	-6,096	11034	,000	-,024
	Equal variances not assumed	-5,867	8489,979	,000	-,024
reward	Equal variances assumed	1,269	11034	,205	,004
	Equal variances not assumed	1,284	10491,572	,199	,004
risk	Equal variances assumed	-2,030	11034	,042	-,004
	Equal variances not assumed	-1,974	8937,649	,048	-,004

Figure 25.

Independent Samples Test

		t	df	Sig. (2-tailed)	Mean Difference
focuspast	Equal variances assumed	1,959	11034	,050	,007
	Equal variances not assumed	1,989	10583,959	,047	,007
focuspresent	Equal variances assumed	7,244	11034	,000	,049
	Equal variances not assumed	7,407	10763,195	,000	,049
focusfuture	Equal variances assumed	-,132	11034	,895	,000
	Equal variances not assumed	-,132	10037,066	,895	,000