JOSIP JURAJ STROSSMAYER UNIVERSITY OF OSIJEK
FACULTY OF HUMANITIES AND SOCIAL SCIENCES

Dragana Božić Lenard

GENDER DIFFERENCES IN THE POLITICAL SPEECHES FROM THE 113TH UNITED STATES CONGRESS

DOCTORAL THESIS

Osijek, 2016

SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU
FILOZOFSKI FAKULTET

Dragana Božić Lenard

RODNE RAZLIKE U POLITIČKIM GOVORIMA 113. AMERIČKOG KONGRESA

DOKTORSKI RAD

Osijek, 2016.

JOSIP JURAJ STROSSMAYER UNIVERSITY OF OSIJEK
FACULTY OF HUMANITIES AND SOCIAL SCIENCES

DRAGANA BOŽIĆ LENARD

GENDER DIFFERENCES IN THE POLITICAL SPEECHES FROM THE 113TH UNITED STATES CONGRESS

DOCTORAL THESIS

Supervisor:
Marija Omazić, Ph.D., Full Professor

Osijek, 2016

SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU
FILOZOFSKI FAKULTET

DRAGANA BOŽIĆ LENARD

RODNE RAZLIKE U POLITIČKIM GOVORIMA 113. AMERIČKOG KONGRESA

Mentor:
prof. dr. sc. Marija Omazić

Osijek, 2016.

ACKNOWLEDGEMENTS

I am deeply indebted to my advisor, Full Professor Marija Omazić, PhD, for her fundamental role in my work. She provided me with guidance and expertise, encouraged my research and allowed me to grow as a researcher. I would also like to thank the British Scholarship Trust for awarding me with a research scholarship and Middlesex University London for accepting me as a visiting researcher. I would particularly like to acknowledge Senior Lecturer Sylvia Shaw, PhD for inviting me to Middlesex University London in the fall of 2015 and providing me with her time and valuable feedback on a preliminary version of this thesis. I am also thankful to my alma mater - the Faculty of Humanities and Social Sciences Osijek, which played a crucial role in my academic career. I am also grateful to my employer the Faculty of Electrical Engineering, Computer Science and Information Technology Osijek and my colleagues for their support and encouragement to strive towards my goals.

My special thanks go to my family – my late mother Marija, my father Dragan and my sister Maja for their love, support, sacrifices and financial support. Without them, this journey would not have been possible. This last word of acknowledgment I have saved for my husband Ivan who spent sleepless nights with me supporting me in joyful and difficult moments and made this journey memorable.

Declaration

I declare that this thesis was composed by myself and that the work contained herein is my own except where explicitly stated otherwise in the text. This work has not been submitted for any other degree or professional qualification except as specified.

Dragana Božić Lenard

Abstract

The main objective of the thesis is to study gender differences in language use in a setting of political speeches. The 113th United States Congress, a legislative branch meeting which lasted from January 2013 until January 2015, was selected. In addition to the meeting being the most recent one, the fact that it was composed of the record number of female participants, namely 450 male and 103 female politicians, was the reason for choosing the 113th Congress. The speech transcripts, downloaded from the official repository *The Library of Congress Thomas*, were included in the corpus. The corpus was composed of all 100-word uninterrupted speeches. The technical prerequisite for analyzing the speech is for it to be composed of at least 100 words, hence the choice. 672 speeches by the female and 2,983 by the male politicians were included in the analysis. The corpus was analyzed with the text analysis software *Linguistic Inquiry and Word Count*, which calculated the degree to which the politicians used 70 language categories ranging from the word count and grammatical categories to different topics, spoken and punctuation categories. The computational analysis results were uploaded in the software for the statistical analysis *Statistical Package for the Social Sciences* which was used to do the Mann-Whitney U, independent sample t-test, Kruskal-Wallis, one-way ANOVA and two-tailed Spearman correlation tests. The statistical tests were used to study the differences in language use by the male and the female politicians, i.e. to calculate if the differences were statistically significant. They were also used in examining some intragroup differences and correlations between variables. The quantitative analysis results were interpreted and the possible underlying reasons for the gender differences elaborated on.

The selected tools for the computational and statistical data analysis were proven to be useful. The text analysis software LIWC is a useful tool for the fast and accurate analysis of a large corpus. The pre-established independently rated categories, containing 4,500 words and word stems, provide an unbiased word categorization. Due to their compatibility, LIWC

output results are easily uploaded in SPSS, which provides a wide choice of statistical tests, thus making SPSS an excellent supplemental tool. The tools are sufficient for a quantitative data analysis, yet less helpful in results interpretation. To put it differently, the tools provide the information about the difference in frequency and not the contextual usage itself.

In the analysis of the corpus, several major gender differences emerged. The female politicians were shown to be more formal, critical and task-focused, while the male politicians were more socially oriented and elaborative, occupying the floor more than the female politicians. While the female politicians worked on establishing themselves as independent politicians, the male politicians embraced their collective identities. Also, the female politicians focused on raising the awareness of different health issues and providing support for patients and their families, the male politicians focused on the consequences and possible solutions to the problems.

Key words: gender differences, language use, 113th United States Congress, computational analysis, statistical analysis, political discourse, underlying reasons for the gender differences

Sažetak

Ciljevi i hipoteza

Glavni je cilj ove doktorske disertacije bio utvrditi postoje li razlike u jezičnome izražavanju političara i političarki koji su sudjelovali u 113. sazivu američkoga Kongresa. Osim pobrojavanja i klasificiranja rodnih razlika, cilj je bio izračunati statističku značajnost u uporabi određene jezične kategorije. Drugim riječima, cilj je bio odrediti koja je od dviju proučavanih skupina ispitanika upotrebljavala svaku od 70 jezičnih kategorija više i je li ta razlika u uporabi bila statistički značajna ili ne. Također, jedan je od ciljeva bio istumačiti dobivene statističke rezultate nudeći moguća objašnjenja za rodne razlike u jezičnome izražavanju, odnosno u uporabi pojedine jezične kategorije. Osnovna je hipoteza da postoje statistički značajne rodne razlike u jezičnome izražavanju političara i političarki koji su sudjelovali u 113. sazivu Kongresa. Također, pretpostavilo se da će političarke, s ciljem uspješnijega etabliranja, biti pripremljenije, ozbiljnije i formalnije u iznošenju svojih ideja, dok će političari govoriti duže, biti ekspresivniji i koristiti agresivnije reference.

Metodologija

Korpus istraživanja sastoji se od transkripata govora iz 113. saziva američkoga Kongresa preuzetih sa službenoga repozitorija govora iz Kongresa *Thomas*. Repozitorij se sastoji od svih govora iz Doma i Senata, naknadnih replika i objašnjenja političara, rasprava, dnevnih obavijesti, poziva na sjednice, amandmana i načina glasovanja. U korpus su istraživanja uključeni svi neprekinuti govori pojedinoga političara koji ispunjavaju tehnički uvjet od najmanje 100 riječi. Točnije, 672 govora političarki i 2 983 govora političara, transkribiranih na ukupno 5 504 stranice, uključeni su u korpus istraživanja. Prije analize, transkripti su pročišćeni slijedeći službene tehničke upute.

Istraživanje se sastojalo od dvaju dijelova – kvantitativnoga i kvalitativnoga. U kvantitativnome su se dijelu istraživanja provele dvije analize. Prvo se provela računalna analiza korpusa primjenom programskoga alata *Linguistic Inquiry and Word Count*.

Programski alat analiza svaku pojedinačnu riječ u korpusu i uspoređuje ju sa svojim unutarnjim rječnikom koji se sastoji od 4 500 riječi i korijena riječi. Unutarnji je rječnik sastavio panel nezavisnih stručnjaka, a odabir su i kategorizacija riječi za uvrštavanje u rječnik prošli nekoliko procjenjivačkih etapa. Nakon što programski alat usporedi svaku riječ s pojavnicama u svome rječniku, svrstava ih u 80 kategorija koje se kreću od prebrojavanja riječi i gramatičkih kategorija do tematskih i interpunkcijskih kategorija. Za potrebe se ovoga istraživanja odabralo 70 kategorija koje nudi spomenuti programski alat.

Rezultati računalne analize učitani su u programski alat za statističku analizu podataka *Statistical Package for the Social Sciences*. Kako bi se odabrali adekvatni testovi za statističku analizu, prvo je proveden test distribucije podataka kontinuiranih varijabli. Nadalje, s obzirom na postavljena istraživačka pitanja, u programskome su se alatu koristili Mann-Whitney U i t-test za nezavisne uzorke. Također, korišteni su i jednosmjerna analiza varijance ANOVA, Kruskal-Wallis te Spearmanov koeficijent korelacije kako bi se napravila podloga i predložio smjer budućih istraživanja u području.

U kvalitativnome su se dijelu istraživanja interpretirali rezultati dobiveni statističkom analizom te objašnjavali razlozi rodnih razlika u korištenju pojedine varijable uzimajući u obzir kontekst govora i koristeći kritičku analizu diskursa.

Rezultati i zaključci

Korišteni programski alati pokazali su se korisnima za kvantitativnu analizu podataka. LIWC je koristan programski alat koji omogućava brzu, točnu i nepristranu kategorizaciju riječi i računalnu analizu. Posebno je pogodan za analize na velikome korpusu. S obzirom na kompatibilnost dvaju programskih alata, rezultati računalne analize lako se mogu učitati u programski alat za statističku obradu podataka SPSS koji omogućava širok izbor parametrijskih i neparametrijskih testova. Međutim, programski alat SPSS može samo dati odgovor na pitanje o frekvenciji korištenja kontinuiranih varijabli i statističkoj značajnosti, ali ne i o kontekstualnoj uporabi pojedine varijable. Drugim riječima, ukoliko je istraživačko pitanje koja od dviju ili više promatranih skupina upotrebljava pojedinu varijablu više i je li razlika u uporabi statistički značajna, programski alati LIWC i SPSS mogu se koristiti kao

glavna sredstva u istraživanja. No, ukoliko, pored spomenutoga, istraživačko pitanje uključuje i razloge za uporabu pojedine varijable, programski se alati mogu koristiti samo kao pomoćna sredstva u istraživanju.

Po provedbi računalne, statističke i kontekstualne analize korpusa, donijelo se nekoliko glavnih zaključaka u vezi s rodnim razlikama u jezičnoj uporabi na korpusu govora iz 113. saziva američkoga Kongresa. Političarke su imale ozbiljniji pristup formalnije se obraćajući predsjedavajućem te obrazlažući razloge svoga govora, dok su političari svoje govore češće započinjali tehnikom *in medias res*. Političarke su bile kritičnije u svojim govorima, koristile složeniji vokabular i više primjera vezanih uz posao, dok su političari proveli više vremena za govornicom i nerijetko iznosili primjere iz privatnoga života. Kontekstualnom se analizom uporabe zamjenica može zaključiti da političarke rade na etabliranju sebe kao nezavisnih sudionika političke scene, dok političari spremnije prihvaćaju kolektivni identitet pri čemu češće spominju doprinose svoje grupacije. Političari su ekspresivniji i češće izražavaju svoje osjećaje na različite načine tijekom čega su im govori postajali nesigurniji, dok su političarke usmjerenije na izvršavanje zadatka. I jedna i druga skupina ispitanika daju veći javi prostor osobama muškoga roda. Nadalje, ciljevi govora političarki često su bili orijentirani na podizanje svijesti o određenome problemu i pružanje podrške obiteljima oboljelih ili ratnih žrtava, dok su političari djelovali proaktivno ukazujući na posljedice i nudeći rješenja za određene probleme. I političari i političarke bavili su se pitanjima postignuća i uspjeha, no političari su uspjeh promatrali kroz prizmu poražavanja protivnika, dok su se političarke fokusirale na rezultat koji nije uključivao nečiji poraz.

Metodološki se okvir i rezultati ovoga istraživanja mogu primijeniti u daljnjim istraživanjima rodnih razlika u političkome diskursu na način da se proširi korpus, napravi komparativna analiza više saziva Kongresa, istraže jezična izražavanja u kongresima drugih zemalja, uvrste nove kategorijske varijable poput stranačke pripadnosti, sudjelovanja u izvršnoj ili zakonodavnoj vlasti i sl.

Ključne riječi: rodne razlike, jezično izražavanje, 113. saziv američkoga Kongresa, računalna analiza, statistička analiza, politički diskurs, razlozi rodnih razlika

List of Tables

List of Figures

List of Abbreviations:

LIWC – Linguistic Inquiry and Word Count

SPSS – Statistical Package for Social Sciences

CDA – Critical Discourse Analysis

i.e. – *id est*; that is; in other words

etc. – *et cetera*; and other things

e.g. – *exempli gratia*; for example

EFL – English as a Foreign Language

ESL – English as a Second Language

ESP – English for Specific Purposes

Table of Contents

1. Introduction

1. 1. Subject matter

Language is one of the most important means of humans' expression of thoughts. Guided by thoughts, choices people make in the forms of expression can be paralleled to their perception of things from the real world and consequently the way they express themselves about those things. Two people may be speaking about the same thing with their descriptions being utterly unrelated. To put it differently, linguistic choices may be a diagnostic of people's both overt and covert feelings about things from the real world. By studying people's linguistic choices, we may learn a lot about their desires, feelings, perceptions and thoughts. There is a consensus among scientists that personality and language used in a variety of contexts – everyday speech (Mehl et al., 2006), interviews (Fast & Funder, 2008), broadcast news speech (Alam & Riccardi, 2013), guided written assignments (Pennebaker & King, 1999; Hirsh & Peterson, 2009), e-mails (Oberlander & Gill, 2004) – are intertwined. We may categorize people according to their linguistic choices or speech styles. To exemplify, the American linguist William Labov (1966) studied the speech of employees from three department stores in New York: *S. Klein* (a discount working-class store), *Macy's* (a moderately priced middle-class store) and *Saks Fifth Avenue* (an expensive upper-middle class store). Asking questions which should elicit the answer *fourth flour*, Labov aimed to study the pronunciation of /r/. The results pointed to a social stratification, i.e. the pronunciation of /r/ depended on the employees' social-class.

Another category based on people's linguistic choices is gender. Men and women have been alleged to differ in every area of psychological functioning at one point or another, so language use is not an exception. Believing for the seat of the intellect to be situated in the brain, differences in verbal ability were tried to be explained by the differences between the brains of men and women (Halpern, 1994). Despite the fact that the phrenologists provided considerable evidence as to the differences in the physical features of men's and women's brains (different frontal lobe and brain tissue) (Walker, 1850 cited in Caplan et al., 1997), the question on gender differences in linguistic choices was not successfully answered.

Gender differences mirrored in language have been extensively investigated by sociolinguists since the 1960s. Robin Lakoff's pioneering work *Language and Women's Place* from 1975 has initiated numerous sociolinguistic research and explanations regarding the origins of gender differences in language use. Aiming at studying the origins of inequalities, researchers examined the earliest speech patterns of boys and girls. The research results indicated that even kindergarten children use different gender-related linguistic expressions (Tannen, 1990). Furthermore, researchers discovered that different social roles are attributed to children based on their gender. If trying to violate pre-attributed social roles or employ other gender's means of linguistic expression, children are warned and instructed to use the gender proper means. The instructions are especially given to girls in order for them to be unquestionably accepted as a part of society (Tannen, 1997). As children grow into adulthood, the instructions on all language levels as well as the ones with respect to nonverbal behavior are continued. One may draw a conclusion that different social roles result in different means of linguistic expression employed by men and women.

Maturity and a higher level of education, among other variables, may trigger one's independence and consequently cause changes and the disturbance of the prearranged social and linguistic hierarchy. Speakers may start to use gender-free linguistic expressions despite risking possible disapproval. The organizational order in both private and public sphere has gradually been violated with men and women taking the roles freely. On the one hand, by receiving an aspired level of education, women are no longer limited to solely perform the housewives' and mothers' role. The number of men performing it tends to moderately increase. On the other hand, being educated, women can do the same jobs as men reducing the aforementioned differences to a minimal level. That being said, the process of a social hierarchy becoming gender-free comes naturally.

Gender-based discrimination has not been eradicated yet, so much so that even in the 21st century, there are job sectors primarily taken by men and the ones reserved for women. Even if employed in the other gender's field, jobs positions are not equally distributed. High-rank decision-making job positions are usually occupied by men compared to low-rank positions commonly occupied by women. Gender-based jobs are grounded in personality traits possessed by men and women, so job fields requiring caring, nurturing, collaboration,

active listening, patience, etc. are generally occupied by women in comparison to jobs occupied by men, which require competition, leadership skills and imposing one's ideas.

Potential problems to personality traits occur when one decides to do the other gender's jobs. In order to be successful, one needs to develop and display preferred personality traits for that job, even if they clash with personality characteristics usually associated with men or women. Costa et al. (2001: 328) studied differences in self-rated personality traits across different cultures. Their results show that women score high on neuroticism (depression, anxiety), agreeableness (altruism, tender-mindedness, confidence) and openness to feelings. In comparison, men score high on assertiveness and openness to new ideas. These personality differences are consistent with gender stereotypes portraying women as more caring and emotional than men and men as more rational and assertive than women (Best & Williams, 1982). Such differences can drive differences in attitudes towards education, poverty, use of force and money management (Schwartz & Rubel, 2005, Eagly et al., 2004) which might influence a job choice.

Whatever one does, he/she will be criticized. One the one hand, by displaying job preferred personality traits, one will be professionally successful, yet criticized for gender treason. On the other hand, if preserving personality traits for specific gender, one may disqualify himself/herself from professional advancement. The same attitude is applied to the use of language. If a man uses "women's language", he is labeled as effeminate or womanly making him a linguistic anomaly and an outcast (Hall, 2003: 355). Emasculation is also articulated in Regina Flannery's (1946: 133) article *Men's and Women's Speech in Gros Ventre* where she claims that "the expressions used by women are more modest and that if a woman used men's words she would be considered mannish, and likewise a man who used women's words would be considered effeminate."

Identified as the struggle for power and imposing one's ideas, politics is the field naturally occupied by men whose personality traits (strength, knowledgeability, assertiveness, directness) are a prototype for it (Huddy & Terkildsen, 1993; Kahn 1996). The fact that the number of men in politics still prevails comes as no surprise. In other words, even in the modern era, women are still under-represented in politics. When running for office, women tend to hold lesser value offices that include education, environment, social and health care

services. Women get to hold offices not that rigorously associated with masculinity traits. Voters associate women candidates with solidarity issues (education, children, the elderly, social affairs, health care and the environment), while men candidates are associated with business, economy, military and agriculture (Huddy & Terkildsen, 1993; Alexander & Andersen, 1993; Leeper, 1991). Furthermore, when women run for office in one of men's sectors, they will be prejudiced and receive less votes (Dolan, 2008). If voters reject stereotypes and trust women to hold offices in men's sectors, they will be depoliticized, womanized and maternalized by media (Bengoechea, 2011). Therefore, entering the world of politics – the world that has always been claimed by men - causes certain changes for women. They need to acquire some of men's personality traits, which might initiate other changes including the linguistic ones.

The means of linguistic expressions used by men have been considered as a norm by researchers of deficit (Lakoff, 1975; Holmes, 2006), dominance (Zimmerman & West, 1975), difference (Tannen, 1990) and communities of practice (Eckert & McConnell-Ginet, 2003) approach, i.e. despite different approaches to gender differences, they study women's speech in comparison to men's. This particularly applies to the linguistic expressions in the field often labeled as the men' field - the field of politics. Hence, if wanting for their views and actions to be taken into serious consideration, women in politics might want to adjust their linguistic styles. That being said, one may expect that politicians, regardless of their gender, use linguistic expressions employed exclusively by men. Linguistic expressions used by women are not to be used in the field whose fundamental concepts are commonly metaphorically mapped from the domain of war and best summarized into three words – struggle for power.

This dissertation's main objective is to study the linguistic practices of the American politicians and their relationship to gender in the last completed 113th American Congress. The following specific objectives are pursued:

- To investigate whether the linguistics practices of the men and the women in political discourse differ;
- To inspect if there is a statistically significant difference between the men and the women in the usage of 70 variables;

- To determine which of the two respective groups use a specific variable more than another group;
- To interpret the results by providing underlying reasons of the variable usage difference.

The outlined specific objectives are relevant to three principal intertwined areas in language and gender research. Firstly, many researchers (Eckert & McConnell-Ginet, 1992; Coates, 1993; Holmes, 1995; Weatherall, 2002) have claimed that men are more likely to use competitive and women cooperative style of speech. Even though these claims can be criticized for overgeneralization, it is indeed a perception among politicians themselves that men adopt aggressive and women more consensual style in the political setting. Secondly, this thesis builds on a growing body of research into gender differences in language, especially gender differences in language used by men and women in public speaking settings (Mulac et al., 1986; West 1990; Holmes, 1992; Baxter 1999a; 1999b; McElhinny, 1998; Burns et al., 2006; Griswold, 2007; Wodak, 2008; Yu, 2013) where the men's speaking style is considered as a norm to be conformed to. Thirdly, the thesis will contribute to debates about women bringing changes into language or assimilating to dominant men's styles, i.e. whether they work towards changing preset practices monopolized by men, accept it or balance between these two positions (Lovenduski, 1996; Childs, 2000; Walsh, 2001). Stereotypical believes about gender differences in language and the hypothesis that female politicians bring a different voice and style to the political arena require systematic investigation. Since there is no current research into gender differences in language on the corpus of the speeches from the 113th United States Congress, this thesis represents an original contribution to sociolinguistic and political studies.

The following argument is suggested:

- Linguistic choices of the female politicians are being assimilated to practices appropriate for the political setting monopolized by male politicians' styles

1.2. Dissertation structure

The introductory chapter presents the subject matter of the research – gender differences in language. It provides a brief overview of the most influential theories and scholars dealing with the issue. It also sets the direction this research will take by listing the specific objectives and the research hypothesis.

The second chapter provides a theoretical overview for the dissertation. Since the concepts of sex and gender are frequently interchangeably used, the chapter begins with the definitions and comparisons of the concepts resulting in an educated decision on the future terminology usage. The chapter further provides a historical overview of the most important gender linguistic theories; lists and presents their representatives' main ideas and criticisms.

The third chapter presents the methodological matters of the research. It provides a detailed stepwise description of the research design and presents the socioeconomic characteristics of the participants from the 113th United States Congress. Furthermore, the procedures and reasons for the corpus design are elaborated on. Also, the chapter presents and exemplifies the working principles of the text analysis software used in the research. Finally, it illustrates the normality distributions tests resulting in the choice of (non)parametric tests used for the statistical data analysis.

The fourth chapter consists of the analysis of the research findings. The research includes both the quantitative and the qualitative analysis. First, each of 70 variables is analyzed with the software for the statistical analysis SPSS and the quantitative analysis results are presented in their respective subsections. Then, in the same subsections, the results are interpreted by providing the underlying reasons for the linguistic choices and gender differences.

The concluding chapter evaluates the software used in the research and summarizes the research results. It also provides a review of the objectives and research questions; points

to the strengths and limitations of the research and gives several recommendations for future research.

2. Theoretical background

This chapter presents a survey of the fields the thesis is drawn from. The chapter begins with a definition of the concept of gender. Upon defining the concept of sex, the two concepts are contrasted resulting in an educated decision on the further usage. The chapter further provides a brief historical overview of the most important 20th century linguistic theories with an earlier theory causing reactions and influencing a later one. Naturally, the focus of attention is placed on the gender theories within the field of sociolinguistics beginning with the Lakoff's hypothesis and the deficit approach. Simultaneously, Zimmerman and West developed the dominance approach whose drawbacks led to establishing the difference approach. The following phase in gender research addresses the limitations caused by essentialist interpretations of gender thus putting an emphasis on discourse. Moreover, it is discourse and social context that determine which community of practice an individual will belong to. Finally, the theoretical overview is concluded with the critical discourse analysis approach suggesting a new variable to correlate with language – that of power.

2.1 Gender versus sex

To begin with, we need to make a distinction between the two types of gender. On the one hand, grammatical gender is a property of nouns which affects grammatical agreement between a noun and an accompanying adjective, article, number and other basic sentence parts (Cruse, 2006: 77). Natural gender, on the other hand, is determined by features of a referent. There is only a partial correlation between these two concepts. Swiss linguist and semiotician Ferdinand de Saussure (1916) believed that a linguistic sign is composed of two parts – signified and a signifier. Given the arbitrary nature of signs, there is no natural relationship between the signified and the signifier; consequently, their relationship is based

on a convention. This notion can be applied to the relationship between grammatical and natural gender – grammatical gender is attributed arbitrarily and conventionally, whereas natural gender may or may not be based on biological traits. Given the nature of this dissertation, we will only deal with natural gender.

Feminist theorists believe that a distinction between the concepts of *sex*, *sexuality* and *gender* needs to be made. The terms *sex* and *gender* are often used interchangeably but incorrectly. According to Medilexicon medical dictionary, *sex* or *biological gender* is "the biologic character or quality that distinguishes male and female from one another as expressed by analysis of the person's gonadal, morphologic (internal and external), chromosomal, and hormonal characteristics." To paraphrase, based on biological characteristics, sex is assigned to an individual at birth; therefore, there is *male sex* and *female sex* (Holmes, 2001; Trudgill, 2000). Gender, on the other hand, is a more complicated concept. The term *gender* implies a socio-cultural construct. It is used when referring to "social, cultural and psychological constructs that are imposed upon these biological differences" (Shapiro, 1981 cited in Yanagisako & Collier 1990: 139). Similarly, Lipman-Blumen (1984: 3) states that gender addresses "all those cultural expectations associated with masculinity and femininity that go beyond biological sex differences". Biological sex is attributed to an individual at birth. While growing up, an individual is raised to adopt the set of gender-labeled social rules which are built upon sex. To paraphrase, sex refers to biology and physiology, sexuality points to sexual preferences, orientation and practices, while gender deals with social roles and status (Dovi, 2008: 154). Shapiro's model has been criticized because of the polarity gender is based on (Cameron, 1997), overstating similarities within the categories and understating similarities across the categories (Nicholson, 1994).

Social constructivists offer a radical critique of biological determinism. They believe that instead of viewing sex as primary and biological while gender as secondary and social, the order is reversed. A constructionist view is that social and cultural beliefs are primary and cannot be separated from biological *knowledge* (Weatherall, 2002: 81). To support this theory, social constructivism uses Martin's (1991) study of the fertilization process. Furthermore, according to a social constructionist approach, gender is not a stable set of traits; rather, it is a social process created and renegotiated in interpersonal relationships and maintained through social activities. Applying this to speech styles, we may talk about *feminine* or *masculine*

speech styles thus referring to cultural associations with being a woman or a man and not to innate characteristics of being a female or a male. The social constructionists' thesis is supported by Hall's (1995) study of telephone-sex work and speech styles. Hall investigated the language used mainly by women in pre-recorded telephone-sex messages. Since the industry demanded for a sexy feminine person, feminine speech styles that were reminiscent of Lakoff's (1975) women's language were used. In addition, Hall interviewed telephone-sex workers who reported that customers were more satisfied when they used feminine speech styles. Nota bene, not all telephone-sex workers were women; however, both female and male workers used feminine speech styles without customers noticing any difference. Therefore, workers speech style was not a reflection of their gender identity; rather, their speech style created their gender identity.

Studying Hillary and Bill Clinton, Muir and Taylor (2009) reconceptualized gender by taking a number of facets into consideration. They (2009: 4) believed that "genders are constructions of social and cultural groups. They are institutions, consisting as all such entities do of boundaries, rules (prescriptions, proscriptions, built-in penalties and rewards) barriers and channeled interactions." They thought that gender is created in interaction; consequently, gender depends on relationships rather than one's characteristics. That being said, a speaker and his audience co-create the speaker's gender while communicating. Moreover, created gender may not be attributed to a single speaker. The analysis of Hillary and Bill Clinton results in Muir and Taylor (2009) suggested a joint gender relationship. In spite of individual acknowledged political careers, the Clintons have created a far more successful political brand reflecting a gendered political team that, as the name suggests, needs to be studied as a unit.

To summarize, women's or men's language is symbolic rather than a descriptive category. It is based on a complex interrelationship between one's sex and a gender identity, i.e. one's sense of self. People can develop their gender identity to match their biological sex. Females can adopt a set of social roles, behaviors and activities that are universally labeled as feminine gender roles, as well as males can adopt masculine gender roles. Or, they may negotiate and recreate their gender identity with respect to numerous factors such as audience, topic, communication aim, situation, etc. Identity can be created and expressed in different ways. Language is one of them, though a very powerful and productive. We will focus on studying the relationship of language and gender.

2.2 Language and gender theories

In order to situate this thesis within a theoretical framework, this chapter will provide a general overview of the main phases in the study of language and gender. Firstly, the deficit model introduced by Robin Lakoff in 1975, identifying women's language as powerless in comparison to men's, will be explained. Secondly, we will elaborate on Zimmerman and West's (1975) dominance approach built on the women's subordinate position in society. As a reaction to the dominance approach, gender differences in language were explained with the cultural difference approach (Maltz & Borker, 1982; Tannen, 1990) viewing men and women as two subcultures that developed different communication styles. Finally, critical rethinking has resulted in an array of anti-essentialist approaches viewing speakers as negotiators of their identities.

2.2.1 Deficit approach

Since the 1920s, linguists have shown a notable interest in the relationship of language and gender. One of the first who studied the issue was Otto Jespersen. He collected the information on how men and women spoke in terms of pronunciation, vocabulary and syntax and published the findings in his book entitled *Language; Its Nature, Development and Origin* (1922). He recorded gender differences on the case study of people of Caribbean descent. Socializing with immigrants, men were more successful in acquiring new vocabulary; hence, men's vocabulary was more extensive than women's. Jespersen believed that women receiving less technical education than men was the underlying reason for this difference. Secondly, he believed that women were more conservative and used traditional language. While men favorized course language between themselves, women used euphemistic expressions and even restrained themselves from using certain expressions which contain body parts. Also, men used alliterations, whereas women did not pay attention to acoustic properties. Women using half-finished sentences was explained with them speaking before thinking it through, hence, men were described to be more intelligent than women. To sum up, Jespersen characterized women's language as inferior compared to men's standard or

normal language. His study and male-centered, sexist and patronizing viewpoint was criticized by feminists.

Similar ideas were put forward in 1975 by Robin Lakoff. Her pioneering work *Language and Woman's Place* (1975) was extensively criticized because it lacked the empirical basis, i.e. her findings were based on her intuitions and peer anecdotal observations. Also, introspection was done on the corpus of white middle-class American women, which was inadequate for generalizations. The features she categorized as typical of women's language continued to appear in numerous later research. During her unsystematic observations, Lakoff recorded a number of phonological, lexical and grammatical features characteristic of women's language. Women used rising intonations in utterances where men used falling ones. Such sentences typically took the form of answers to questions but had the rising intonation typical of yes/no questions (e.g. *A: When will you be ready? B: Hm…around 5 o'clock…?*). Lakoff believed that such intonations required approval and confirmation from others. Secondly, when making lexical choices (especially colors and adjectives), men tended to use categories at the basic or generic level, while women used categories at the subordinate level. For example, women discriminated between the shades of *beige, lavender, aquamarine* which were absent from men's vocabulary. Women would use a different set of adjectives (*charming, divine, adorable*) than men (*cool, great, terrific*) to express their opinion on a subject. Further, women used hedges (*kind of, sort of*), polite forms (*would you mind, I would appreciate if you*) and wh-questions for imperative structures (*why don't you open the door?*) all of which was evidence for women's hesitancy. Also, women overused qualifiers (*I think, I mean*) and intensifiers (*so, very*). In terms of grammatical differences, women were said to use more question tags which were associated with tentativeness and insecurity.

The last hypothesis was challenged by several researchers (Dubois & Crouch, 1975; Cameron et al., 1989; Holmes, 1992). In Dubois and Crouch's dataset (1975), men used more question tags than women; however, it was not suggested that they were less confident speakers because of that. Moreover, Cameron et al. (1989) found that in some contexts, the usage of question tags was a marker of powerful rather than tentative speech. Holmes (1992) believed that question tags can serve as devices to maintain discussion or be polite. Furthermore, in their courtroom cases and witnesses' speeches study, O'Barr and Atkins (1980) studied the majority of Lakoff's hypotheses in a specific institutional context. They

suggested that the differences proposed by Lakoff were not necessarily a result of gender but of power. In order to prove their hypothesis, they used three men and women. The first man, an ambulance driver, and a 68-year-old housewife extensively used the features of women's language. The third pair, a female doctor and a policeman (expert witnesses), scored low on the mentioned features showing more power in their professional and private lives. Based on the results of their study, O'Barr and Atkins concluded that the features of women's language were neither features of all women nor limited solely to women. Rather, the cluster of those features was related to powerless people. Very frequently, a woman equaled a powerless person. However, since that might not always be the case, O'Barr and Atkins suggested for women's language to be changed into powerless language.

Lakoff was one of the first researchers to claim that the social role of women was to *talk like ladies* which included hypercorrectness, euphemisms, no joke telling, confirmation by nodding, etc. She claimed that girls were raised to learn special linguistic uses, i.e. a gendered way of communication. Women were not rewarded with acceptance in society, rather, this special speech style was later on used to keep them in a demeaning position (Lakoff, 1975: 5). If women tried to adopt linguistic features of the stronger group (men), they would be rejected by both men and women, which is a case of double-bindness. Therefore, Lakoff labelled women's speech style as a deficient and inferior to neutral men's style, hence the name of the approach.

Lakoff examined the representation of women in language. Women were more frequently referred to as *girls*, regardless of their age, than men as *boys*. While *master* had positive connotations, *mistress* usually invoked sexual (negative) connotations. The same applied to *bachelor*, which had desirable, and *spinster*, which had undesirable connotations. Men were always addressed as *Mr*., whereas women were defined in relation to their marital status *Miss*./*Mrs*. In terms or professional addressing, women were more likely to be addressed by their first name or by their first and last name, while men were usually referred by their last name or the title and the last name. These pieces of evidence inspired Lakoff to conclude that men were defined in terms of what they did in the world and women with whom they were associated (1975: 64).

Lakoff's *Language and Women's Place* is considered as the cornerstone of feminist linguistics despite the criticism of Lakoff's theory and politics. Lakoff adopted an androcentric viewpoint seeing women's behavior as a deficient variant compared to neutral men's behavior. So, it was implied that something was wrong with women's behavior and required remedies in order for women to be treated more equally in society. Indeed, many researchers who affiliated to the deficit approach (Crawford, 1995; Cameron, 1995a) worked on *language remediation* and tried to reinforce Jespersen's folklinguistic stereotypes. In spite of the methodological criticisms, it should be noted that Lakoff's work arose at a time the field had yet to establish itself and that, as Lakoff herself stated, it was not a definite account of gender differences in language but rather a road to further research.

2.2.2 Dominance approach

While Lakoff was developing the deficit model, other researchers tried to explain gender differences in language in a different way the most famous of which was Zimmerman and West's interruptions study (1975). Zimmerman and West recorded mixed-sex conversations at the University of California in 1975. The subjects were middle-class Caucasian young people. In 11 mixed-sex conversations, they recorded 46 interruptions by men and only 2 by women. Even though the research was done on a small and possibly unrepresentative sample, the authors concluded that more interruptions done by men pointed to men's dominance in conversation thus supporting the idea of men's more powerful position in society.

Inspired by Zimmerman and West's research, Beattie (1981) conducted his own by recording 10 hours of tutorial discussions. He found 557 interruptions compared to 48 Zimmerman and West's. Beattie found that men interrupted more than women; however, the difference was not statistically significant. Criticizing Zimmerman and West's research for an unrepresentative sample and possible skewness of the results (e.g. if one speaker disproportionally interrupted others), Beattie believed that interruptions were caused by status rather than gender.

Similarly, in a study of preschool children, Greif (1980) discovered that fathers were more likely than mothers to interrupt children and/or speak simultaneously with them. Also, both mothers and fathers more frequently interrupted daughters than sons. The research indicated that gender and power relationship from mixed-sex conversations could be replicated in spousal communication and parent-child interaction. Also, the girls' speech hindrance led to girls adopting stereotypical feminine passive communication roles.

The interruptions study was the most prototypical for the dominance approach; yet, not the only representative of it. Pamela Fishman (1977; 1978; 1980) studied some of Lakoff's hypotheses, namely question tags, and came to different conclusions. She asked Caucasian married American couples to record their home conversations and did her question tags analysis on 52 hours of recordings. Like Lakoff, Fishman noted that question tags were more commonly used by women (precisely, four times more); however, she offered a different interpretation. Fishman argued that in order to initiate or keep a conversation with their husbands, women had to do the *interactional shiftwork* – ask questions and use question tags to gain conversational power. Hence, question tags were not a sign of women's tentativeness. The same results were reported in the follow-up study by DeFrancisco (1989; 1991) who additionally interviewed the couples who had been taped and showed that her interpretations were in line with the couples' observations.

Another linguist advocated a radical view of language as structures that sustained men's power. Dale Spender (1980) was highly critical of the deficit approach. On the one hand, language was studied as an abstract system and on the other, it was studied within a given context. This separate analysis of, as she had put it, sexism in language as a system and sex differences in language, i.e. the separation of the form from its function, was inefficient. Spender decided to put these two notions together and provide a systematic analysis believing that an analysis needed a patriarchal order. For the sake of attributing meaning and its interpretation, rules were mandatory. Spender believed that rules were man-made and language was used to limit our world and construct the reality. Therefore, due to men being in the position of power and control to monopolize language, the myth of male superiority was created. The most vivid example of the encoded sexism was *he/man* language. Even though the generic *he* and *man* applied to both men and women, Spender believed that people automatically thought of a male person thus making women invisible. She also analyzed the

semantic aspect of the way men and women were addressed and provided evidence for masculine determiners seen as positive and feminine as negative or marked. For example, while *Sir* kept its associations with high society, *Madam* lost its dignity. The expression *She is a professional* had different (negative) connotations compared to *He is a professional*. Furthermore, Spender was critical of research that presented women as talkative. She believed that the desired state for women was silence. Hence, it was not that women were talkative in comparison to men; they were talkative in comparison to the desired state. Spender concluded that language needed to be liberated from men's control, which could be done with consciousness raising and women generating new meanings on the basis of multidimensional reality. However, this men bullying oppressed women view was criticized for its monolithic view of male power and ignoring other variables such as race and class which, in certain contexts, could give women more power (Talbot, 1998; Black & Coward, 1998; Goddard & Patterson, 2000).

One of the main criticisms of the dominance approach was that it portrayed women as powerless victims fighting against aggressive and powerful men when in fact those characteristics could be seen as successful communicative strategies (Coates, 1994: 73). Consequently, researchers started reassessing women's language searching for its strengths. Secondly, the dominance approach provided evidence and interpretation of gender differences in mixed-sex conversation; yet, the concepts of dominance and coercion were not as applicable in same-sex interaction. Based on the criticism of the dominance approach and the need for reevaluation of women's language, the difference approach arose.

2.2.3 Difference approach

The difference or subcultural approach arose as a reaction to the dominance approach. The first ideas were put forward by linguistic anthropologist John Gumperz (1981; 1982a; 1982b) who proposed a framework for studying issues in interethnic and intercultural communication. This approach assumed that individuals participated in communicative activities as cooperative agents, who were mutually interested in the accomplishment of the interaction. Hence, any miscommunication was explained in terms of differences in shared

understandings. However, this approach was criticized for its overly simplistic view (Kandiah, 1991).

Stemming from Gumperz's framework, Maltz and Borker (1982) explained six differences in conversational styles underlying miscommunication. One of the examples that best illustrated the basic idea of the gender differences in conversations was the different interpretation of positive minimal responses. Positive minimal responses included comments like *yes*, *yeah*, *aha*, *mm-hmm* or nodding. The responses were used by both men and women, though differently. Consequently, the differences might lead to miscommunications. For women, minimal positive responses meant *I am listening to you*, while for men, they had the meaning of *I agree with you* or *I follow your argument so far*. Hence, women used minimal responses more often than men. Misunderstandings may occur in mixed-sex conversations. Infrequent minimal responses by a male listener, a woman could interpret as a sign of him not listening to her, whereas to him, it only meant that he did not agree with her on everything. On the other hand, by giving frequent responses, a woman indicated listening, while a man would interpret that as agreeing with him. So, if later on a woman changed her mind, a man would see that as her frequent change of an opinion. This example explained one of the most common problems in mixed-sex communication – men could not conclude what women thought and women got upset with men who rarely listened to them. The second example was related to the meaning of questions. While men raised questions requesting for information, women used them as conversation maintenance devices. Thirdly, men frequently ignored the demand to link their utterance to the previous one thus underrecognizing another person's contribution. Men often ignored conversational flows and made abrupt topic shifts, which could be interpreted as a prerogative of power. Men perceived mentioning a problem as an opportunity to act as experts and offer advice, whereas women sympathized and shared their problems. To summarize, Maltz and Borker characterized women's speech as friendly and men's as uncooperative.

In addition to finding gender differences, Maltz and Borker explained the reasons which caused them. They believed that men and women formed sub-cultures with different sets of interactional rules. However, these sub-cultures were not formed in adulthood but in childhood, i.e. between the ages of 5 and 15, boys and girls socially interacted with members of their own sex. Girls played in small groups creating and maintaining relationships of

equality, intimacy and cooperation. Boys, on the other hand, played in larger, hierarchically organized groups asserting their position of dominance with a clear emphasis on verbal posturing. They also paid attention to the audience because the success of their performance was proportional to the size of their audience. Hence, the communicational patterns adopted in childhood were carried over into adulthood. Women's speech was of the collaborative and men's of the competitive nature because their conversational aims differed.

The cultural difference approach was popularized by Deborah Tannen's books *That's not what I meant* (1986) and *You just don't understand* (1990), which contained everyday examples to explain the hypothesis of miscommunication between men and women. We will use some of them for the illustration purposes. Similarly to Maltz and Borker's (1982) positive minimal response examples, miscommunication between men and women happened because of different underlying meaning of utterances. For example, if a wife during a car ride asked her husband if he would like to stop for a coffee and his answer was no, they would not stop. The wife, who had wanted to stop, would be annoyed believing her suggestion had not been taken into consideration. Simultaneously, the husband would be angry with his wife because she did not say that she wanted to stop. The reason for miscommunication was a different interpretation of the same interchange. The wife asked the question to initiate a negotiation process and not to get an instant decision, while the husband expressed his preference not making the final decision.

Based on the everyday examples, Tannen (1986; 1990) set up an essentialist approach, which viewed gender as fundamental part of an individual. This identity-oriented approach was anti-assimilationist, i.e. it did not assume that women wanted to be like men. Believing that the gender differences started in childhood, Tannen (1986; 1990) came up with six major gender differences, which we will elaborate on and exemplify in the following lines.

One of the dichotomies suggested by Tannen was *status versus support*. She used her own example for the illustration. Tannen and her husband worked in different cities and people often expressed their sympathy believing that that kind of life was difficult. While Tannen peacefully accepted people's sympathies, her husband was irritated and deemphasized the inconvenience giving a number of reasons which they benefitted from. He perceived people's sympathies as if they were looking down on them. So, in a world of a hierarchical

social order, a man needed to acquire and maintain a status since life was a struggle for independence. Tannen, on the other hand, perceived the world as a network of connections where people sought confirmation and support aiming at preserving intimacy.

Another dichotomy intertwined with the previous one was *independence versus intimacy*, which was reflected in men and women having different views of the same situation. Tannen described a case of a married couple Linda and Josh. An old high-school friend informed Josh that he would be in town the next month and Josh invited him for the weekend. When Josh informed Linda about it, she was upset because he had made plans without discussing it with her. To Josh, discussing the plans would mean seeking permission, which would imply that he was not independent, whereas to Linda, it would mean that her life was intertwined with someone else's. Both of them were upset – Linda for the lack of Josh's courtesy and a sense of a failure in their relationship and Josh for Linda limiting his freedom and controlling him. This happened because of men and women seeking different things – while women looked for closeness and support (intimacy), men were more concerned with status thus focusing on independence.

The third dichotomy was *advice versus understanding* exemplified on Eve's and Mark's story. Eve had a breast surgery and removed a lump from her breast. She believed that the stitches changed the looks of her breast. She found cutting into her body upsetting and shared her thoughts with her sister and a female friend. Both of them sympathized with Eve not offering any solution to her problem. However, when Eve shared her concerns with her husband Mark, he did not sympathize like her sister and the friend, but advised having a plastic surgery, which made Eve upset. She felt as if he was repelled by her looks and asked her to undergo another surgery, whereas he wanted to offer a solution to the problem. Furthermore, while Eve only wanted reassurance that it was normal to feel that way, Mark's suggestion implied that she did not have the right to feel that way but had to fix the problem. The problem was in the different purpose of the conversation – while women talked their problems through seeking for confirmation and support of ideas, men played the role of a problem solver offering pieces of advice. Men perceived problems as challenges, whereas women saw them as means of empathy.

The fourth dichotomy was *public versus private speaking* also known as *information versus feelings*. Public speaking was usually associated with men and private with women. To exemplify the idea, Tannen used a letter from an anonymous woman to a psychologist. The woman could not understand her husband's behavior of coming home from work and being extremely quiet. She explained that her husband was not a silent person especially when they had guests around when the husband was the life of the party. Moreover, during parties, the husband would tell jokes and retell work stories the wife wanted to hear and asked about. The psychologist explained that men rarely talked after coming home from work, while women, despite being equally tired, felt the need to share their thoughts, feelings and events that had happened at work. Tannen believed that men felt more comfortable speaking in public compared to women who enjoyed the private speaking. The underlying reason for this gender difference was in the purpose of their talks. For most women, a conversation's purpose was to establish rapport, i.e. to establish connections and negotiate relationships by sharing experiences. In comparison, men perceived talk as a means for preserving independence and maintaining a status in a hierarchy. This was done by storytelling, joking and showing different skills, i.e. by reporting. To paraphrase, women shared feelings and thoughts and men reported relevant information. Both women's and men's verbal behavior started in childhood – while girls criticized peers who wanted to stand out, boys learned how to get and keep attention in larger groups. To summarize, the crucial element was the perception of home. For men, home meant freedom from (verbally) proving themselves so they frequently remained silent. On the other hand, women perceived home as a means of sharing with their loved ones without worrying of being judged. So, women could not understand men who avoided this unjudged sharing, while men could not understand talking just to talk and not to share relevant information.

Furthermore, the next dichotomy, *orders versus proposals*, referred to the gender differences in the usage of the illocutionary speech act directives (Austin, 1962; Searle, 1969). Tannen noticed that women frequently started their sentences with *Let's* thus suggesting their husbands to do something. However, husbands did not interpret those as suggestions but rather as orders, which jeopardized their status in a family. This again could be related to a habitual conversational style of boys and girls. Researchers (Sachs et al., 1984; Andersen, 1984) studying preschool children during role-play of a doctor and a patient found gender patterns. While girls who played doctors gave their patients suggestions (*Let's sit down and*

use the medicine), boy doctors gave orders (*Lie down. Give me your arm.*). Similarly, in Smith's (1993; cited in Tannen, 1990: 75) sermons study at a Baptist seminary, men often gave orders to the audience (*Listen carefully as I read Luke, chapter seventeen*), whereas women used suggestions (*Let's go back to verses fifteen and sixteen*) inviting the audience to participate. Gleason's (1975) study showed that parents talked to their children in a different way. Precisely, fathers issued more commands than mothers and they issued them more to sons than daughters, i.e. boys were raised to be given more commands. The act followed by carrying out an action, men perceived as an order. Since men gained a status by telling others what to do and resisting being told what to do, they felt that their status and a dominant position were threatened. Trying to avoid conflicts, women used requests or suggestions, which men perceived as manipulation or a hidden directive, so the conflict was inevitable.

Finally, the last dichotomy was *conflict versus compromise* exemplified on Dora and Hank's car situation. Dora had to commute to work using cars that Hank had chosen and bought. She never complained even though she did not like some of the cars. After Dora almost died in a car accident, they were looking for a new second-hand car. Dora did not like Hank's choice and tried to persuade him to buy another car. Prior to the accident, she agreed on compromises but now was determined to get her way. Despite Dora's expectations, Hank did not say a word, which made Dora realize that occasional conflict and argumentative discussion were necessary. Trying to avoid conflict and agreeing to compromise, women gave men the right to think they were always right. Not being challenged and opposed to, men continued with their habitual style evoking feelings of frustration and dissatisfaction in women all of which could be resolved by flexibility of both men and women.

Tannen's hypotheses were confirmed in later research. Pilkington (1992) studied same-sex conversations in a bakery during a nine-month period. She found that women frequently agreed thus building on and completing each other's utterances, while men often disagreed challenging each other. Also, focusing on feelings and relationships, women talked to maintain relationships and affirm solidarity as opposed to men who engaged in verbal sparring. In comparison, Kupier's (1991) study on male solidarity proved that men indeed showed solidarity; however, they used insults to express it. Furthermore, Christine Howe's (1997) review confirmed the previous findings of boys' verbal predominance. She believed that gender differences began at socialization (ages 3 and 4). She confirmed Maltz and

Borker's (1992) minimal responses theory of women being more engaged listeners, which was realized by their responses, such as *uh huh* and *oh, dear*, i.e. back-channeling. Also, her research showed that girls requested help more often than boys, who were more likely to express their disagreeing views. Weatherall (2002) concluded that women's talk could be characterized as cooperative and men's as competitive. She confirmed some of Lakoff's hypotheses, namely that women used hedges, question tags and were less likely to interrupt a speaker.

The cultural difference approach was criticized for viewing miscommunication as misunderstanding, which was nobody's fault, and for failing to acknowledge power relations (Troemel-Ploetz, 1991; Uchida, 1992; Freed, 1992). Secondly, it failed to recognize gender similarities. In her construction of *genderlect*, Tannen (1990) worked on the populist genre and individual examples, which was criticized for overgeneralization and simplification. Further, Henley and Kramarae (1991) believed that the cultural difference approach could not explain all language differences and miscommunications. Rather, the approach was a powerful tool to maintain the male supremacy structure. Believing that the concepts of gender, language and power were intertwined, Uchida (1992) suggested the combination of the dominance and the deficit approach in constructing gender.

2.2.4 Anti-essentialist approaches

The fourth stage of gender and language research stemmed from the criticism of the essentialist approaches, hence the name. There were two sets of reasons for the anti-essentialist approaches. Firstly, instead of using one approach to interpret gender differences in language, a combination of poststructuralist approaches was applied. Secondly, researchers rethought the nature of gender and dismissed the polarization of gender.

The anti-essentialist approaches were built upon the criticism of the previous approaches that viewed gender as an integral part of an individual and separated it from interaction and social contexts of one's life (Bohan, 1993). One of the most influential anti-essentialist approaches was social constructionism which viewed gender as a central factor in building social identities (Fairclough, 1989; Davies & Harre, 1990; Ochs, 1993; Swan, 1993;

Crawford, 1995; Freeman & McElhinny, 1996) and discourse as a fundamental concept of social processes (Shotter & Gergen, 1994). A number of approaches were labelled as constructionist including ethnomethodology (Garfinkel, 1967; Cicourel, 1974; Sacks, 1992) and discursive psychology approaches (McKinlay & Potter, 1987; Edwards & Potter, 1992; Potter & Billig, 1992; Potter & Wetherell, 1987; 1998). Anti-essentialist researchers, drawing upon the ethnometodological approach developed in the conversation analysis, believed that gender is something one *does* rather than a fundamental trait one was born with. They interpreted conversation in terms of social contexts speakers emphasized in their talks. Sharing that view, discursive psychology also focused on the means mental phenomena was constructed and oriented in people's practices (Potter & Edwards, 2001: 90). Hence, the discursive psychology approach extended the ethnometodological view of orientation to action into cognitive states descriptions.

Judith Butler (1990; 1993) defined gender as a performative, social construct, i.e. gender was constituted by individual's acts. She believed that gender was "a set of repeated acts within a rigid regulatory frame which congeal over time to produce the appearance of substance, of a natural *"kind of being*" (1990: 33). Butler's performativity approach shared the social constructionist notion of individuals creating their social identities. However, that implied a certain degree of freedom, denied contexts and power relations and limited a possibility of power relations and social identities to be co-constructed during interaction, which it was criticized for (Cameron, 1997: 30-31).

In addition to Butler's (1990; 1993) ideas, Eckert (1989) and Eckert and McConnell-Ginet (1995) argued that gender constructs were embedded in other aspects of social life and correlated with other variables such as race, ethnicity, region, class, etc. To paraphrase, the idea of *the whole woman* (Eckert, 1989) could be realized with not isolating gender from other aspects of one's identity. This clearly indicated that the notion of gender polarization was abandoned in favor of the performativity nature of gender (Bergvall, 1996; Cameron, 1995b; 1996; Sunderland, 2004).

Another approach that addressed the problem of the isolation of gender from other aspects of social identity were *communities of practice* developed by Jane Lave and Etienne Wenger (Lave & Wenger, 1991; Wenger, 2000). According to the initiators, communities of

practice were formed by people who engaged in a process of learning based on some shared experience, e.g. doctors working on a medicine to cure cancer, engineers trying to build a robotic arm, a clique of students defining their identities, etc. However, not all communities are communities of practice. Three conditions need to be met for a community to become a community of practice. Firstly, a community of practice has an identity characterized by a shared domain of interest. To put it differently, members of a community of practice share competencies that distinguish them from other people. Secondly, members engage in activities, share information and learn from each other's experiences with the aim of pursuing the interest of their domain. Thirdly, a community of practice is not based solely on a shared interest. Rather, members share experiences, tools and knowledge with the final aim of using them in practice, i.e. they are practitioners. However, communities of practice differ from speech communities which view heterogeneity as structured by essential social categories (class, race, gender, age, ethnicity) and based in a geographically determined population (Labov, 1966; Wolfram & Shilling-Estes, 1969; Trudgill, 1974; Dittmar, 1976; Macaulay, 1977; Romaine, 1982; Coates, 1993; Kerswill, 1994; Mougeon & Beniak, 1996; Durant, 1999). Lave and Wenger's (1991) model was further developed by sociolinguists Eckert and McConnell-Ginet (1992). According to them, an individual who belongs to multiple communities of practice and gender are determining factors of one's membership within a community. In her study of communities of practice in secondary school, Eckert (1998) found that girls were more adjustable than boys, i.e. they used non-standard linguistic forms in communities of practice where they were more valued and standard in communities where they were more valued. She believed that the reason for this was that women were usually seen as interlopers and had to put an extra effort to prove themselves.

2.3 Critical Discourse Analysis approach

Influenced by the Frankfurt School and Michael Halliday's systemic functional linguistics, researchers decided to shift their research focus from single linguistic units to studying more complex social phenomena, which led to a new approach. Formerly known as Critical Linguistics (Fowler et al., 1979; Kress & Hodge, 1979; Trew, 1979), the Critical Discourse Analysis approach (hereinafter CDA approach) stemmed from text linguistics, sociolinguistics, rhetoric, pragmatics, anthropology and many other disciplines. The CDA

approach emerged in the early 1990s when the respected researchers van Dijk, Fairclough, Kress, Leeuwen and Wodak got together at the symposium and discussed different theories, methodologies and approaches. According to Fairclough and Wodak (1997: 258), CDA perceives language as social practice putting a special emphasis on the context of language use. Discourse is a form of social practice, which implies a relationship between a discursive event and an institution or a situation, i.e. the discursive event is shaped by them and it simultaneously shapes them. In contrast to traditional theories focused on detecting and explaining issues, CDA should work towards critiquing and changing society as a whole. Furthermore, CDA approaches are problem-oriented and interdisciplinary. They are trying to demystify ideologies and power by examining social domination, that is, the power (ab)use. CDA defines power as a systematic and constitutive element of society (Foucault, 1975, cited in Wodak & Meyer, 2001; Giddens, 1984). Power and language are connected since language can be used to assert and challenge power.

The most influential scholars affiliated to the CDA approaches are van Dijk, Foucault and Wodak and their main ideas will be presented in the following lines. One of the most cited CDA practitioner is Teun van Dijk who first started to apply the discourse analysis theory on media texts focusing on minorities and ethnic groups in Europe (1983). Van Dijk (1988) argued that a thorough analysis should not only be conducted on the structural and textual level but also include production and comprehension levels. As for the structural analysis, in addition to the study on phonological, morphological, syntactic and semantic level, van Dijk (1988: 2) called for the study of coherence, overall topics and themes and schematic forms and rhetorical dimensions of texts. To van Dijk, the production processes level includes institutional processes of news-making on the one and social and economic processes of structuring media discourse on the other hand. Finally, comprehension processes deal with memorizing and reproducing news information. A proper analysis would include all mentioned levels.

Furthermore, according to van Dijk (1995: 17), a discourse analysis is ideal for an ideology analysis because ideologies (including non-verbal semiotic messages) are usually expressed in discourse. Van Dijk suggested three aspects for ideologies analysis: discourse analysis (primarily text-based), social analysis (context-based) and cognitive analysis. For van Dijk, cognitive analysis consists of both social (mental representation of group members) and

personal cognition. Hence, ideologies influence the personal cognition of group members (van Dijk, 1995: 18) eventually leading to the establishment of *we versus they* dichotomy, which has been the central issue of the most van Dijk research (1988; 1993; 1995; 1998).

The second main approach in CDA is Faiclough's (1989). He believed that CDA should unite linguistics and social science into a theoretical and analytical framework. Like most CDA practitioners (Fowler et al., 1979; Hodge & Kress, 1979), central to Fairclough's analytical framework is systematic functional linguistics. Similar to van Dijk's three aspects for the analysis (discourse, cognitive and social analysis), Fairclough's three aspects are text, discourse and sociocultural practice. To put it differently, while van Dijk perceived social cognition as the mediator, Fairclough believed that the mediator was discourse practice. Fairclough's first analytical focus, text, involved linguistic analysis regarding sound system, vocabulary, grammar, semantics and cohesion organization (1995: 57). Secondly, discourse practice "straddles the division between society and culture on the one hand, and discourse, language and text on the other" (1995: 60). Intertextual analysis, which analyzed the text from the discourse perspective, was on the borderline between text and discourse practice. Therefore, descriptive linguistic analysis was supplemented with interpretative intertextual analysis. The third aspect, sociocultural practice, consisted of three parts, namely, economic (economy of the media), political (ideology) and cultural (values); however, the analysis should be carried out on the level relevant to a particular event (1995: 62).

The third CDA approach is associated with Ruth Wodak and the Vienna School of Discourse Analysis. CDA approach practiced by Wodak was discourse sociolinguistics which studied text in context and attributed both factors equal importance. Wodak examined various institutional settings (schools, courts, hospitals, parliaments, etc.) on a number of issues (sexism, racism, Semitism). Her 1991 anti-Semitism study led to the development of the *discourse historical method*, which integrated all available background information into the analysis. The usage of the historical perspective in the analysis was the feature that distinguished Wodak from other CDA practitioners. According to Wodak and Ludwig (1999: 12), by using the discourse historical method, language manifested and constituted social processes and interaction. This entailed three things: firstly, discourse involved ideologies and power, which were responsible for interaction; secondly, discourse was both synchronically and diachronically related to communicative events; thirdly, the interpretation of

communicative events depended on listeners' background knowledge and information. Therefore, the right interpretation did not exist (Wodak & Ludwig, 1999: 12).

The principles of the CDA approach could be summarized as follows:

- language is a social practice which represents the world;
- as a form of social practice, discourse constitutes other social practices;
- texts are assigned their meanings based on the relationship between texts and social subjects;
- texts acquire their meanings stemming from cultural, social and ideological contexts;
- in addition to the interpretation, CDA explains texts;
- discourse is used to produce and exercise power relations and ideologies.

In addition to the mentioned CDA practitioners, CDA has developed in new directions such as *Duisburg approach* (Link, 1983; Jäger & Maier, 2009), *Lesarten approach* (Maas, 1984; Januschek, 1992), *Loughborough approach* (Billig, 2003), *visual grammar* (Kress, 1993; Scollon & Scollon, 2002; Lemke, 2003), *corpus linguistics* (Caldas-Coulthard, 1996; Mautner, 2005) and *socio-cognitive approach* (Chilton, 2004; Koller, 2005). Since the other approaches are not related to our topic, we will provide a brief overview of Chilton's work focused on the analysis of political discourse.

Chilton's (2004) discourse analyses dealt with parliamentary language, politicians' speeches and political interviews in both British and international politics. He criticized the previous CDA approaches for underrecognizing cognition in analyses. Upon examining Grice's cooperative principle, Sperber and Wilson's relevance theory and Chomsky's generative linguistics, Chilton has come up with three principles. Firstly, language and political behavior can be based on cognitive endowment of the human mind rather than social practices. Language and social behavior are intertwined in innate mechanisms of the human mind. Finally, language is related to the cognitive ability to engage in criticism (2004: 28-29). Furthermore, Chilton introduced new concepts to explain his cognitive approach to political discourse. According to Chilton, *frame* was an area of expertise in a particular culture and was especially useful in studying metaphors. Another Chilton's concept was *discourse worlds*, i.e. mental spaces realized by an array of propositions. Chilton proposed a three-dimensional

analytical framework composed of *time*, *space* and *modality*. Everyone processing information would use those three dimensions. In spite of providing a valuable framework mixing discourse analysis, cognitive science and social theory, Chilton's work was criticized for being overly speculative and tentative in formulating his ideas.

2.4 Gender and language in the workplace

In this section, we will provide a brief overview of the most influential research of gender and language in the workplace and public contexts. Additional research will be used in subsections in the analysis part. Coates (1989) and Holmes (1992) described men's speech as competitive and verbally aggressive and women's as cooperative and other-oriented. However, conversation differs from speech in public contexts, which is more formal, has institutional procedures and is frequently oriented towards a particular goal. Holmes (1992: 134) believed that the talk strategies used by men (challenging, disruptive and assertive utterances) served the purpose of asserting power, which made them more confident speakers in public contexts.

One of the characteristics mainly associated with men's speech was speaking in longer turns and interrupting more. Even though the research was done in different contexts, such as faculty meetings (Westbrook Eakins & Eakins, 1976; Edelsky, 1981), school managers meetings (Case, 1988), televised political debates (Edelsky & Adams, 1990, Adams, 1992), committee hearings (Kathlene, 1994; 1995) etc., the studies showed that men tended to occupy the floor and interrupted others more than women, thus asserting their dominance. The researchers (Hall, 1985; Lakoff, 1990; Gal, 1991; Coates, 1994; Tannen, 1997) explained that the reason for the men's greater participation in public context speeches was in interactional norms–norms made according to men's speech since only men participated in public speeches until the 20th century. Therefore, men's speech was not the better but the only way to talk in public contexts (Lakoff, 1990: 210). In terms of Freed's (1996) gendered spaces theory, public contexts were male and private female spaces.

Public contexts were usually associated with asserting power and dominance, so some researchers studied the gender differences in enacting authority in public contexts. In the

analysis of doctor-patient interaction, West (1990) found that female doctors used suggestions thus reducing the status difference, whereas male doctors used orders thus reinforcing status differences. Additionally, Ainsworth-Vaughn's (1992) research showed that female doctors negotiated topic shifts, while male doctors did not. The research indicated that in addition to status, gender played an important role.

Judith Baxter (1999a; 1999b) was interested in behavioral differences between teenage boys and girls in classroom discussions. She found that dominant speakers (boys) spoke more, took longer turns and used humor frequently. They also verbally interrupted others and diverted attention from a speaker (usually a girl) by clapping hands or heckling, which eventually resulted in that girl not participating in the discussion anymore. Further, the dominant speakers were supported by the audience (positive minimal responses and nodding) they entertained with their jokes. Baxter (1999b) also found evidence of girls behaving competitively and boys cooperatively. Girls were more confident in informal discussion. To put it differently, girls' confidence was inversely proportional with the speech setting formality. Baxter believed that boys' dominant and girls' submissive behavior were expected. Moreover, if misbehaving, girls were penalized, whereas boys paradoxically might be rewarded.

2.5 Gender and language in political speeches

Politics has been described as the sphere of public life limited to men and more intensely masculine than other social practice (Brown, 1988: 13). The fact that men's linguistic style has been considered as a norm comes as no surprise. Nevertheless, with the exception of Wendy Brown's (1988) *Manhood and Politics* and Jeff Hearn's (1992) *Men in the Public Eye*, there have been few studies researching men's behavior in politics. Studies, mostly done by feminists, tended to focus on women's behavior and language trying to demarginalize women's position in politics.

Many researchers (Carrol et al., 1991; Thomas, 1994; Lovenduski, 1996; Norris, 1996; Hansen, 1997) claimed that women stressed different issues than men (children and the elderly, education, social issues), hence, the number of women in decision-making positions

needed to be increased for women's issues to be dealt with. Secondly, men outnumbered women in organizations, committees and assemblies in the majority of countries. In the study of the representation of women in politics in the European countries, Ruth Wodak (2008) found that Finland and Luxembourg had the highest and Italy, Greece and Spain the lowest number of female representatives. She believed that southern countries were male-oriented, while northern countries have had a long tradition of equality. The same findings were reported in Solheim's (2000) Scandinavian politics and Tamale (2000) South African politics study.

Numerous research have proven that male and female politicians behaved differently in various situations. As a follow up study of Edelsky and Adams (1990), Adams (1992) studied the gender differences in turn taking behavior in televised political debates. She found that women obeyed the debate rules (time and topic restrictions) more than men, who talked more than they were permitted and made uninvited interventions. Adams concluded that women appreciated obeying the rules thus presenting themselves as mannered politicians. What the author did not explain was how this obeying the rules benefitted women. This strategy of abiding by the rules might have served for the audience to create a better impression about female politicians.

Similarly, Kathlene (1994; 1995) studied the gender differences in turn taking and interruptions in Colorado State committee hearings. Her results showed that women spoke less, made fewer interruptions and took fewer turns than men. Additionally, she extended her research into the party affiliation, one's political years and interest in a topic. The results indicated that when in a chair position, men took the floor away from a speaker and made substantive comments more than women chairs did, which made Kathlene conclude that men chairs used their positions to assert power, while women chairs acted as facilitators (1995: 178). In comparison to women chairs who were first interested in witnesses' testimonies, men chairs first questioned the legislation sponsors. Also, female sponsors were questioned more than male sponsors.

Furthermore, by using a content analysis approach, Broughton and Palmieri (1999) examined the types of arguments male and female politicians used in euthanasia debates and discovered gender differences. The tone of the debate was against euthanasia. When speaking

for the bill against euthanasia, women based their arguments on the need to improve the palliative care and men on morality. Also, women used personal examples and references more than men did.

We provided a theoretical overview of the most influential language and gender theories in this section. Further analytical research on grammatical, lexical and punctuation categories and their results will be presented in individual subsections in section four.

3. Methodology

Following the explanation of the dissertation subject matter and the distinction between the concepts of sex and gender, we presented a theoretical overview of the dominance and the difference approaches. The follow-up anti-essentialist approaches, linguistic communities of practice and gender spaces approaches as well as the critical discourse analysis approach were elaborated on. Having covered the relevant theoretical approaches, let us focus on the empirical data of the research subject. This chapter presents the methodology of the research, followed by the chapters dealing with the research findings, data analysis and discussion thereof. In Section 3.1, an explanation as to why this research is important is provided. Section 3.2 presents a detailed stepwise description of the research design. Furthermore, Section 3.3 deals with the socioeconomic characteristics of the 113th Congress participants. The procedures of the corpus design are elaborated on in Section 3.4, accompanied by the tools used for the data analysis in Section 3.5. The normality distribution and graphical methods tests resulting in the choice of (non)parametric tests use for data analysis are provided in Section 3.6. Finally, the final research questions are outlined in Section 3.7.

3.1 The rationale

There are numerous reasons for conducting this research. Firstly, there are a number of studies on gender differences in language. These studies are of two kinds; they are either

quantitative or qualitative. Quantitative research papers in the said field provide statistical evidence of men and women differing in linguistic choices; however, they very often lack the qualitative part, i.e. they rarely explain the motivation for gender differences in language (Newman et al., 2003; Yu, 2013; Mulac et al., 2013). Furthermore, qualitative research papers usually fail to provide evidence of statistically significant gender differences in linguistic choices. The most famous study repeatedly criticized for basing assumptions on an insight rather than a systematic empirical approach is Lakoff's (1975). However, recent studies (Kacewicz et al., 2014; Pennebaker, 2013; Bell et al., 2012; Tausczik & Pennebaker, 2010; Pennebaker et al., 2003; Turner et al., 1995, etc.) successfully combine both approaches to analyze gender differences in language. Having said that, this dissertation will be composed of both the quantitative part that will provide statistical evidence of gender differences and similarities in language practices of American politicians, and the qualitative part that will elaborate on underlying reasons of the differences. Secondly, many research papers in the field are conducted on a small sample size, which renders them inadequate for making valid generalizations for the whole population. This research is conducted on a large sample of 3,655 speeches; 672 by female and 2,983 by male American politicians. Therefore, the generalizations of recorded language practices by the female and male politicians in the 113th American Congress should be considered as valid. Thirdly, larger sample research frequently takes one speech per speaker chosen by a simple random sampling method. Even though a simple random sampling method should create a representative view of the entire population, there is always a question what the results would have looked like had the sample been different. In order to leave no such questions unanswered, all the speeches by a single speaker containing at least 100 words were included in the analysis of this dissertation. Studying gender differences in language use is not a new topic in sociolinguistic studies, but we expect that a combination of the older dominance and difference approaches with a critical discourse analysis approach have the potential to shed some new light on the subject. Last but not least, this research will contribute to the previous and ongoing similar research in sociolinguistics by providing an analysis of the gender-based use of 70 linguistic categories, ranging from parts of speech such as pronouns, verbs, adverbs, prepositions, etc. to lexical items related to various categories such as money, religion, success, family, etc.

3.2 Research design

Table 1: *Research steps*

1) Selecting the field
2) Reviewing the literature
3) Establishing the objectives of the research
4) Defining the initial research questions
5) Compiling a corpus for a pilot study
6) Refining the corpus
7) Selecting a text analysis software program
8) Conducting the pilot study
9) Expanding and refining a final corpus
10) Conducting a computational data analysis
11) Identifying the variables to be studied
12) Uploading the results in the program for statistical analysis
13) Coding the data
14) Conducting a quantitative analysis of the data
15) Conducting a qualitative analysis of the data
16) Contextualizing the results
17) Evaluating the results
18) Drawing conclusions

Step 1) Selecting the field

The issue of exploring similarities and differences between men and women has always been an intriguing one and has inspired us to examine whether it extends to their linguistic behavior by investigating the differences in male and female political discourse of the 113th American Congress. We were interested in finding out which linguistic categories do the men and women working in the predominantly male field use differently and whether the recorded differences in the usage are statistically significant.

Step 2) Reviewing the literature

The thesis set off from a critical summary of the theoretical background on gender-based language use, which has led to pinpointing specific open questions and issues related to the main research task. Suggestions for further research from earlier studies and their limitations were taken into serious consideration during the research design stage. Having studied the current state of the art in the field of language and gender, we identified some desiderata and research gaps, which this dissertation hopes to fill.

Step 3) Establishing the objectives of the research

Starting from the hypothesis that linguistic practices of men and women differ, a working hypothesis claiming that linguistic choices and practices do not solely depend on gender but on working spheres and the corresponding discourses has been established. In order to test the working hypothesis, the following specific objectives were set up:

- To investigate whether the linguistics practices of the men and women in political discourse differ;
- To inspect if there is a statistically significant difference between the men and women in the usage of 70 variables;
- To determine which of the two respective groups use a specific variable more than another group;
- To interpret the results by providing underlying reasons of the variable usage difference.

Step 4) Defining the initial research questions

Upon establishing the specific objectives of the research, the initial set of research questions were formulated as follows:

- Is there a statistically significant difference in the usage of 70 linguistic categories tracked by LIWC between the male and female speakers on the corpus of the political speeches made in the 113th American Congress?
- Which linguistic categories are predominately used by the male politicians?
- Which linguistic categories are predominantly used by the female politicians?
- How can the determined differences be interpreted?

Step 5) Compiling a corpus for a pilot study

In order to conduct a pilot study, the official library of congressional record *Thomas* containing speech transcripts accessed at http://thomas.loc.gov/home/thomas.php was searched. The repository was searched and the speeches official transcripts downloaded from March to July 2014. The prerequisite for a speech to be analyzed with the software tool LIWC is that a speech is composed of at least 100 words. Since the search engine does not allow searching by word count, the repository had to be searched and checked for the number of words in each speech manually. For a more detailed description on the corpus compilation, see Section 3.4.

Furthermore, the pilot study was done in order to check the feasibility, identify potential methodological problems and try to come up with solutions before conducting a large-scale quantitative research. The 113th United States Congress was a legislative branch meeting of the United States federal government composed of 541 members. More precisely, at the time of conducting the pilot study, it was composed of 102 female and 441 male members. 102 women and randomly selected 102 men were chosen to be the participants in the pilot study. For the corpus of the pilot study, we chose one uninterrupted 100-word speech per speaker, i.e. 102 speeches by the female and 102 speeches by the male politicians were included in the pilot study. A random sampling method was used to choose the speeches.

Step 6) Refining the corpus

The first step before conducting the pilot study was to search for and download speech

transcripts that meet the pre-set requirement. However, the transcripts as such were not ready to be analyzed, i.e. they had to be adjusted for a computational analysis. The second step was to clean and adjust the corpus. The corpus cleansing procedure and examples are presented in Section 3.4.

Step 7) Selecting a text analysis software program

For the purpose of conducting a computational analysis of a large amount of data, a text analysis software program Linguistic Inquiry and Word Count was chosen. The software tool analyses a text on a word-by-word basis and calculates the degree to which people use different categories of words. LIWC analyses words categorized in 80 output variables presented as one line of data in a designated output file which can be further used in other software programs. The output with the computational analysis results from LIWC was uploaded in the program for statistical analysis SPSS where further statistical tests were done. The development, psychometric properties, framework, text processing module and a detailed list of words categorized in 80 variables are presented in chapter 3.5.

Step 8) Conducting the pilot study

The pilot study, conducted on a limited corpus, aimed to check the study feasibility, corpus size and identify potential research gaps. Since the men in the 113th Congress outnumbered the women, it was decided to level the number of participants. The corpus for the pilot study was composed of 204 speech transcripts.

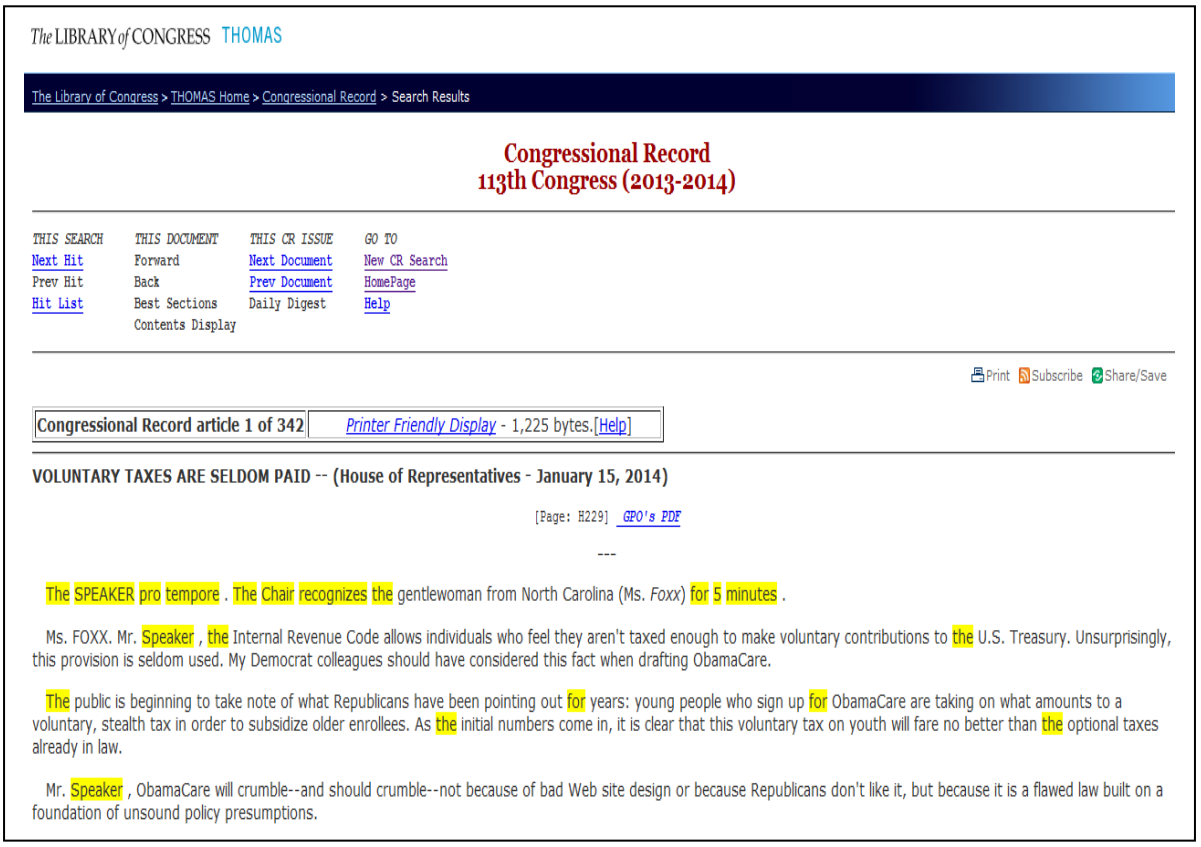
We decided to analyze the corpus for the pilot study using 70 out of 80 possible variables. The processing module, compilation of the dictionary and rating stages are elaborated on in chapter 3.5 and the relevant subchapters. The punctuation category was excluded from the study, i.e. periods, commas, colons, semicolons, question marks, exclamation marks, dashes, quotation marks, apostrophes, parentheses and other punctuation marks, because we chose to focus on grammatical and lexical categories rather than punctuation. Table 2 outlines the variables analyzed in the pilot study.

Table 2: *Output variable categories*

Category
1. Word count
2. Words per sentence
3. Dictionary words
4. Words longer than six letters
5. Total function words
6. Total pronouns
7. Personal pronouns
8. First person singular pronoun
9. First person plural pronoun
10. Second person pronoun
11. Third person singular pronouns
12. Third person plural pronouns
13. Impersonal pronouns
14. Articles
15. Common verbs
16. Auxiliary verbs
17. Past tense
18. Present tense
19. Future tense
20. Adverbs
21. Prepositions
22. Conjunctions
23. Negations
24. Quantifiers
25. Numbers
26. Swear words
27. Psychological processes
28. Social processes
29. Family
30. Friends
31. Humans
32. Affective processes
33. Positive emotion
34. Negative emotion
35. Anxiety

Category
36. Anger
37. Sadness
38. Cognitive processes
39. Insight
40. Causation
41. Discrepancy
42. Tentativeness
43. Certainty
44. Inhibition
45. Inclusive
46. Exclusive
47. Perceptual processes
48. See
49. Hear
50. Feel
51. Biological processes
52. Body
53. Health
54. Sexual
55. Ingestion
56. Relativity
57. Motion
58. Space
59. Time
60. Current concerns
61. Work
62. Achievement
63. Leisure
64. Home
65. Money
66. Religion
67. Death
68. Assent
69. Nonfluencies
70. Fillers

Conducting the pilot study was a valuable experience in learning about the organization of the official library of congressional speech transcripts. Firstly, each of the politicians participating in the 113th Congress had his/her own folder. A speaker can ask for the permission to address the House of Representatives in duration of either one or five minutes. All one-minute speeches had to be discarded because they were not long enough. This allowed for a more automatic identification of speeches. Namely it was possible to narrow down the search procedure, as shown in Figure 1, by searching for *The SPEAKER pro tempore. The Chair recognizes Mr./Ms. X for 5 minutes.*


The LIBRARY of CONGRESS THOMAS

The Library of Congress > THOMAS Home > Congressional Record > Search Results

Congressional Record
113th Congress (2013-2014)

THIS SEARCH	THIS DOCUMENT	THIS CR ISSUE	GO TO
Next Hit	Forward	Next Document	New CR Search
Prev Hit	Back	Prev Document	HomePage
Hit List	Best Sections Contents Display	Daily Digest	Help

Print Subscribe Share/Save

Congressional Record article 1 of 342 | Printer Friendly Display - 1,225 bytes.[Help]

VOLUNTARY TAXES ARE SELDOM PAID -- (House of Representatives - January 15, 2014)

[Page: H229] GPO's PDF

The SPEAKER pro tempore . The Chair recognizes the gentlewoman from North Carolina (Ms. *Foxx*) for 5 minutes .

Ms. FOXX. Mr. Speaker , the Internal Revenue Code allows individuals who feel they aren't taxed enough to make voluntary contributions to the U.S. Treasury. Unsurprisingly, this provision is seldom used. My Democrat colleagues should have considered this fact when drafting ObamaCare.

The public is beginning to take note of what Republicans have been pointing out for years: young people who sign up for ObamaCare are taking on what amounts to a voluntary, stealth tax in order to subsidize older enrollees. As the initial numbers come in, it is clear that this voluntary tax on youth will fare no better than the optional taxes already in law.

Mr. Speaker , ObamaCare will crumble--and should crumble--not because of bad Web site design or because Republicans don't like it, but because it is a flawed law built on a foundation of unsound policy presumptions.

Figure 1: *Screenshot of the adjusted search engine*

However, despite narrowing down the search and consequently speeding up the search process, all the transcripts in the speech repository were opened with the aim of testing the validity of the search methodology.

Furthermore, after the computational analysis with LIWC, the data were uploaded in SPSS for the further analysis. IBM SPSS Statistics version 20 provides a quick summary option of testing the null hypothesis stating that the distribution of a variable is the same between two groups of participants and suggesting to either reject or accept the null hypothesis. The corpus from the pilot study was used to run this test. The results suggested, with an exception of a small number of examples, that the null hypothesis should be accepted. To paraphrase, the results pointed out to no statistically significant difference in using 62 variables between the men and women. Such results raised the question – would the results have been different had the corpus been designed differently? We conducted a separate analysis using a different corpus yet applying the same methodology as described above. Precisely, we downloaded new speech transcripts by randomly chosen 102 different men and compared it to speech transcripts of 102 women. The results showed no statistically significant difference in the usage of 50 variables, i.e. 20 out of 70 variables were used differently by the men and women with a statistical difference. As the parallel screenshot in Figure 2 indicates, two different corpora we compiled for our pilot study reported different results regarding the number of variables used by the men and women with a statistical significance and the variables themselves.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Word count is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,061	Retain the null hypothesis.
2	The distribution of Words per sentence is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,253	Retain the null hypothesis.
3	The distribution of Six letter words is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,086	Retain the null hypothesis.
4	The distribution of Dictonary words is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,824	Retain the null hypothesis.
5	The distribution of Total functional words is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,047	Reject the null hypothesis.
6	The distribution of Pronouns is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,901	Retain the null hypothesis.
7	The distribution of Personal pronouns is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,101	Retain the null hypothesis.
8	The distribution of Pronoun I is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,026	Reject the null hypothesis.
9	The distribution of Pronoun we is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,228	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of WC is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,058	Retain the null hypothesis.
2	The distribution of WPS is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,297	Retain the null hypothesis.
3	The distribution of Sixltr is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,003	Reject the null hypothesis.
4	The distribution of Dic is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,672	Retain the null hypothesis.
5	The distribution of Funct is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.
6	The distribution of Pron is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,017	Reject the null hypothesis.
7	The distribution of Ppron is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,407	Retain the null hypothesis.
8	The distribution of I is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,832	Retain the null hypothesis.
9	The distribution of We is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,390	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
10	The distribution of Pronoun you is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,782	Retain the null hypothesis.
11	The distribution of Pronoun he she is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,386	Retain the null hypothesis.
12	The distribution of Pronoun they is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,400	Retain the null hypothesis.
13	The distribution of Impersonal pronouns is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,020	Reject the null hypothesis.
14	The distribution of Articles is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,005	Reject the null hypothesis.
15	The distribution of Verbs is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,180	Retain the null hypothesis.
16	The distribution of Auxiliary verbs is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,125	Retain the null hypothesis.
17	The distribution of Past tenses is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,111	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
10	The distribution of You is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,051	Retain the null hypothesis.
11	The distribution of He is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,489	Retain the null hypothesis.
12	The distribution of They is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,322	Retain the null hypothesis.
13	The distribution of Ipron is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.
14	The distribution of Articl is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,390	Retain the null hypothesis.
15	The distribution of Verb is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,030	Reject the null hypothesis.
16	The distribution of Auxverb is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,018	Reject the null hypothesis.
17	The distribution of Past is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,018	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
18	The distribution of Present tenses is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,798	Retain the null hypothesis.
19	The distribution of Future tenses is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,260	Retain the null hypothesis.
20	The distribution of Adverbs is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,124	Retain the null hypothesis.
21	The distribution of Prepositions is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,319	Retain the null hypothesis.
22	The distribution of Conjunstions is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,893	Retain the null hypothesis.
23	The distribution of Negations is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,872	Retain the null hypothesis.
24	The distribution of Quantifiers is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,762	Retain the null hypothesis.
25	The distribution of Numbers is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,312	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
18	The distribution of Present is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,183	Retain the null hypothesis.
19	The distribution of Future is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,785	Retain the null hypothesis.
20	The distribution of Adverb is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,021	Reject the null hypothesis.
21	The distribution of Preps is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,991	Retain the null hypothesis.
22	The distribution of Conj is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,027	Reject the null hypothesis.
23	The distribution of Negate is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,290	Retain the null hypothesis.
24	The distribution of Quont is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,690	Retain the null hypothesis.
25	The distribution of number is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,216	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
26	The distribution of Swears is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,317	Retain the null hypothesis.
27	The distribution of Social is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,107	Retain the null hypothesis.
28	The distribution of Family is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,039	Reject the null hypothesis.
29	The distribution of Friends is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,510	Retain the null hypothesis.
30	The distribution of Humans is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,122	Retain the null hypothesis.
31	The distribution of Affective processes is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,017	Reject the null hypothesis.
32	The distribution of Positive emotion is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,281	Retain the null hypothesis.
33	The distribution of Negative emotions is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,258	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
26	The distribution of Swear is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,434	Retain the null hypothesis.
27	The distribution of Social is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,626	Retain the null hypothesis.
28	The distribution of Family is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,002	Reject the null hypothesis.
29	The distribution of Friends is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,163	Retain the null hypothesis.
30	The distribution of Humans is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,201	Retain the null hypothesis.
31	The distribution of Affect is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,016	Reject the null hypothesis.
32	The distribution of Posemo is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,124	Retain the null hypothesis.
33	The distribution of Negemo is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,022	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
34	The distribution of Anx is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,180	Retain the null hypothesis.
35	The distribution of Anger is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,477	Retain the null hypothesis.
36	The distribution of Sad is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,002	Reject the null hypothesis.
37	The distribution of Cogmech is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,349	Retain the null hypothesis.
38	The distribution of Insight is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,114	Retain the null hypothesis.
39	The distribution of Cause is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,085	Retain the null hypothesis.
40	The distribution of Discr is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,236	Retain the null hypothesis.
41	The distribution of Tent is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,003	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
34	The distribution of Anxiety is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,175	Retain the null hypothesis.
35	The distribution of Anger is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,122	Retain the null hypothesis.
36	The distribution of Sadness is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,255	Retain the null hypothesis.
37	The distribution of Cognitive processes is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,513	Retain the null hypothesis.
38	The distribution of Insight is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,483	Retain the null hypothesis.
39	The distribution of Causation is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,225	Retain the null hypothesis.
40	The distribution of Discrepancy is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,157	Retain the null hypothesis.
41	The distribution of Tentative is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,020	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
42	The distribution of Certainty is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,152	Retain the null hypothesis.
43	The distribution of Inhibition is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,383	Retain the null hypothesis.
44	The distribution of Inclusive is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,437	Retain the null hypothesis.
45	The distribution of Exclussive is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,215	Retain the null hypothesis.
46	The distribution of Perceptual processes is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,611	Retain the null hypothesis.
47	The distribution of See is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,305	Retain the null hypothesis.
48	The distribution of Hear is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,395	Retain the null hypothesis.
49	The distribution of Feel is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,930	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
42	The distribution of Certain is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,680	Retain the null hypothesis.
43	The distribution of Inhib is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,184	Retain the null hypothesis.
44	The distribution of Incl is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,478	Retain the null hypothesis.
45	The distribution of Excl is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,020	Reject the null hypothesis.
46	The distribution of Percept is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,947	Retain the null hypothesis.
47	The distribution of See is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,300	Retain the null hypothesis.
48	The distribution of Hear is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,689	Retain the null hypothesis.
49	The distribution of Feel is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,347	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
50	The distribution of Biological processe is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,283	Retain the null hypothesis.
51	The distribution of Body is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,371	Retain the null hypothesis.
52	The distribution of Health is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,249	Retain the null hypothesis.
53	The distribution of Sexual is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,110	Retain the null hypothesis.
54	The distribution of Ingestion is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,146	Retain the null hypothesis.
55	The distribution of Relativity is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,133	Retain the null hypothesis.
56	The distribution of Motion is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,362	Retain the null hypothesis.
57	The distribution of Space is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,775	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
50	The distribution of Bio is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,216	Retain the null hypothesis.
51	The distribution of Body is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,986	Retain the null hypothesis.
52	The distribution of Health is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,278	Retain the null hypothesis.
53	The distribution of Sexual is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,281	Retain the null hypothesis.
54	The distribution of Ingest is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,395	Retain the null hypothesis.
55	The distribution of Relativ is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,815	Retain the null hypothesis.
56	The distribution of Motion is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,276	Retain the null hypothesis.
57	The distribution of Space is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,326	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Hypothesis Test Summary

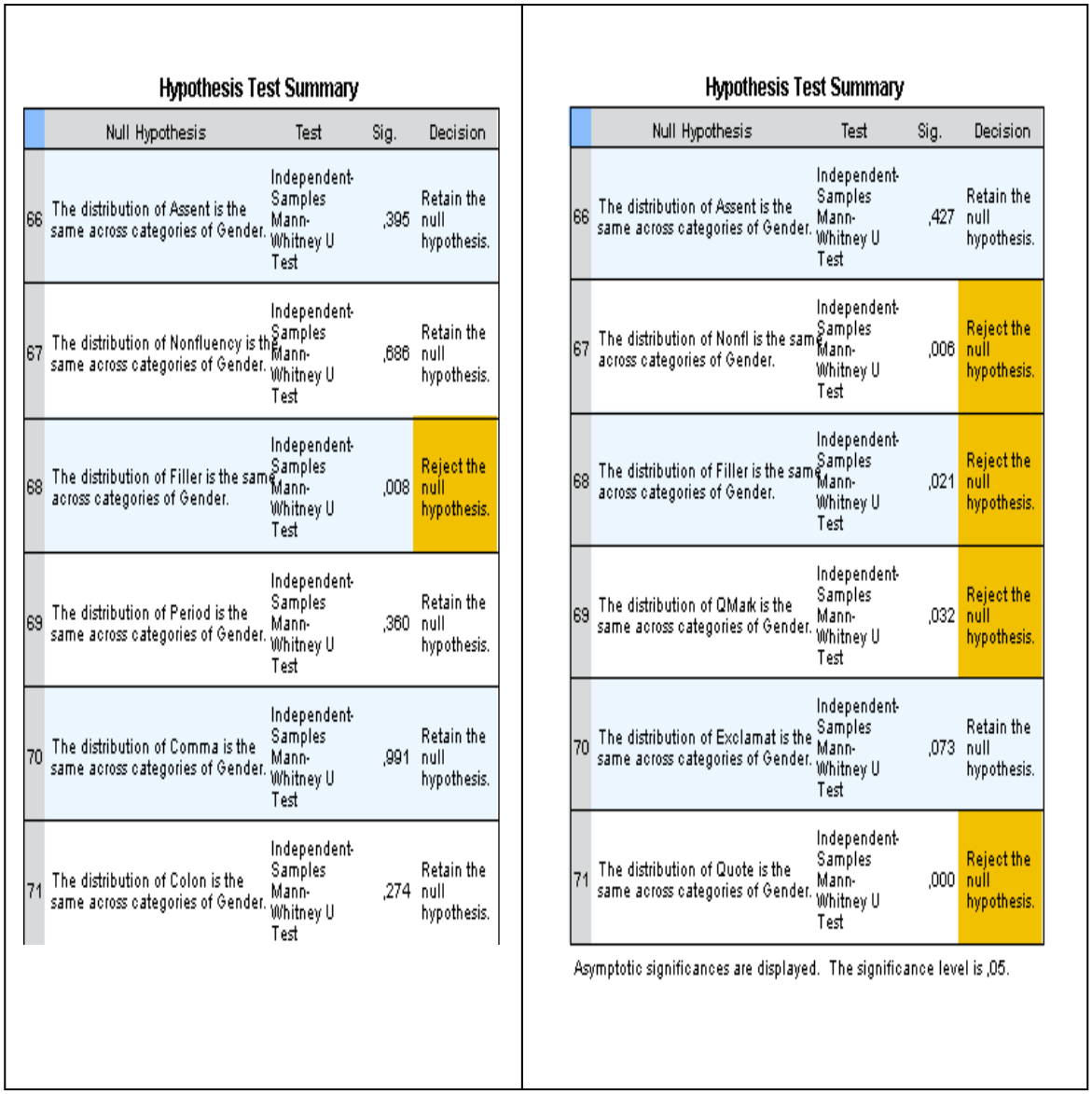
	Null Hypothesis	Test	Sig.	Decision
58	The distribution of Time is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,223	Retain the null hypothesis.
59	The distribution of Work is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,992	Retain the null hypothesis.
60	The distribution of Achieve is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,703	Retain the null hypothesis.
61	The distribution of Leisure is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,244	Retain the null hypothesis.
62	The distribution of Home is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,004	Reject the null hypothesis.
63	The distribution of Money is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,959	Retain the null hypothesis.
64	The distribution of Religion is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,664	Retain the null hypothesis.
65	The distribution of Death is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,837	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
58	The distribution of Time is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,985	Retain the null hypothesis.
59	The distribution of Work is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,147	Retain the null hypothesis.
60	The distribution of Achiev is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,032	Reject the null hypothesis.
61	The distribution of Leisure is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,130	Retain the null hypothesis.
62	The distribution of Home is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,007	Reject the null hypothesis.
63	The distribution of Money is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,221	Retain the null hypothesis.
64	The distribution of Religion is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,108	Retain the null hypothesis.
65	The distribution of Death is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,211	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.



Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
66	The distribution of Assent is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,395	Retain the null hypothesis.
67	The distribution of Nonfluency is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,886	Retain the null hypothesis.
68	The distribution of Filler is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,008	Reject the null hypothesis.
69	The distribution of Period is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,360	Retain the null hypothesis.
70	The distribution of Comma is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,991	Retain the null hypothesis.
71	The distribution of Colon is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,274	Retain the null hypothesis.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
66	The distribution of Assent is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,427	Retain the null hypothesis.
67	The distribution of Nonfl is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,006	Reject the null hypothesis.
68	The distribution of Filler is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,021	Reject the null hypothesis.
69	The distribution of QMark is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,032	Reject the null hypothesis.
70	The distribution of Exclamat is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,073	Retain the null hypothesis.
71	The distribution of Quote is the same across categories of Gender.	Independent-Samples Mann-Whitney U Test	,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Figure 2: *Screenshot of the hypothesis summary*

To summarize, the results of the pilot study suggested that the application of a simple random sampling method was not appropriate for this kind of research thus pointing out to the necessity of applying a different methodology when designing a corpus for the main research.

Step 9) Expanding and refining a final corpus

Having realized the limitations of the random sampling method for this kind of research, expanding the corpus was the next logical step. We applied a type of purposive sampling technique, namely total population sampling, and included all the speeches transcripts that meet the technical condition of at least 100 words in the final corpus. The organizational structure of the official library of transcripts accompanied by the selection methodology and the data cleansing procedures will be elaborated on in Section 3.4.

Step 10) Computational data analysis

After completing the design of the final corpus for the main research and data cleansing, the transcripts of the speeches made by each politician, stored in each politician's individual file in .doc file type, were saved in a special folder. The analysis was done by running the LIWC software, selecting *File* ⟶ *Process text* ⟶ *Select all* (documents stored in the folder) options as shown in Figure 3.

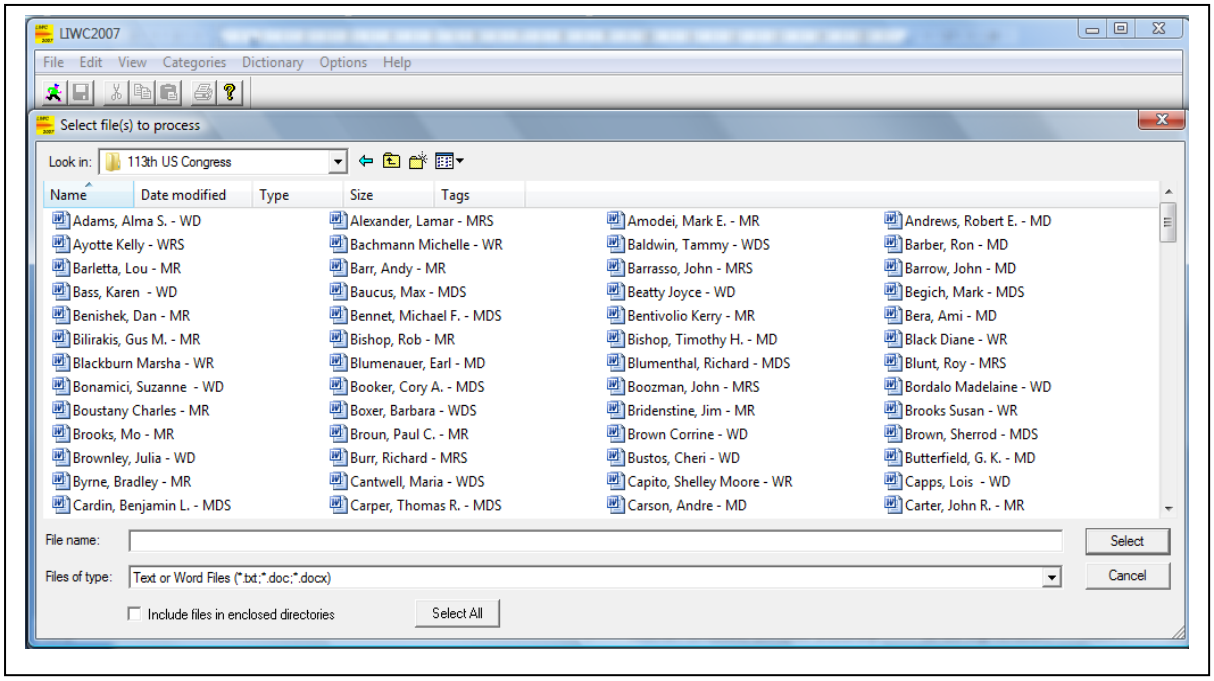


Figure 3: *Screenshot of the LIWC computational analysis*

Given the fact that a total number of pages from the individual files was 5,504, a computational data analysis with the LIWC software lasted for approximately half an hour. The output in .txt file type was afterwards uploaded in SPSS.

Step 11) Identifying relevant variables

LIWC analyses a text and provides an output in 80 variables which include 4 general descriptor categories (total word count, words per sentence, percentage of words captured by the internal dictionary, and percent of words longer than six letters), 22 standard linguistic dimensions (percentage of words in a text that are pronouns, articles, verbs, prepositions, numbers, etc.), 32 word categories tapping psychological constructs (social, affective, cognitive, perceptual and biological processes), 7 personal concern categories (work, achievement, home, money, religion, leisure activities, etc.), 3 paralinguistic dimensions (assents, nonfluencies and fillers), and 12 punctuation categories (periods, question marks, parenthesis, etc.) (Pennebaker et al., 2007: 4). The pilot study and further corpus design for the main research inspired us to include some of the variables from the punctuation category in the analysis. We wanted to study which of the two groups of participants (men or women) asked more questions, gave more orders or expressed strong feelings about something and supported their claims by quoting different sources. Therefore, in addition to the variables analyzed in the pilot study outlined in Table 2, question mark, exclamation mark and quotation mark from the punctuation category were included in the analysis.

Step 12) Uploading the corpus results in SPSS

The output was then saved in .csv file type and uploaded in *SPSS Statistics version 20.* A number of steps needed to be done in order to upload the data correctly. The steps in questions and the instructions to be followed written in italic are as follows:

1. Does your file text match a predefined format? – *No* ⟶ *Next*
2. How are your variables arranged? – *Delimited*
3. Are variable names included at the top of your file? – *No* ⟶ *Next*

4. How are your cases represented? – *Each line represents a case*
5. How many cases do you want to import? – *All of the cases* ⟶ *Next*
6. Which delimited appear between variables? – *Tab, Semicolon*
7. What is the text qualifier? - *None* ⟶ *Next*
8. Variable name: - *V1*
9. Data format – *String* ⟶ *Next* ⟶ *Finish*

Step 13) Coding the data

The data can be categorized into two types of variables, namely categorical and continuous variables. A categorical variable (also known as a nominal variable) has two or more categories which do not have intrinsic ordering. For example, gender is a categorical variable which has two categories (men or women). The parties American politicians affiliate themselves with (Democrats and Republicans) are also a categorical variable. A categorical variable is the one which allows a researcher to assign categories without ordering them. Hence, the participants' gender from our study is a categorical variable which requires coding. Value 1 was assigned to the men and value 2 to the women, after which the respective values were attributed to each politician manually. Other categorical variables (House, party, education level and ethnicity) were also coded and the codebook is provided in Table 6.

In comparison, continuous variables (also known as quantitative variables) can be further classified as interval or ratio variables. Interval variables have a numerical value and are measured along a scale. To compare, ratio variables are similar to interval variables with an addition condition of zero, i.e. zero means that there is none of that variable. For example, the temperature in Celsius is an interval, whereas the one in Kelvin is a ratio variable because zero in Celsius do not mean there is no temperature while in Kelvin it does. The variables from the LIWC analysis are numerical, more precisely ratio scale. The output is provided in a two-digit number with two decimal points. Even though these variables can be automatically coded, ratio variables should not be coded because they provide more detailed information when being used in their numerical form. If we code LIWC results, we might get as many codes as we have participants (since a code is attributed to equal values), which will make the

analysis extremely difficult and probably skew the data. Therefore, we did not code the LIWC data output.

Finally, the string data format had to be chosen to upload the data correctly into SPSS, as mentioned in step 9 of the previous subsection. However, once the data was uploaded into SPSS, the string data format had to be converted into the numerical data format in order to perform further tests.

Step 14) Quantitative data analysis

The Mann-Whitney U and the Spearman correlation tests were done presenting the results in the forms of tables and charts. The main questions to be answered by the quantitative analysis are: Is there a difference between the male and the female politicians from the 113th American Congress in using different word categories? Is the difference statistically significant? Does a certain word categories usage correlate with the other word categories usage? Is the correlation statistically significant? Is the correlation positive or negative?

Step 15) Qualitative data analysis

The main aim of the qualitative data analysis is to look for the motivation behind the different word categories usage by the male and the female politicians from the 113th American Congress. In order to do that, the following questions are to be answered: Why do the male or the female politicians from the 113th American Congress use a certain word category significantly more than the other gender group? How can we interpret the statistically significant differences in a word categories usage? What is the difference between these research findings and previous research findings in the respective field? Which conclusions can we draw about language use and gender in political discourse?

Step 16) Contextualizing the results

Since the field abounds in studies of gender-related language use, the results of the previous related studies will be compared to the results of this study. We will examine whether the results differ and how, and try to offer a possible explanation for the differences, thus setting our study in the wider context of related research.

Step 17) Evaluating the results

Before making conclusions, the results of the study will be evaluated. The evaluation can be done by checking whether the research questions have been answered and consequently whether the aims of the study have been fulfilled. By doing that, we will be able to decide whether the study has been successfully conducted or not.

Step 18) Drawing conclusions

Quantitative and qualitative research results will be used to draw conclusions on gender and language use in political discourse. We will also make desiderata and provide recommendations for future research.

3.3 Participants

A summary of the 113th Congress participants categorized in several groups will be presented in Section 3.3 while a full version of demographic characteristics listed for each participant in Congress will be provided on a CD attached to this dissertation.

The United States Congress is composed of 435 voting Representatives and 100 Senators, 2 from each state. However, the number and participants themselves may change

during 2-year Congress duration due to various reasons such as death, maternity/paternity leave, retirement, etc. We decided to include, if their speeches meet a technical prerequisite, all participants who at one point served in the 113th Congress.

The 113th American Congress was composed of 553 individuals from 50 states including American Samoa, Guam, Puerto Rico, the District of Columbia, the Northern Mariana Islands and the U.S. Virgin Islands. The list of all politicians from the 113th Congress sitting in the Senate categorized by their respective states is provided in Figure 4 followed by the politicians sitting in the House of Representatives in Figure 5.

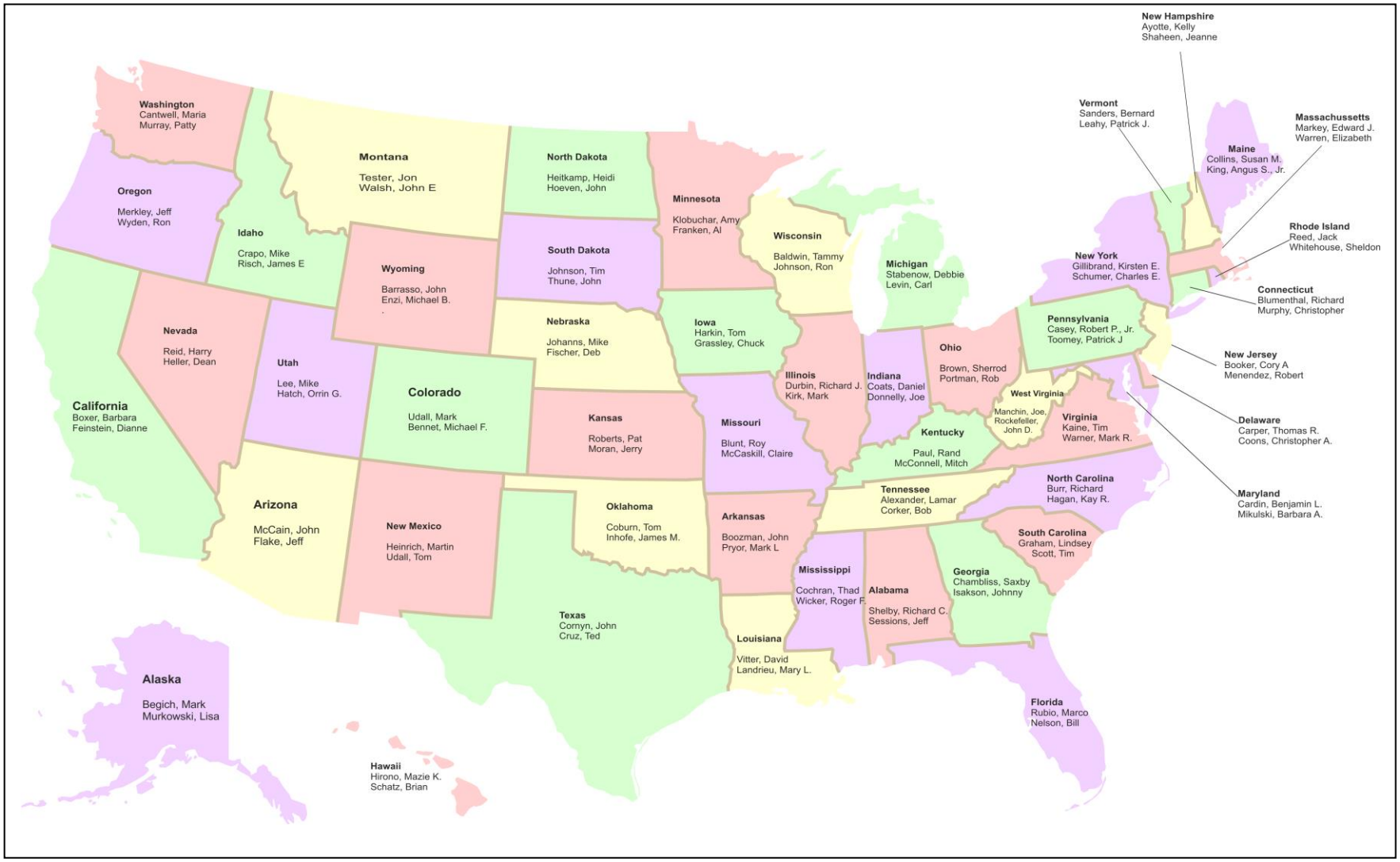


Figure 4: *Geographical distribution of politicians from the Senate*

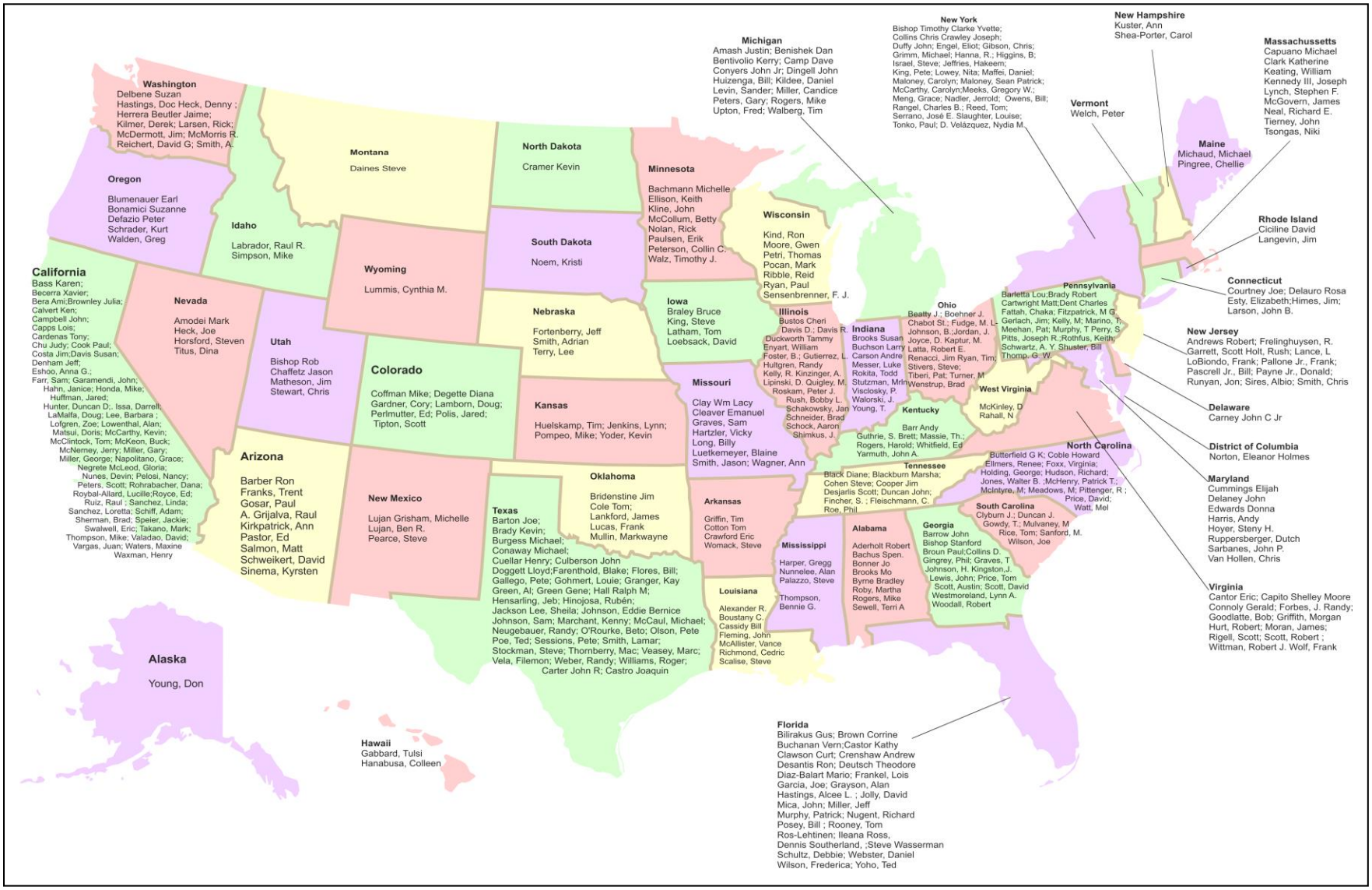


Figure 5: *Geographical distribution of politicians from the House of Representatives*

The 113th US Congress was composed of a record number of female participants. More precisely, 103 women and 450 men at one point participated in the 113th Congress, as illustrated in Figure 6.

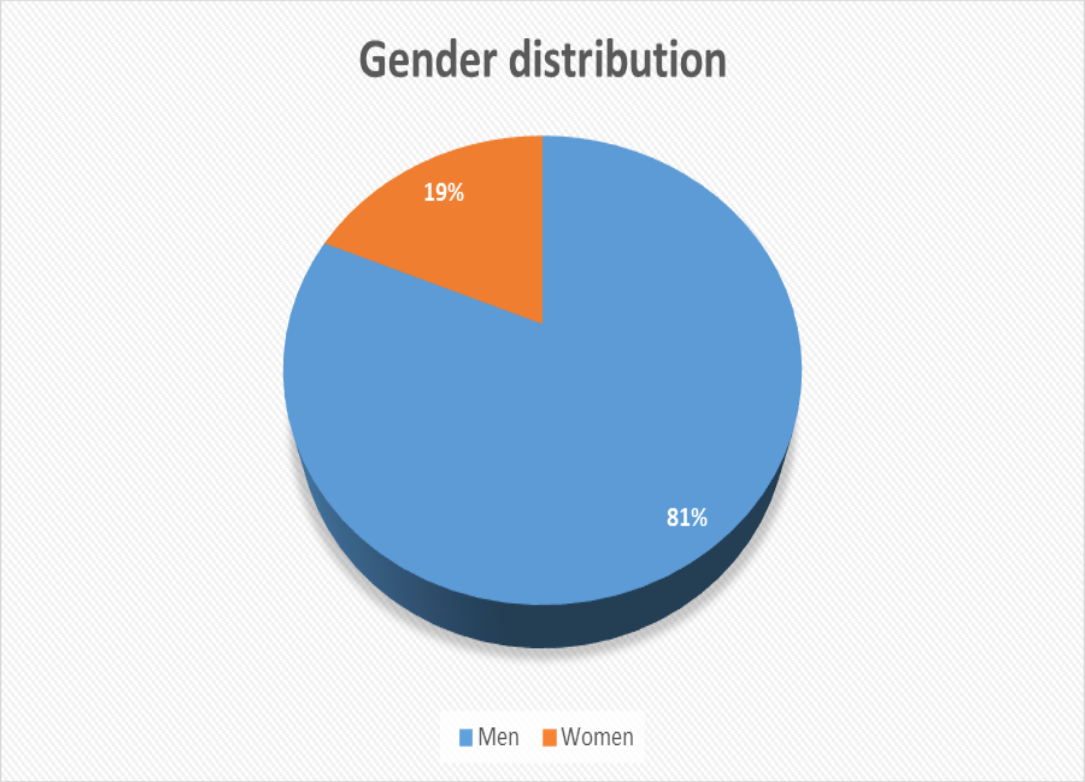


Figure 6: *Distribution of politicians based on gender*

Based on the 2012 election results, the House of Representatives majority was held by the Republican and the Senate majority by the Democratic Party. As illustrated in Figure 7, 80 women Democrats, 187 men Democrats, 23 women Republicans and 263 men Republicans were sitting in the 113th Congress.

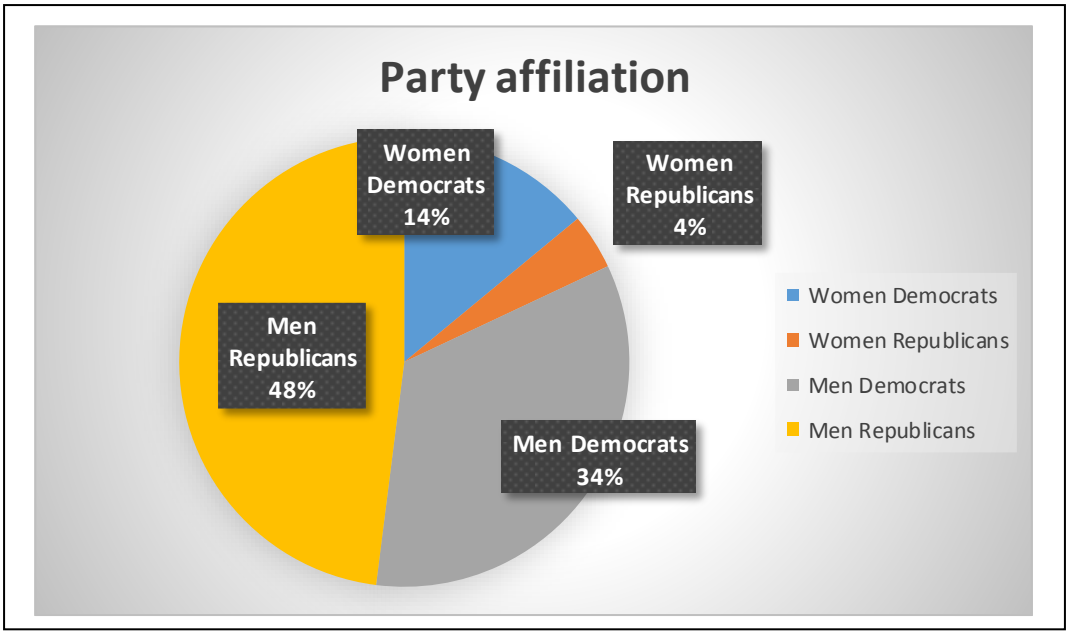


Figure 7: *Distribution of politicians based on party affiliation*

As a bicameral legislature, the United States Congress is composed of two chambers, namely the House of Representatives and the Senate. As shown in Figure 8 below, the House of Representatives was composed of 64 women Democrats, 146 men Democrats, 20 women Republicans and 220 men Republicans. Furthermore, the Senate was composed of 16 women Democrats, 41 men Democrats, 4 women Republicans and 42 men Republicans.

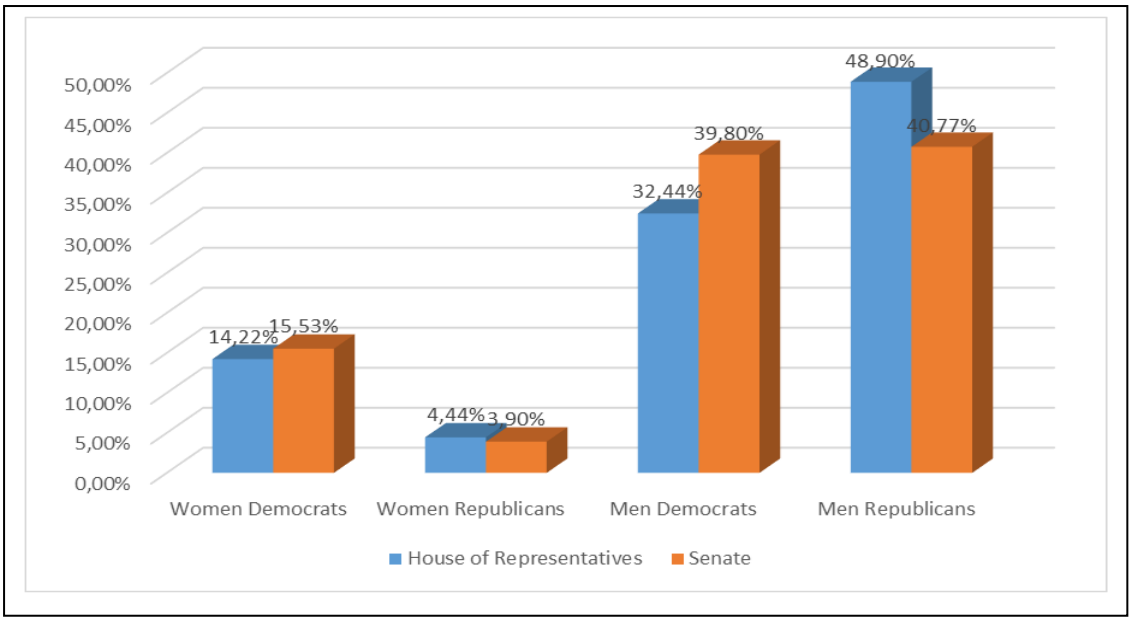


Figure 8: *Distribution of politicians based on chambers seats*

Furthermore, there are certain limitations regarding the age of politicians. The US Constitution requires for a Representative to be at least 25 years old at the time of taking office. At the beginning of the 113th Congress, the youngest Representative was 29-year-old Patrick Murphy. In comparison, the oldest Representative in both the 113th Congress and US history was 89-year-old Ralph Hall. When taking office, a Senator has to be at least 30 years old. The youngest Senator in the 113th Congress was the 39-year-old Christopher Murphy and the oldest 79-year-old Dianne Feinstein. The average age at the beginning of the 113th Congress was 57 years in the House of Representatives and 62 years in the Senate being among the highest of any US Congress.

As reported in the biographies on the politicians' official websites, 1 woman and 16 men had only secondary education. 21 women and 66 men obtained a Bachelor's degree and 72 women and 348 men a Master's degree. There were 9 women and 20 men who had a PhD.

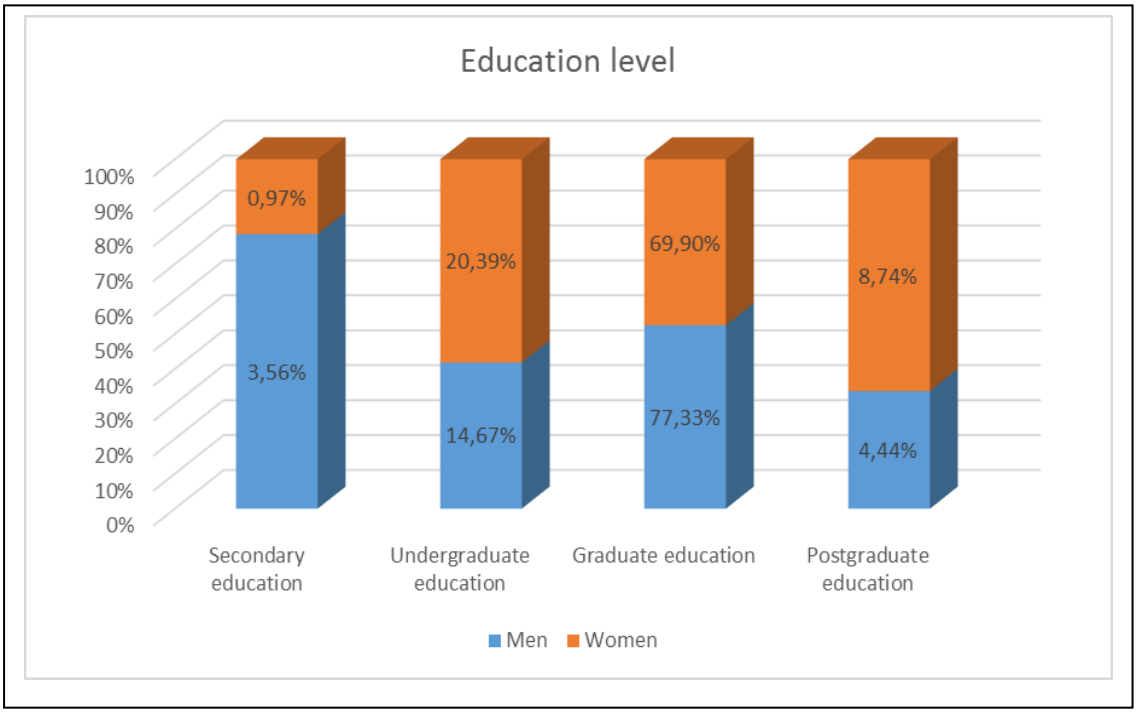


Figure 9: *Distribution of politicians based on education level*

According to the report of the *CQ Roll Call Guide to the New Congress* (Manning, 2014), the main professions of the politicians in the 113th Congress were business, education, law and public service/politics, as shown in Figure 10. However, it should be noted that most

politicians listed more than one profession and that the listed professions were not necessarily the ones practiced prior to entering the 113th Congress.

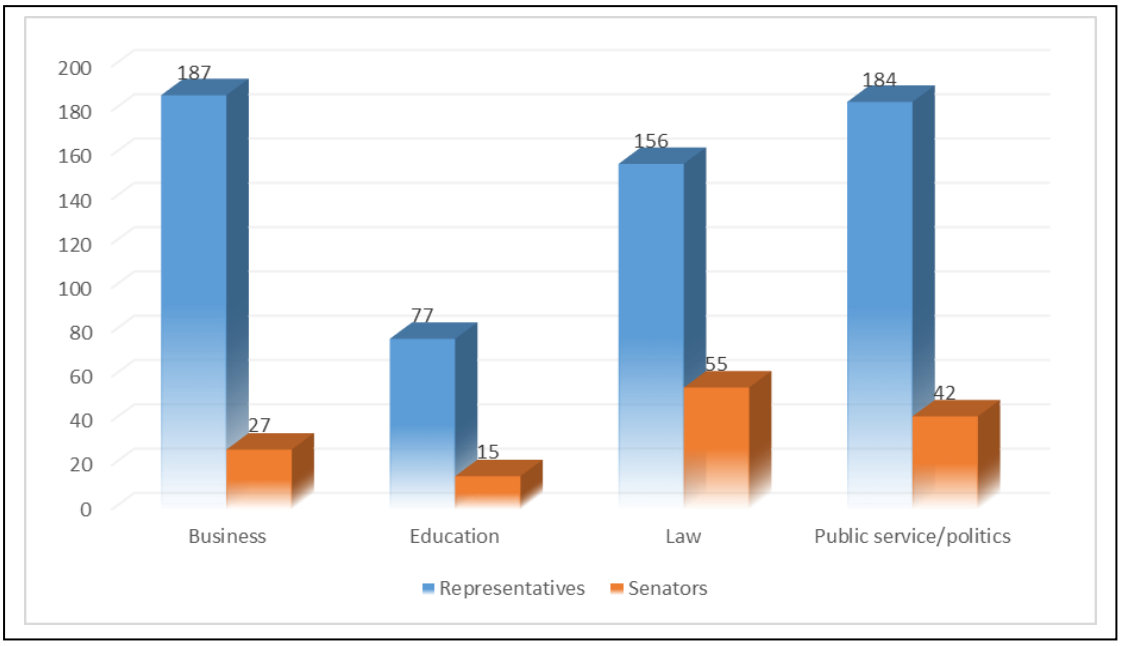


Figure 10: *The most frequently listed occupational categories*

The 113th Congress has so far been the most demographically diverse. Not taking into account the Caucasians, who are the majority, ethnic minorities are categorized in Figure 11.

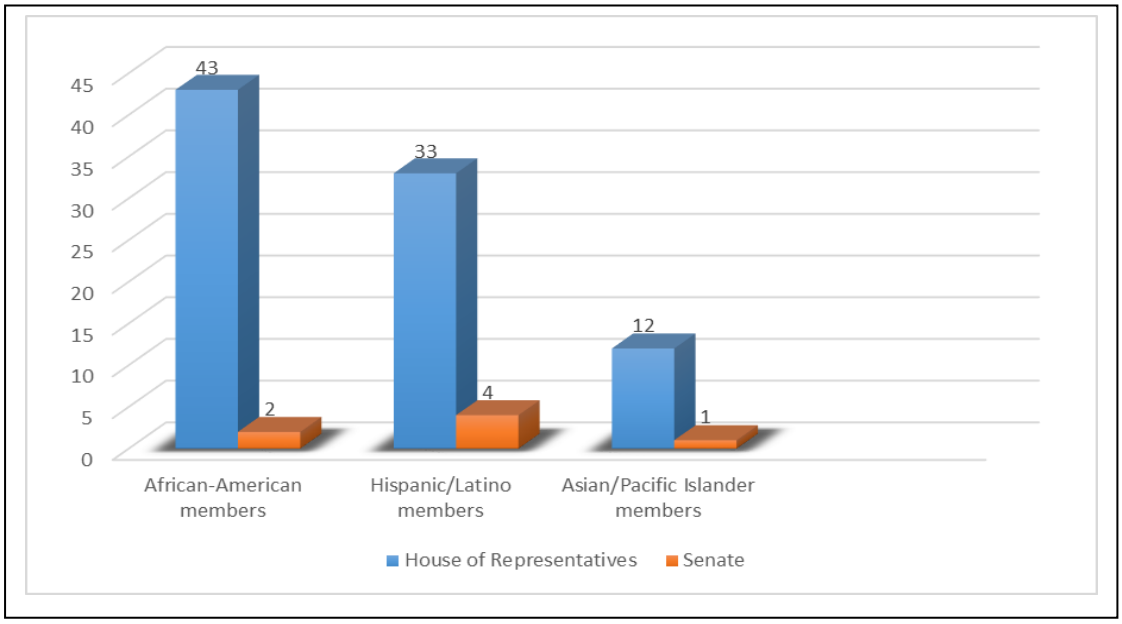


Figure 11: *Distribution of politicians based on ethnic minority*

3.4 Corpus

The official record of the proceedings and debates of the United States Congress *Thomas* found at http://thomas.loc.gov/home/thomas.php was browsed from March, 2014 till January, 2015. The Congressional Record contains a full record of proceedings from both chambers of the United States Congress – the House of Representatives and the Senate. The Government Printing Office publishes new issues of the record daily, which become available on *Thomas* the following morning. Since the 113th Congress was the meeting of the legislative branch from January 3, 2013 until January 3, 2015, the last access and download could not have been done before January 4, 2015.

As provided in the About section on the Thomas website, each daily issue in the Congressional Record consists of proceedings grouped in four categories, namely House of Representatives, Senate, Extension of Remarks and Daily Digest. House of Representatives and Senate contain transcripts of debates and statements made on the floor of each of the two chambers. They also contain roll call votes, petitions, amendments, memorials and records of various parliamentary actions. Furthermore, Extension of Remarks consists of additional statements not made on the floor accompanied by speeches delivered outside Congress, letters, tributes and various articles. Daily Digest is a section which provides a summary of each day's activities. It typically contains Highlights, Senate Chamber Actions, Senate Committee Meetings, House Chamber Actions, House Committee Meetings and Committee Meetings Scheduled for the Following Day. Therefore, given the nature of the texts in Daily Digest, i.e. summaries written in a headlines form with no specific author/speaker, the texts were not included in the analysis. Since the texts in the Extension of Remarks section are not transcripts of speeches made on the floor, there is a possibility that they were written by someone else (e.g. a politician's spokesperson or a secretary). Moreover, the Extension of Remarks section includes letters, newspaper and magazines articles written by common citizens or scientists. Those texts cannot be attributed to any politician, so the transcripts from Extension of Remarks were not included in the analysis. As far as the transcripts from the House of Representatives and the Senate sections are concerned, every uninterrupted 100-word speech transcript attributed to a certain politician was included in the analysis.

In addition to the sections of the Congressional Record presented above, there are two additional categories, namely the one for the politicians sitting in the House of Representatives and another for those sitting in the Senate with the politicians categorized respectively. Each politician sitting in Congress has his/her own folder where all the speeches, debates, votes, amendments, etc. After studying the repository, a total of 56,360 hits were recorded; 11,306 by the female and 45,054 by the male politicians. However, as stated above, those hits included everything related to a politician's name. Therefore, each of 56,360 hits had to be opened with the aim of selecting uninterrupted speech transcripts among other hits previously excluded from the analysis and explained.

Politicians sitting in the House of Representatives can ask Mr./Madam Speaker for the permission to address the House in duration of either 1 or 5 minutes. Since none on the one-minute speeches contained 100 words, they were not included in the final corpus. Furthermore, every five-minute speech had to be manually opened and checked for the condition of interruption, on the one hand, and the number of words, on the other after which all uninterrupted 100-word five-minute speeches were included in the corpus. Compared to the politicians sitting in the House of Representatives, the politicians sitting in the Senate ask Mr./Madam President for the permission to address the Senate; however, there is no time limitation. The procedure is the same: every hit had to be opened manually to check for the word-count and interruption. Table 3 shows the number of the speeches from both the House of Representatives and the Senate with respect to gender and party affiliation included in the analysis.

Table 3: *The number of the speeches included in the corpus*

Gender	**Party**	**Number of the included speeches from the House of Representatives**	**Number of the included speeches from the Senate**	**Total number of the included speeches**	**Total number of the included speeches**
Female	Democrats	255	217	672	3,655
Female	Republicans	110	90		
Male	Democrats	546	1,029	2,983	
Male	Republicans	568	840		

Every speech is marked for the official name attributed to it according to a topic it covers, specifies a chamber where it was made and the exact time when it was delivered. Due to length restrictions, a list of all 3,655 speeches included in the analysis with the relevant information is provided on the CD attached to the dissertation.

We started from the premise that all 553 participants will have at least one 100-word speech that will be included in the corpus of the study. However, after reviewing the repository and checking the validity of each speech transcript, we ended up with 395 participants and their 3,655 respective speeches (details of which can be found on the CD).

3.4.1 Data cleansing procedures

Once a speech transcript is downloaded from the repository, it needs to undergo data cleansing procedures. Firstly, LIWC software tool cannot discriminate between lower and upper case graphemes; therefore, those adjustments are not required. One needs to bear that in mind when setting up goals of a computational LIWC analysis. Secondly, misspellings, colloquialisms, foreign words and abbreviations are usually not in the internal LIWC dictionary, hence spelling errors should be corrected to a standard American or British English spelling as suggested by a *Word* spellchecker. Meaningful abbreviations such as *Dec* for *December* should be spelled out. Common verb contractions such as *I'm, she's, aren't, doesn't, couldn't, we'll, shouldn't*, etc. are integrated in the LIWC dictionary.

One of the categories LIWC analyses is a *Words per sentence* category which, as the name suggests, counts a number of words in a sentence based on the end of a sentence markers (periods, exclamation marks and question marks). This is the most problematic LIWC category because all abbreviations (e.g. Mr., Ph.D., U.N.) are counted as multiple sentences unless periods are removed. One needs to carefully go through a text and remove all unnecessary periods. 3,655 speech transcripts, i.e. a total of 5,504 pages were reviewed and adjusted accordingly. The list of the adjustments done in the corpus is provided in Table 4. Naturally, one may find different examples, though of the same kind, in one's own corpora and should clean data as instructed.

Table 4: *List of adjustments*

Original entry	Adjusted entry
Mr./Mrs./Ms.	Mr/Mrs/Ms
Jr./Sr.	Jr/Sr
Dr./Ph.D.	Dr/PhD
L.A./D.C./U.N.	LA/DC/UN
U.S.[1]	USA
H.R./S.Res.	HR/S Res
Rep./Lt./Col./Sgt.[2]	Representative/Lieutenant/ Colonel/Sergeant
No. 5	No 5
$/%/&	Dollar/percent/and
5 a.m.[3]/5 p.m.	5am/5pm
gotta/cause	got to/because

In order to demonstrate the data cleaning procedure, a comparative overview of an original speech transcript and an adjusted one is given in Table 5. The adjustments are highlighted for easier tracking.

Table 5: *Comparative overview of an original and adjusted speech transcript*

Original speech transcript	Adjusted speech transcript
Mr. President, I rise today to recognize David J. Thatcher, a remarkable Montanan and American. On Apr 18, 1942, Thatcher was one of 80 Doolittle Raiders who carried out the first air raid on Japan during World War II. The unit was named for their commander, Lt. Col. Jimmy Doolittle, who planned and led the mission that dealt a devastating psychological blow to the Japanese Empire in	Mr President, I rise today to recognize David J Thatcher, a remarkable Montanan and American. On April 18, 1942, Thatcher was one of 80 Doolittle Raiders who carried out the first air raid on Japan during World War II. The unit was named for their commander, Lieutenant Colonel Jimmy Doolittle, who planned and led the mission that dealt a devastating psychological blow to the

[1] The abbreviation *U.S.* should not be changed into *US* because the software would recognize that as a first person plural pronoun; hence, it has to be changed into *USA* or a full name

[2] A simple removing periods would result in meaningless words; therefore, a full form words need to be used

[3] Time markers *ante meridiem* (AM) and *post meridiem* (PM) should be spelled as one word with a number preceding them. Otherwise, the software would recognize *AM* as the verb *TO BE*

the wake of the Pearl Harbor attacks.	Japanese Empire in the wake of the Pearl Harbor attacks.
I ask my colleagues in the Senate to join me in honoring Mr. Thatcher and his comrades for their heroic deeds, carried out 71 years ago today.	I ask my colleagues in the Senate to join me in honoring Mr Thatcher and his comrades for their heroic deeds, carried out 71 years ago today.
Staff Sgt Thatcher was born on Jul 31, 1921 in Bridger, MT and entered the Army in Dec 1940. He volunteered for the secret mission that later became known as the Doolittle Raid and was assigned as an engineer/gunner to Crew 7 of the ``Ruptured Duck."	Staff Sergeant Thatcher was born on July 31, 1921 in Bridger, MT and entered the Army in December 1940. He volunteered for the secret mission that later became known as the Doolittle Raid and was assigned as an engineer/gunner to Crew 7 of the ``Ruptured Duck."
On Apr 18, 1942, the Doolittle Raiders launched their B-25 bombers off the USS *Hornet* aircraft carrier, 250 miles further out than planned because they had been discovered by a Japanese fishing boat. During their approach to Tokyo, the crew of the ``Ruptured Duck" spotted a formation of enemy planes, but because of their special training and unique flying tactics, the Japanese formation never detected the ``Ruptured Duck." Crew 7 successfully bombed the Nippon Steel Factory in Tokyo.	On April 18, 1942, the Doolittle Raiders launched their B-25 bombers off the USS *Hornet* aircraft carrier, 250 miles further out than planned because they had been discovered by a Japanese fishing boat. During their approach to Tokyo, the crew of the ``Ruptured Duck" spotted a formation of enemy planes, but because of their special training and unique flying tactics, the Japanese formation never detected the ``Ruptured Duck." Crew 7 successfully bombed the Nippon Steel Factory in Tokyo.
Following their airstrikes, all 16 aircraft either ditched at sea or crash landed because they did not have enough fuel to make it to their intended landing sites on the Chinese mainland. The commander of Crew 7, LT T.W. Lawson, attempted to land the ``Ruptured Duck" on a beach, but instead struck the water a quarter mile off the Chinese coastline. The crew was forced to swim to shore.	Following their airstrikes, all 16 aircraft either ditched at sea or crash landed because they did not have enough fuel to make it to their intended landing sites on the Chinese mainland. The commander of Crew 7, Lieutenant TW Lawson, attempted to land the ``Ruptured Duck" on a beach, but instead struck the water a quarter mile off the Chinese coastline. The crew was forced to swim to shore.
Staff Sgt Thatcher, the only member of Crew 7 who was unharmed, cared for the injured until the Chinese arrived to help. Sadly, 11	Staff Sergeant Thatcher, the only member of Crew 7 who was unharmed, cared for the injured until the Chinese arrived to help.

Doolittle Raiders were killed or captured by the Japanese during the raid but, remarkably, 69 of them were eventually rescued.	Sadly, 11 Doolittle Raiders were killed or captured by the Japanese during the raid but, remarkably, 69 of them were eventually rescued.
Staff Sgt Thatcher went on to serve in England and became an engineer/gunner on a B-26 for the invasion of North Africa. He was discharged from the service on Jul 11, 1945.	Staff Sergeant Thatcher went on to serve in England and became an engineer/gunner on a B-26 for the invasion of North Africa. He was discharged from the service on July 11, 1945.
Today, I would like to honor the four courageous Doolittle Raiders who remain with us: Richard E. Cole, Robert L. Hite, Edward J. Saylor and David J. Thatcher. Let us also take a moment to honor the 76 others who have passed.	Today, I would like to honor the four courageous Doolittle Raiders who remain with us: Richard E Cole, Robert L Hite, Edward J Saylor and David J Thatcher. Let us also take a moment to honor the 76 others who have passed.
The success of the Doolittle Raid marked a turning point in the war. It provided a morale boost for the U.S. and it proved to the Japanese people that they were no longer invulnerable.	The success of the Doolittle Raid marked a turning point in the war. It provided a morale boost for the United States and it proved to the Japanese people that they were no longer invulnerable.
The Doolittle Raiders have earned a hallowed place in our American history, and today I commend Mr. Thatcher and his comrades for their courage and sacrifice.	The Doolittle Raiders have earned a hallowed place in our American history, and today I commend Mr Thatcher and his comrades for their courage and sacrifice.

3.5 Tools

The quantitative part of this research consisted of two subparts; firstly, a computational analysis with the LIWC software tool was carried out followed by a statistical analysis of the data with SPSS. The working principles of both software tools will be described in the following sections.

3.5.1 LIWC

Numerous studies (Pennebaker & Chung, 2011; Fratteroli, 2006; Lepore & Smith, 2002; Pennebaker, 1997; Stiles, 1992; Rosenberg & Tucker, 1978; Gottschalk & Glaser, 1969) have provided evidence suggesting that people's emotional and cognitive worlds, i.e. people's physical and mental health correlate with their linguistic expressions. Being a part of an exploratory linguistic study, Pennebaker and Francis designed the first LIWC (pronounced "luke") application in 1993 with the aim of providing an effective tool for studying the influence and correlation of one's emotional and cognitive state with linguistic expression. The main idea behind LIWC was that the words people use in a variety of genres would reflect their feelings and that we could get insights into people's emotional states by counting words (Pennebaker, 2013: 21). The authors believed that happy people would use happiness words; angry people would use anger-related words, etc. In order to develop LIWC, they compiled a series of dictionaries (anger, happiness, sadness, anxiety dictionary, etc.) to capture different psychological processes. In 2001, Pennebaker, Francis and Booth updated the original application by expanding the internal dictionary and developing a more modern design. Finally, in 2007 the most recent evolution was done with the dictionary being significantly expanded and the software options upgraded. The newest LIWC2007 software[4] was used in this analysis.

3.5.1.1 Processing module

Written or transcribed verbal texts stored in individual files in systematic and meaningful way in ASCII, Unicode or standard .doc files can be processed by the LIWC software. Based on a word-count approach, the software compares grapheme patterns in an input text with the patterns incorporated in the internal dictionary. The software accesses each file individually and compares each target word (a word from a text) with dictionary words (words in the LIWC dictionary file). If a target word matches a dictionary word, the appropriate word scales for that word are incremented writing the output to a single file.

[4] For an online tryout and details on purchasing a license, check the official LIWC website http://www.liwc.net/

Pennebaker (2013: 23-24) provided an example of the software's processing module by using the first two sentences of Lewis Carroll's *Alice's Adventures in Wonderland*.

> Alice was beginning to get very tired of sitting by her
> sister on the bank, and of having nothing to do: once or
> twice she had peeped into the book her sister was
> reading, but it had no pictures or conversations in it,
> "and what is the use of a book," thought Alice "without
> pictures or conversation?"
>
> So she was considering in her own mind (as well
> as she could, for the hot day made her feel very sleepy
> and stupid), whether the pleasure of making a daisy chain
> would be worth the trouble of getting up and
> picking the daisies, when suddenly a White Rabbit
> with pink eyes ran close by her.

Firstly, LIWC would count the words in the text, which is, in this case, 113. Secondly, it looks at each word and checks whether it is in any of the internal subdictionaries. To exemplify, LIWC would first check the word *Alice*; however, it would not find it in the dictionary. Therefore, the word *Alice* would only be placed in the *word count* category. The software would then process the word *was* and find it in several dictionaries – the auxiliary verb, verb and past-tense verb dictionary – and count it as one in each of those dictionaries. The word *beginning* would be placed in the time and verb dictionaries, the word *to* into the preposition dictionary and so on. LIWC would calculate the percentage of total words associated with each of the dictionaries. For example, in these two sentences, 9% of all words are articles, 7% are personal pronouns and 3.6% are emotion-related words.

Processing time for a page of a text is a fraction of a second. In our case, given the fact we have 395 participants' document files with a total of 5,504 pages, it took approximately half an hour for the software to process the data.

3.5.1.2 Dictionary

The internal dictionary is the core of the LIWC analysis. It is composed of 4,500 words and word stems. For example, the dictionary contains the stem *happ** which allows for any word containing these four graphemes (e.g. happier, happiest, happiness, happily, etc.) to be counted as a positive emotion word. The asterisk denotes the acceptance of all graphemes, hyphens or numbers following its appearance. Each word or word stem may be categorized into several LIWC categories simultaneously. For example, the word *grieved* is part of five LIWC categories: sadness, negative emotion, overall affective process, verb and past tense verb. To put it differently, if the verb *grieved* is found in a target text, it will be placed in each of these five categories. Most of the LIWC2007 categories are arranged hierarchically placing a word in several naturally connected categories. To exemplify, the category of pronouns is the sum of personal and impersonal pronouns, all of which are placed in the category of function words (Pennebaker et al., 2007: 4).

The authors (Pennebaker et al., 2007: 5-6) classified 4,500 words and word stems into 80 categories organizing them in four major groups. The first group includes numerous *linguistic processes*, e.g. *pronouns*, *articles*, *verbs*, *adverbs*, etc. and other categories manifesting the way something is said (*negations*, *quantifiers*, *swear words*, etc.). In addition, it also provides *general text descriptors*, e.g. *word count*, *words per sentence*, *words longer than six letters* and *the percentage of an input text covered with the dictionary*. The second group is composed of 32 hierarchically organized *psychological categories* with several superordinate categories – *social*, a*ffective*, *cognitive*, p*erceptual* and *biological processes* – all of which include several subordinate categories. The third group *current concerns* contains the most frequent topics (*work*, *achievement*, *leisure*, *home*, *money*, *religion* and *death*) and their respective lexical representations in a wide variety of texts. Finally, the fourth group involves *spoken categories* (*assents*, *nonfluencies* and *fillers*) useful for the oral production analysis. Additionally, the *punctuation* category provides the frequency of different punctuation signs.

The selection of words included in the LIWC2007 dictionary underwent four major steps during several years reported in the LIWC2007 manual (Pennebaker et al., 2007: 7-8).

Step 1) Word collection

In order to design the LIWC categories, sets of words were generated for each category scale. Numerous sources such as standard English dictionaries and thesauruses were used for *linguistic processes* and *current concerns* categories. In addition to dictionaries and thesauruses, *Positive and Negative Affect Schedule* (Watson et al., 1988) was used to draw on common emotion word items. It took almost three years to get LIWC running due to a painstaking process of building dictionaries during which an army of students evaluating each word was employed. Panels of judges of students had to agree for a certain word to be included in a particular dictionary. Upon completion of preliminary word lists, brain-storming sessions among 3-6 judges were organized with the aim of generating word items relevant to appropriate scales and adding them to initial lists.

Step 2) Rating stages

In order to confirm or reject a word's previous categorization, two rating stages were organized. In the first rating stage, several independent judges reviewed the word lists indicating whether a word should or should not be included in a certain category. They were simultaneously instructed to suggest any additional words to be included in a category. Following the reviews and suggestions, all category word lists were updated in accordance to the following set of principles: a) a word remained in a category if the majority of judges agreed to it; b) a word was excluded from a category if the majority of judges agreed to it; c) a word was added to a category if the majority of judges decided that way. However, considering the objective nature of items in the *linguistic processes* category (numbers, prepositions, articles, pronouns, etc.), the rating process was not applied to most subcategories in this category. In the second rating stage, judges were given category level alphabetized word lists (e.g. *cognitive process* category) and were asked to decide which subcategory a word belongs to. Like in the first rating stage, words were included in or excluded from a subcategory based on the majority judges' votes. The final percentages of judges' decision in the second rating stage ranged from 93% for *insight* to 100% for *ingestion, death, religion, friends, relatives* and *humans* subcategories (Pennebaker et al., 2007: 5).

Step 3) Psychometric evaluation

The initial LIWC rating took place in 1992 followed by a significant revision in 1997. In addition to text file documents from several dozen studies with over 8 million words analyzed with the LIWC1997 version, a word count program *WordSmith*, usually used in discourse analysis, was used for evaluation purposes. Categories used at very low rates or the ones being rated with poor validity and reliability were excluded. Simultaneously, new categories such as *social processes, personal concerns* and the *relativity dimensions* were added upon passing both judges' rating stages as described in step 2. Finally, low frequency words (0.005%) or the ones not listed in Francis and Kučera's *Frequency Analysis of English Usage: Lexicon and Grammar* (1982) were omitted from the LIWC dictionary.

Step 4) Updates and Expansions

Since 1997, LIWC has experienced substantial changes. Several hundred thousand text file documents made up of several hundred million words of both written and spoken language samples were analyzed searching for common words not included in the previous LIWC version all of which were subjected to two rating stages. New word lists and categories were added. The final version of the categories with a detailed word list can be found on the CD attached to this paper.

3.5.1.3 Application of LIWC

A large body of evidence suggests that a computerized text analysis is useful for studies in a wide variety of fields. In clinical psychology, some verbal production features were proven to be related to negative affectivity, schizophrenia, depression and anhedonia (Watson & Pennebaker, 1989; Cohen et al., 2008; Cohen et al., 2009). Furthermore, LIWC proved its usefulness in social psychology studies in the research of lying and deception (Newman et al., 2002), attitudes (Lee, 2009), interpersonal relationships (Ireland et al., 2011; Kramer et al., 2014)), political views (Graham et al., 2009), etc. Researchers also proved

LIWC effectiveness in educational (Carrol, 2007; Robinson et al., 2013) and personality traits studies (Yarkoni, 2010; Hirsh & Peterson, 2009; Mairesse et al., 2007; Pennebaker & King, 1999). Finally, LIWC has been successfully used in linguistic researches in various discourses and contexts, e.g. comparative linguistic (Li et al., 2014), marital conflicts (Bell et al., 2012), social networks (Danescu-Niculescu-Mizil et al., 2011; Lin & Qiu, 2013; Steinberg, 2012; DeWall et al., 2011), political speeches (Slatcher et al., 2007; Kangas, 2014; Duan et al. 2014), gender differences (Schwartz, 2013; Bell et al., 2006; Newman et al., 2008; Kapidžić & Herring, 2011, etc.) and many other studies. Considering the proved effectiveness of LIWC in numerous studies, we decided to analyze our corpus with it.

3.5.2 SPSS

Statistical Package for Social Sciences is a licensed software package developed by IBM Corporation to run on the most widely used operating systems. The software is used for various types of statistical analyses ranging from descriptive and bivariate statistics to prediction of identifying groups and numerical outcomes. SPSS graphical user interface is designed to be used by both novice and advanced users. It provides pull-down menus with internal help and tutorial options on the one and reprogrammable command syntax input language on the other hand. Given its worldwide usage, detailed step-wise tutorials explaining how to perform each test are available online. However, one needs to have working knowledge and understanding of basic statistical concepts.

Initially released in 1968, SPSS has undergone numerous changes and upgrades. The version used in this research is IBM SPSS Statistics version 20. The steps and instructions of uploading the LIWC output in .csv file type into SPSS are provided in step 12) of chapter 3.2.

3.5.2.1 Preparing a codebook

Before entering information from a questionnaire, survey or an experiment into SPSS, it is necessary to prepare a codebook. A codebook is a complete list of one's data transformable to a format the software can understand. Preparing a codebook involves

deciding on labeling variables and assigning numbers to possible answers. It is essential to prepare a codebook because data files contain only abbreviations and numbers which might become meaningless even to an author after longer time of not using it. A codebook should include a full list of variables used in an analysis, their abbreviations input in SPSS and assigned numerical codes.

Our data set consists of two types of variables. We have categorical variables of *gender*, *house*, *party*, *education level*, *religion* and *ethnicity*. For each of the variables, we entered appropriate labels (e.g. men and women) and assigned them numerical values (e.g. 1 for the men and 2 for the women). Numerical values are assigned randomly, i.e. we could have used 0 for the men and 1 for the women. Based on the demographic characteristics of the participants attached on the CD, we went through a list of 395 politicians and manually entered the numerical values for each of the categorical variables accordingly. The codebook with the variable names, labels and numerical codes is shown in Table 6.

Table 6: *A codebook of the categorical variables*

Variable	**Label**	**Numerical code**
Gender	Men	1
	Women	2
House	House of Representatives	1
	Senate	2
Party	Democrats	1
	Republicans	2
Education level	Secondary education	1
	Undergraduate	2
	Graduate	3
	PhD	4
Ethnicity	Caucasian	1
	African-American	2
	Hispanic	3
	Asian	4

Numerical LIWC variables are ratio variables. As such, they are very informative and do not require coding. Moreover, since they are expressed as a number with two decimal places, coding the LIWC variables would result in numerous codes thus being destroyed and

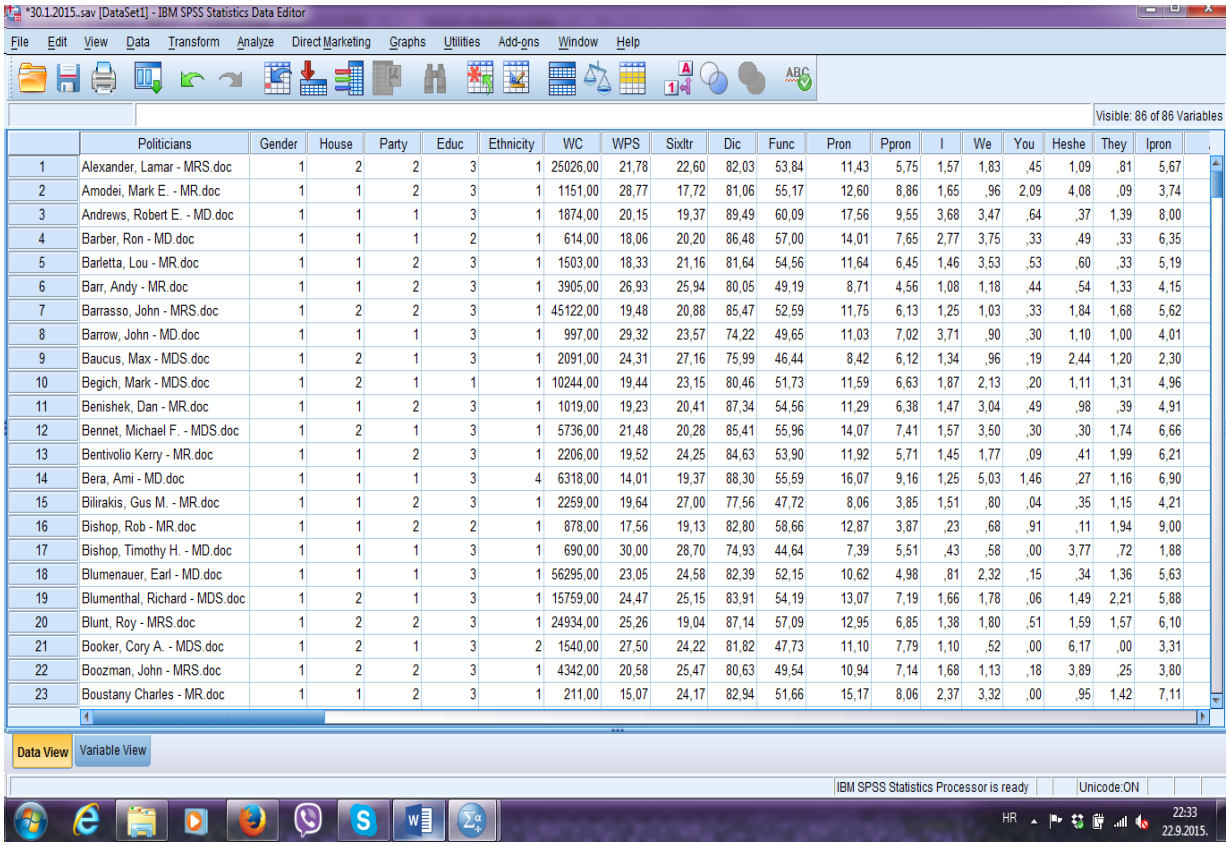
inadequate for processing. Therefore, we did not assign them numerical codes, i.e. we used abbreviations and full variables names as illustrated in Table 7.

Table 7: *A codebook of the LIWC variables*

LIWC variable	Label	LIWC variable	Label
WC	Word count	Cogmech	Cognitive processes
WPS	Words per sentence	Insight	Insight
Sixltr	Words longer than six letters	Cause	Causation
Dic	Dictionary words	Discr	Discrepancy
Func	Function words	Tent	Tentativeness
Pron	Total pronouns	Cert	Certainty
Ppron	Personal pronouns	Inhib	Inhibition
I	First person singular pronoun	Incl	Inclusive
We	First person plural pronoun	Excl	Exclusive
You	Second person pronoun	Percept	Perceptual processes
He/she	Third person singular pronouns	See	See
They	Third person plural pronoun	Hear	Hear
Ipron	Impersonal pronouns	Feel	Feel
Art	Articles	Bio	Biological processes
Verb	Common verbs	Body	Body
Auxverb	Auxiliary verbs	Health	Health
Past	Past tense	Sex	Sexual
Present	Present tense	Ingest	Ingestion
Future	Future tense	Relativ	Relativity
Adv	Adverbs	Motion	Motion
Preps	Prepositions	Space	Space
Conj	Conjunctions	Time	Time
Negate	Negations	Work	Work
Quant	Quantifiers	Achiev	Achievement
Numb	Numbers	Leisure	Leisure
Swear	Swear words	Home	Home
Social	Social processes	Money	Money
Family	Family	Relig	Religion
Friends	Friends	Death	Death
Humans	Humans	Assent	Assent
Affect	Affective processes	Nonfl	Nonfluencies
Posemo	Positive emotions	Fill	Fillers

Negemo	Negative emotions	Period	Periods
Anx	Anxiety	Qmark	Question marks
Anger	Anger	Exmark	Exclamation marks
Sad	Sadness	Quote	Quote

Each variable from an analysis must have a unique name. There are certain rules for naming a variable. A variable name must begin with a letter and not a number. It cannot include periods, blank spaces or other characters and it cannot exceed 64 characters. The first variable in any data set is commonly an identification one. For example, it can be a name of a participant or a number assigned to a case. In our data set, we used the politicians' names as an ID variable. In addition to the input LIWC variables, we added (right mouse click on the top of a column and choose *insert variable* option) the categorical variables of *gender*, *house*, *party*, *education level* and *ethnicity* shown in Figure 12.



	Politicians	Gender	House	Party	Educ	Ethnicity	WC	WPS	Sixltr	Dic	Func	Pron	Ppron	I	We	You	Heshe	They	Ipron
1	Alexander, Lamar - MRS.doc	1	2	2	3	1	25026,00	21,78	22,60	82,03	53,84	11,43	5,75	1,57	1,83	,45	1,09	,81	5,67
2	Amodei, Mark E. - MR.doc	1	1	2	3	1	1151,00	28,77	17,72	81,06	55,17	12,60	8,86	1,65	,96	2,09	4,08	,09	3,74
3	Andrews, Robert E. - MD.doc	1	1	1	3	1	1874,00	20,15	19,37	89,49	60,09	17,56	9,55	3,68	3,47	,64	,37	1,39	8,00
4	Barber, Ron - MD.doc	1	1	1	2	1	614,00	18,06	20,20	86,48	57,00	14,01	7,65	2,77	3,75	,33	,49	,33	6,35
5	Barletta, Lou - MR.doc	1	1	2	3	1	1503,00	18,33	21,16	81,64	54,56	11,64	6,45	1,46	3,53	,53	,60	,33	5,19
6	Barr, Andy - MR.doc	1	1	2	3	1	3905,00	26,93	25,94	80,05	49,19	8,71	4,56	1,08	1,18	,44	,54	1,33	4,15
7	Barrasso, John - MRS.doc	1	2	2	3	1	45122,00	19,48	20,88	85,47	52,59	11,75	6,13	1,25	1,03	,33	1,84	1,68	5,62
8	Barrow, John - MD.doc	1	1	1	3	1	997,00	29,32	23,57	74,22	49,65	11,03	7,02	3,71	,90	,30	1,10	1,00	4,01
9	Baucus, Max - MDS.doc	1	2	1	3	1	2091,00	24,31	27,16	75,99	46,44	8,42	6,12	1,34	,96	,19	2,44	1,20	2,30
10	Begich, Mark - MDS.doc	1	2	1	1	1	10244,00	19,44	23,15	80,46	51,73	11,59	6,63	1,87	2,13	,20	1,11	1,31	4,96
11	Benishek, Dan - MR.doc	1	1	2	3	1	1019,00	19,23	20,41	87,34	54,56	11,29	6,38	1,47	3,04	,49	,98	,39	4,91
12	Bennet, Michael F. - MDS.doc	1	2	1	3	1	5736,00	21,48	20,28	85,41	55,96	14,07	7,41	1,57	3,50	,30	,30	1,74	6,66
13	Bentivolio Kerry - MR.doc	1	1	2	3	1	2206,00	19,52	24,25	84,63	53,90	11,92	5,71	1,45	1,77	,09	,41	1,99	6,21
14	Bera, Ami - MD.doc	1	1	1	3	4	6318,00	14,01	19,37	88,30	55,59	16,07	9,16	1,25	5,03	1,46	,27	1,16	6,90
15	Bilirakis, Gus M. - MR.doc	1	1	2	3	1	2259,00	19,64	27,00	77,56	47,72	8,06	3,85	1,51	,80	,04	,35	1,15	4,21
16	Bishop, Rob - MR.doc	1	1	2	2	1	878,00	17,56	19,13	82,80	58,66	12,87	3,87	,23	,68	,91	,11	1,94	9,00
17	Bishop, Timothy H. - MD.doc	1	1	1	3	1	690,00	30,00	28,70	74,93	44,64	7,39	5,51	,43	,58	,00	3,77	,72	1,88
18	Blumenauer, Earl - MD.doc	1	1	1	3	1	56295,00	23,05	24,58	82,39	52,15	10,62	4,98	,81	2,32	,15	,34	1,36	5,63
19	Blumenthal, Richard - MDS.doc	1	2	1	3	1	15759,00	24,47	25,15	83,91	54,19	13,07	7,19	1,66	1,78	,06	1,49	2,21	5,88
20	Blunt, Roy - MRS.doc	1	2	2	3	1	24934,00	25,26	19,04	87,14	57,09	12,95	6,85	1,38	1,80	,51	1,59	1,57	6,10
21	Booker, Cory A. - MDS.doc	1	2	1	3	2	1540,00	27,50	24,22	81,82	47,73	11,10	7,79	1,10	,52	,00	6,17	,00	3,31
22	Boozman, John - MRS.doc	1	2	2	3	1	4342,00	20,58	25,47	80,63	49,54	10,94	7,14	1,68	1,13	,18	3,89	,25	3,80
23	Boustany Charles - MR.doc	1	1	2	3	1	211,00	15,07	24,17	82,94	51,66	15,17	8,06	2,37	3,32	,00	,95	1,42	7,11

Figure 12: *Screenshot of some of the SPSS variables*

3.6 Tests

Once data is imported and coded in SPSS, it is ready for the analysis. Prior to making a statistical inference, there are several assumptions about the data that need to be fulfilled. Most statistical methods assume an underlying distribution of data. By assuming that data have a particular distribution, we take a serious risk of getting invalid results should the assumption prove to be incorrect. Therefore, we need to check the data distribution carefully. A number of authors (Anderson & Darling, 1954; D'Agostino & Stephens, 1986; Ponocny, 2001; Huber-Carol, 2002; Li & Papadopoulos, 2002; Thode et al., 2002; Steele & Chaseling, 2006, etc.) have considered goodness-of-fit tests whose measures can be used to test whether two samples of data are drawn from an identical distribution. Furthermore, several studies have attempted to compare various tests for distribution normality (Shapiro et al., 1968; Mendes & Pala, 2003; Keskin, 2006; Farrel & Stewart, 2006; Razali & Wah, 2010; Yap & Sim, 2011; Noughabi & Arghami, 2011; Marmolejo-Ramos & González-Burgos, 2012). Power comparisons of the most frequently used normality tests, namely Kolmogorov-Smirnov, Shapiro-Wilk, Vasicek, Anderson-Darling, Kupier, Jarqe-Bera and Cramer von Mises were obtained usually via Monte Carlo simulation. The results were contradictory. As noted by Ahad et al. (2011), Razali & Wah (2010), Farrel & Stewart (2006), Keskin, (2006) and Mendes & Pala (2003), Shapiro-Wilk is the most powerful test for all sample sizes and types of distribution in comparison to Kolmogorov-Smirnov which is the least powerful test. Additionally, Howell (2013) strongly discourages from using Kolmogorov-Smirnov as he finds it powerless and consequently worthless. These findings contrast to that of Shapiro & Wilk (1968) and Shapiro et al. (1968) who did a power test on simulated data with a sample size of maximum 50 cases, thus giving preference to Shapiro-Wilk test for small sample sizes (up to N=50) and Kolmogorov-Smirnov for large sample sizes (N > 50). In 1972, Shapiro and Francia modified Shapiro-Wilk normality test to be used with larger samples. Taking all the relevant studies into consideration, we decided to test the distribution of our data using Kolmogorov-Smirnov and Shapiro-Wilk tests because other normality tests (Vasicek, Anderson-Darling, Kupier, Jarqe-Bera and Cramer von Mises) are not implemented in SPSS.

There are three common ways to test normality assumption: *normality tests* (Kolmogorov-Smirnov and Shapiro-Wilk tests), *numerical methods* (skewness and kurtosis) and *graphical methods* (histogram, stem-and-leaf plot, boxplot, probability-probability plot

and quantile-quantile plot). Firstly, we performed Kolmogorov-Smirnov and Shapiro-Wilk tests aiming to check whether the data of our 70 variables were normally distributed. The test results are presented in Table 8.

Table 8: *Kolmogorov-Smirnov and Shapiro-Wilk tests of distribution normality*

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Word count	,318	395	,000	,472	395	,000
Words per sentence	,059	395	,002	,979	395	,000
Six-letter words	,056	395	,005	,996	395	**,380**[5]
Dictionary words	,034	395	**,200***[6]	,990	395	,007
Function words	,037	395	**,200***	,995	395	**,247**[7]
Pronouns	,064	395	,001	,983	395	,000
Personal pronouns	,057	395	,004	,971	395	,000
I	,114	395	,000	,917	395	,000
We	,099	395	,000	,934	395	,000
You	,222	395	,000	,704	395	,000
He/she	,177	395	,000	,819	395	,000
They	,075	395	,000	,927	395	,000
Impersonal pronouns	,033	395	**,200***	,995	395	**,172**
Articles	,050	395	,018	,976	395	,000
Verbs	,044	395	**,070**	,995	395	**,182**
Auxiliary verbs	,037	395	**,200***	,995	395	**,203**
Past tense	,085	395	,000	,932	395	,000
Present tense	,037	395	**,200***	,995	395	**,178**
Future tense	,090	395	,000	,932	395	,000
Adverbs	,049	395	,026	,961	395	,000
Prepositions	,041	395	**,119**	,991	395	,013
Conjunctions	,058	395	,003	,969	395	,000
Negations	,055	395	,006	,928	395	,000
Quantifiers	,044	395	**,060**	,976	395	,000
Numbers	,091	395	,000	,944	395	,000
Swear words	,499	395	,000	,290	395	,000
Social processes	,058	395	,003	,974	395	,000
Family	,180	395	,000	,705	395	,000
Friends	,127	395	,000	,869	395	,000
Humans	,126	395	,000	,839	395	,000

[5] The p-value of the red-colored variable suggests a normal distribution solely by the Shapiro-Wilk test
[6] The p-value of the blue-colored variables suggest a normal distribution solely by the Kolmogorov-Smirnov test
[7] The p-values of the green-colored variables suggest a normal distribution by both Kolmogorov-Smirnov and Shapiro-Wilk tests

Affective processes	,058	395	,003	,982	395	,000
Positive emotions	,066	395	,000	,970	395	,000
Negative emotions	,067	395	,000	,974	395	,000
Anxiety	,145	395	,000	,822	395	,000
Anger	,117	395	,000	,879	395	,000
Sadness	,114	395	,000	,812	395	,000
Cognitive processes	,036	395	**,200***	,986	395	,001
Insight	,064	395	,001	,967	395	,000
Causation	,077	395	,000	,960	395	,000
Discrepancy	,051	395	,016	,984	395	,000
Tentativeness	,064	395	,000	,970	395	,000
Certainty	,110	395	,000	,914	395	,000
Inhibition	,111	395	,000	,855	395	,000
Inclusive	,049	395	,022	,986	395	,001
Exclusive	,034	395	**,200***	,983	395	,000
Perceptual processes	,101	395	,000	,908	395	,000
See	,126	395	,000	,791	395	,000
Hear	,096	395	,000	,927	395	,000
Feel	,185	395	,000	,670	395	,000
Biological processes	,158	395	,000	,760	395	,000
Body	,169	395	,000	,739	395	,000
Health	,187	395	,000	,694	395	,000
Sexual	,337	395	,000	,419	395	,000
Ingestion	,317	395	,000	,467	395	,000
Relativity	,075	395	,000	,947	395	,000
Motion	,097	395	,000	,942	395	,000
Space	,060	395	,002	,984	395	,000
Time	,083	395	,000	,933	395	,000
Work	,082	395	,000	,952	395	,000
Achievement	,066	395	,000	,952	395	,000
Leisure	,146	395	,000	,819	395	,000
Home	,097	395	,000	,908	395	,000
Money	,090	395	,000	,899	395	,000
Religion	,300	395	,000	,459	395	,000
Death	,214	395	,000	,771	395	,000
Assent	,290	395	,000	,592	395	,000
Nonfluencies	,186	395	,000	,820	395	,000
Fillers	,190	395	,000	,792	395	,000
Question mark	,233	395	,000	,734	395	,000
Exclamation mark	,481	395	,000	,104	395	,000
Quote	,317	395	,000	,535	395	,000

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Generally, the null hypothesis (H_0) suggests a normal distribution of data. In order to test the null hypothesis, we refer to Sig. (p-value). If the p-value is less than 0.05, data significantly differ from a normal distribution, hence, we reject the null hypothesis and accept the alternative hypothesis (H_1). As can be seen from Table 8, the p-values suggested that our data significantly differed from the normal distribution except for the colored variables.

Since the normality tests are not the only method of normality distribution testing, we decided to check skewness and kurtosis and also do a visual inspection. We were particularly interested in the variables the two tests, Kolmogorov-Smirnov and Shapiro-Wilk, reported differently. The histograms, stem-and-leaf plots, boxplots, p-p plots and q-q plots showed that our data were normally distributed in the variables either Kolmogorov-Smirnov or Shapiro-Wilk suggested as such. Precisely, the data from *six-letter words, dictionary word, function words, impersonal pronouns, verbs, auxiliary verbs, present tense, prepositions, quantifiers, cognitive processes* and *exclusive* variables were normally distributed. The data from other variables were not normally distributed. To illustrate, Figures 13 and 14 provide a comparative overview of several variables both normally and non-normally distributed by using a histogram and p-p plot.

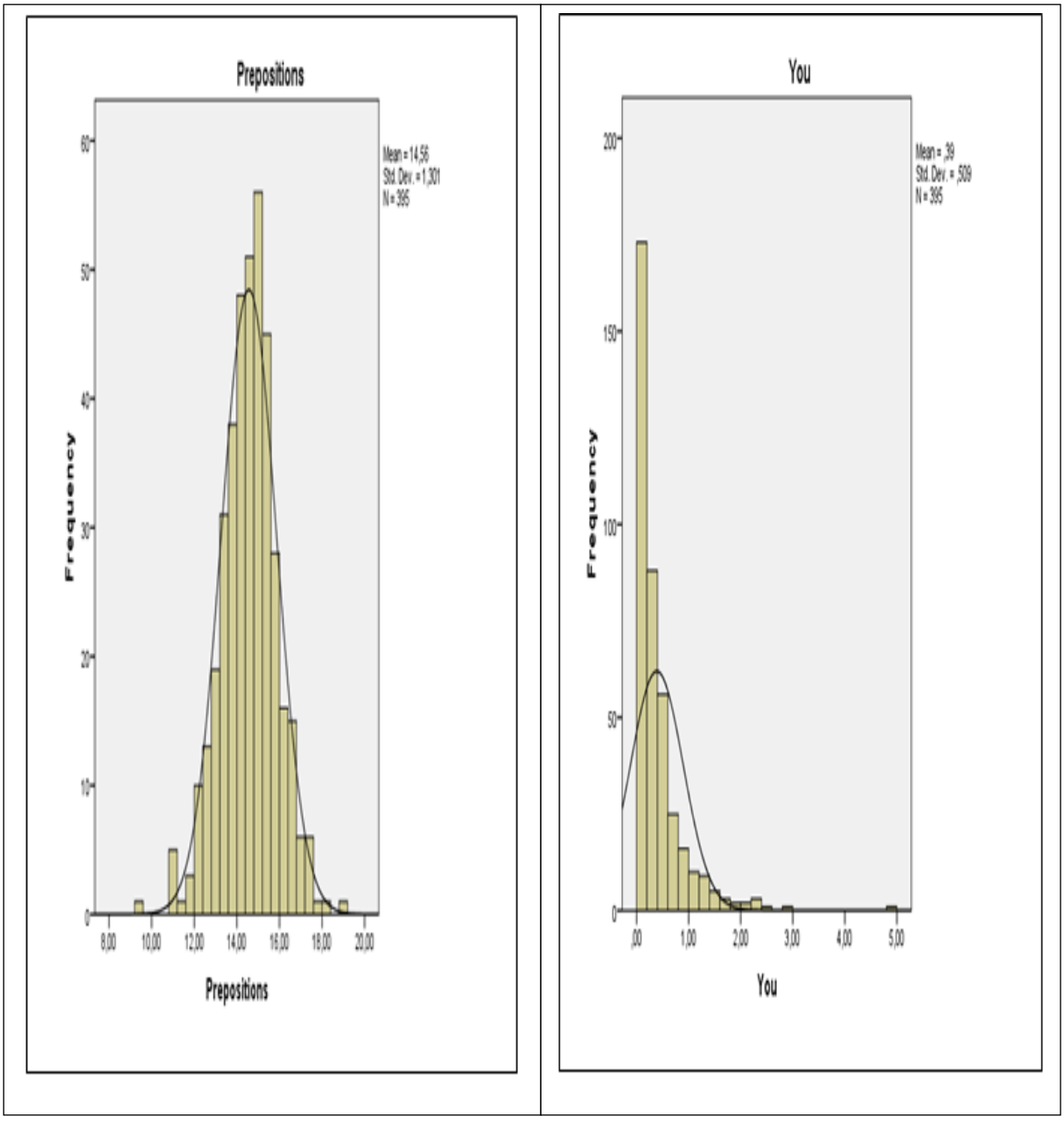


Figure 13: *A comparative overview of normally and non-normally distributed variables using a histogram*

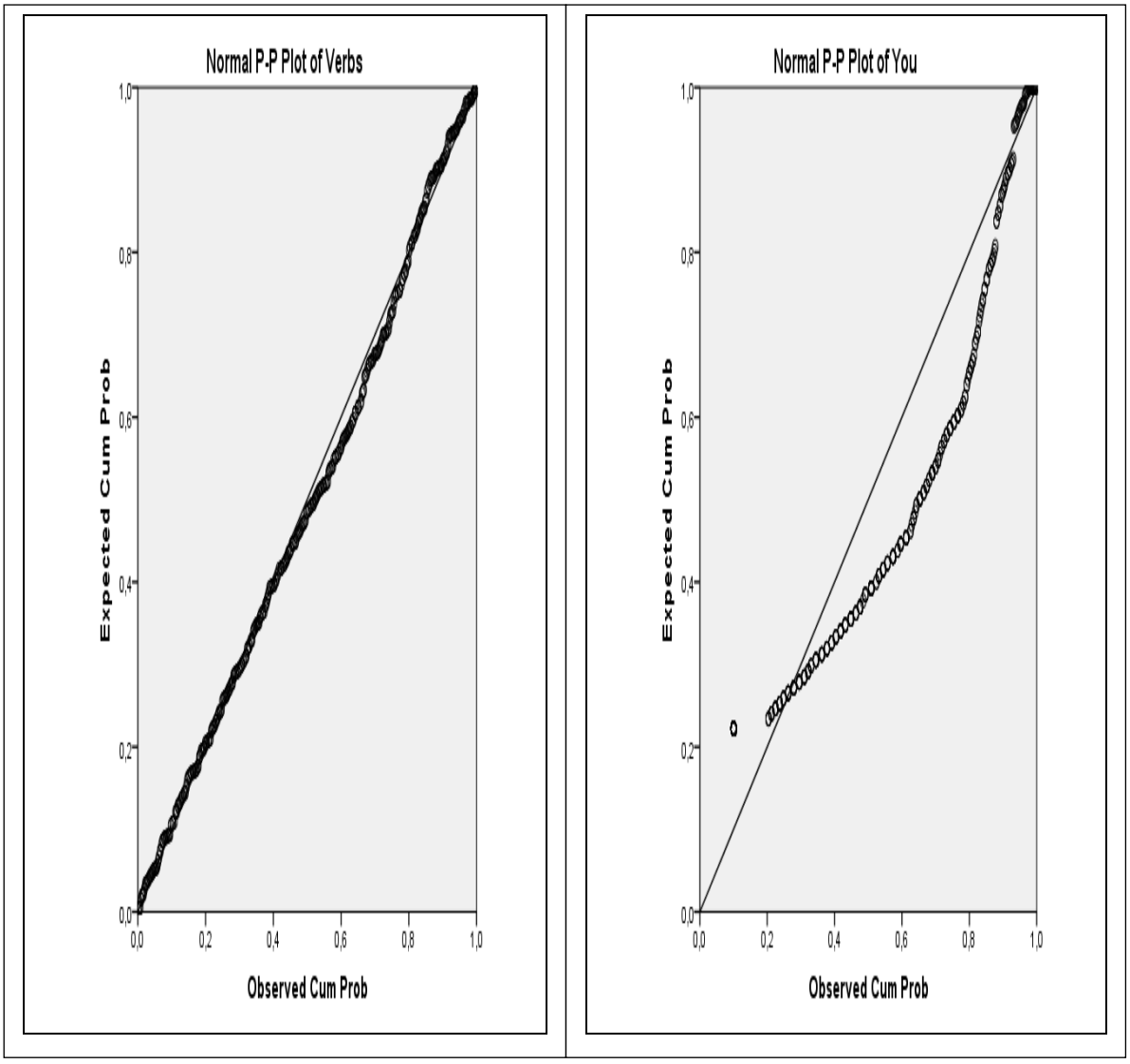


Figure 14: *A comparative overview of normally and non-normally distributed variables using a p-p plot*

Assessing a distribution normality is critical for further tests selection because parametric tests assume a normal distribution of data, i.e. their validity depends on it. Our normality distribution testing results indicated that we should perform parametric tests with normally distributed data. Due to many variables having highly skewed and kurtotic distributions further supported by the graphical methods and normality tests, we followed the previous recommendations (Delucchi & Bostrom, 2004) to use non-parametric tests with such variables.

3.7 Research questions

Since the issue of gender differences in language use has been extensively studied, we decided to compare the previously proven hypotheses and set theories to our research findings.

This dissertation aims to address the following research questions:

1. Is there a statistically significant difference in the usage of 70 linguistic categories between the male and female speakers on the corpus of the political speeches made in the 113th American Congress?
2. Which linguistic categories are predominantly used by the male politicians?
3. Which linguistic categories are predominantly used by the female politicians?
4. How can the determined differences be interpreted?
5. Are the results of the study different from similar studies on gender differences in language use? Which respects do they differ in?
6. Based on the results of the study, which conclusions can we make about gender differences in language in political discourse?

4. Research findings and discussion

Our research consists of two parts, namely a quantitative and a qualitative part. However, in the *Research findings and discussion* section we will not present each of the two parts separately. We will elaborate on each of our 70 variables in a subsection by reporting quantitative research results and providing the qualitative interpretation of the results. We believe that this way of structuring the *Research findings and discussion* section is more effective and easier for a reader tracking-wise.

The variables will be grouped in 5 categories (linguistic category, psychological category, personal concerns category, spoken category and a punctuation category) which will further be subgrouped in their respective subcategories.

4.1 Linguistic category

The linguistic category is composed of the following subgroups: word count, dictionary words, words longer than six letters, function words, pronouns (personal pronouns, first person singular, first person plural, second person, third person singular, third person plural and impersonal pronouns), articles, verbs, auxiliary verbs, past tense, present tense, future tense, adverbs, prepositions, conjunctions, negations, quantifiers, numbers and swear words. Let us now elaborate on each of the subcategories in the following subsections.

4.1.1 Word count

A great deal of previous research into gender differences in language has focused on the issue of verbosity. Marjorie Swacker (1976) investigated gender differences in asking questions and providing answers at academic conferences. Her research found that women contributed only 27.4% with questions asking. Questions asked by women were twice as short as men's questions. The differences in both the structure and length of women's and men's questions were in men introducing the opening question with a statement, asking more than one question and responding to a speaker's answer with additional questions or comments. Additionally, when invited to ask questions, almost exclusively men were first to ask questions; they asked more questions and their questions were longer. The findings led Swacker to conclude that women were less comfortable than men in speaking before a large group of people in a public meeting. Two years later, Westbrook Eakins and Eakins (1978) tape recorded seven university faculty meetings. Their findings supported Swacker's – with only one exception, men spoke more frequently and their speeches were longer than women's. The authors were also interested in turn-takings. Their study recorded that women's turns lasted from 3 to 10 seconds while men's lasted from 10.66 to 17.07 seconds. According

to the research findings, Westbrook Eakins and Eakins concluded that women were reluctant to speak in a public event attended by a larger group of people.

In her book, Dale Spender (1980) suggested that people intuitively believed that women should be seen and not heard. From her point of view, when talking equally, women were perceived as the ones who talked more. She believed that happened because of women speaking in various situations (home, social situations, on a phone), which men could not understand. Similar findings were reported by Sadker and Sadker in 1985. Teachers were shown a videotape of a classroom discussion and asked to conclude who spoke more. They believed that girls talked more when in fact boys talked three times more than girls.

In a ground-breaking paper *Who's Got the Floor?* Edelsky (1981) distinguished between two types of floors: singly developed where one speaker speaks at a time and collaboratively developed which is open to all participants simultaneously. To paraphrase, a collaboratively developed floor, also known as a polyphonic floor (Chafe, 1995), included overlapping speech and co-construction of utterances. Edelsky's research showed that men talked more and took longer turns in a singly developed floor. In comparison, turn length and frequency differences were naturalized and women were more actively engaged in speech in a collaboratively developed floor. These research findings were supported by numerous studies (Falk, 1980; Blitz, 1988; Chafe, 1995; Coates, 1996; Coates, 1997; Coates & Jordan, 1997) conducted in English-speaking communities in Australia, Britain and North America.

Other women's personal experiences of their husbands being talkative at work and life of parties and simultaneously being mute at home inspired the linguist Deborah Tannen to introduce the concepts of report talk and rapport talk in 1990. She believed that men and women used language for different purposes. Women see language as a way of establishing connections and negotiating relationships – a rapport. Men, on the other hand, use language to maintain independence and negotiate status in a hierarchical order – a report. These differences result from different styles of upbringing, expectations and talking to boys and girls. However, in spite of being raised differently, men do not intentionally prevent women from speaking in public settings. Rather, they see women as equals and implicitly invite them, as well as other men, to fight for the floor.

Furthermore, studying gender differences in the workplace, Kendall and Tannen (1997) found that men talked more often, their speeches were longer and they interrupted a person speaking, while women were more interrupted even by a person subordinate to them. Similar findings of men talking more in formal and women in informal settings were reported by James and Drakich (1993) with women's speech often being trivialized and labeled as gossiping (Weatherall, 2002). Starting from the hypothesis that women speak less than men in formal settings, Power and Berardone (1998) carried out an analysis of first speeches in the Australian parliament. Their study showed that there was no statistically significant difference in the amount of men's and women's speech; however, women spoke about a wider range of topics significantly more ($p<.01$) than men.

According to the previous research findings, we expected that the men's and women's speech in the US Congress would differ in terms of word count. Word count and words per sentence categories, unlike others, are not expressed in the percentage form, i.e. the software provides the exact number of used words. The total number of word count in our corpus was 2,615,264 – 2,203,595 words spoken by the men and 411,669 by the women. Even though a conclusion can be drawn from these raw numbers, we did the Mann-Whitney test and found that there was a significant difference in the men's ($M = 205.59$) and the women's ($M = 176.20$) word count ($U = 12719$, $Z = - 2.239$, $p = .025$, two-tailed).

Once we detected a statistically significant difference between the men and women as groups, we were interested in more subtle differences such as differences between and within groups based on the party affiliation, chamber and education level. We then conducted Kruskal-Wallis and the post hoc tests. The results showed that the men Democrats ($M = 216.80$) spoke more than the women Democrats ($M = 168.60$) with a statistically significant difference ($p = .017$). Furthermore, the post hoc tests revealed that the women Representatives ($M = 145.23$) spoke less than the women Senators ($M = 303.15$) as well as the men Senators ($M = 318.71$) with $p = .000$ in both cases. The men Representatives ($M = 165.26$) spoke less than the men Senators ($M = 318.71$) and the women Senators ($M = 303.15$) with $p = .000$ in both cases. There was no statistically significant difference ($p = .112$) in the word count between the men and the women regarding the level of education.

Statistical evidence showed that the Senators, regardless of gender, spoke more than the Representatives which may have happened because of the time limitation in the House of Representatives. Providing statistical evidence that the men spoke more in a public setting, our research contributed to a growing body of research reporting that men speak more than women in various types of public settings thus trying to establish themselves in a hierarchical order.

4.1.2 Words per sentence

The previous studies on gender differences in the length of sentences were contradictory. In 1979, Poole interviewed 96 sixteen-year-olds who were divided in equal social class and gender groups. The verbatim transcripts of the undertaken interviews showed that girls used longer sentences than boys. One of the most productive researchers in the field is Anthony Mulac who, together with Lundell, reported the same results in 1986. They assessed oral descriptions of landscape photographs provided by 40 subjects who were sixth grade students, university freshmen and sophomore students, teaching assistants and older residents from the adjacent town. They audiotaped and later on transcribed the descriptions attributing gender codes to each subject. Compared to Poole's study which was equalized in terms of the education level and social class, demographic characteristics (background, education level, and race) of their subjects varied substantially. They also recorded women using longer sentences than men in oral descriptions.

Furthermore, Mulac and his colleagues (1986) took one-minute speech transcripts by 30 university students and asked 11 trained coders to analyze them linguistically for 35 language features chosen as potential discriminators of speakers' gender. The results of the discriminant analysis showed 100% accuracy of gender prediction based on a combination of 20 linguistic features. One of the discriminant features were longer sentences used by female speakers. The same feature was recorded in Kerstin Thelander's 1986 study on parliamentary language in Sweden (cited in Romaine, 1999). Additionally, Thelander invited her participants to describe each other's language styles. Men's speech was described as abstract, authoritative, impersonal and pompous, whereas women's speech was described as soft,

simple, spontaneous, clear and sensitive. Interestingly, both men and women ascribed negative features to men's speech.

In the second study on the effects of writing, Mulac and Lundell (1994) asked 40 communication class university students to write descriptions of landscape photographs. Their results were consistent with the ones from Mulac and Lundell's study (1986) which recorded women using longer sentences than men. Another study by Mulac et al. (2000) on the differences in language use and effects of men and women managers giving criticisms to their subordinates confirmed that men used more words overall, while women used longer sentences.

The following year Mulac et al. provided empirical support for the Maltz and Borker's (1982) hypothesis that gender differences can be explained by gender-as-culture approach. Mulac et al. (2001) located 16 language features that had consistently indicated communicator gender among which was women's usage of lengthy sentences. In addition, statistically significant difference ($p = .002$) on women using longer sentences when writing about the previous summer in e-mails and letters to male and female friends was reported in Colley et al. (2004). Reporting the same findings, Mulac (2006: 236) stated that men and women "grew up in different sociolinguistic cultural groups and have subtly different styles while accomplishing the same communication task". In a more recent study, Mulac et al. (2013) described men's language as reflecting higher on dynamism, while women's reflected higher socio-intellectual status because women used intensive adverbs, hedges, dependent clauses and longer sentences.

Several studies reported contradictory results. Examining eight-minute problem-solving interactions which involved 108 university students (54 women and 54 men), Mulac (1989) discovered that regardless of their partner's gender, men spoke in longer utterances. Interested in written speeches, Mulac et al. (1990) studied fourth-grade students' essays. The analysis showed that boys used longer sentences than girls.

The previous research reported the difference in the mean length of sentences used by women and men in both written texts and oral speeches in a variety of communicational situations. The majority of them indicated that women used longer sentences than men. Our

Mann-Whitney analysis pointed to no statistically significant difference (U = 14307.5, Z = -.640, p = .522, two-tailed) in the length of sentences, i.e. the men (M = 200.17) and the women (M = 191.77) used equally long sentences. Additionally, the post hoc Kruskal-Wallis analysis revealed no significant difference in neither within nor among groups (for gender and party p = .103, gender and chamber p = .065, gender and education level p = .241).

No statistically significant difference was recorded with respect to the length of sentences used by the female and the male politicians; hence, both the female and the male politicians were equally elaborative when giving their speeches. These results are likely to be related to the formality of a political setting. Another possible explanation is in the participants' preparation for the speech written to be spoken rather than giving one in an impromptu way.

4.1.3 Six-letter words

Several studies investigating gender differences paid closer attention to the length of words. Kučera and Francis (1967) compiled a million-word corpus of present-day English language. Their corpus consisted of 500 samples of about 2,000 words per text. The texts, selected from American publications from 1961, were chosen to include a broad range of styles and topics and were grouped into 15 genres. The computational analysis of the corpus showed that women used less frequent and longer words than men. These results were not supported by similar later research.

Biber et al. (1998) employed a factor analysis technique to study text registers. Their findings showed that male authors used long words and nouns, whereas female authors used more pronouns and present tense verbs. The results led to the conclusion that men used more complex language and their style was informational and uninvolved, while the women's style was more involved. Furthermore, using British National Corpus texts, Kopper et al. (2002) carried on an investigation on predicting a writer's gender simply on words usage. They used a list of 30 words reported in the previous research as having extreme usage variations and possessing distinguishing features across gender and age groups. Reporting 80% accuracy in determining a writer's gender, men's writing was characterized with long words. Newman

and his colleagues (2008) analyzed a database of over 14,000 text files from 70 studies and 22 laboratories from the United States (63 studies), New Zealand (4 studies) and United Kingdom (3 studies). The studies were carried on over a 22-year-period (1980-2002) and contained 93% of written texts and 7% of transcribed speeches. Two-thirds of the participants were college students. Their research findings were consistent with the previous ones reporting men to use longer and more complex words than women. A recent Yu's study (2013) on gender differences conducted on a large corpus of Congressional speeches from the 101st to the 110th Congress (1989-2008) confirmed the consistent usage of long words as a masculine style feature. The same was confirmed by Jones (2015) in an article submitted for publication in *Perspectives on Politics*. She studied speech given by Hillary Clinton as one of the arguably most prominent female American politicians. Analyzing 564 interviews and candidate debates (1992-2013), using LIWC Jones hypothesized that Clinton has been using long words and other masculine style features, changing her language over the years into a more masculine one.

In our study, the independent sample t-test was conducted to compare six-letters words usage by the women and men. There was a significant difference ($p = .000$) in the usage of six-letter words by the men ($M = 23.5$, $SD = 3.61$) and the women ($M = 24.9$, $SD = 3.16$); $t(199) = -3.81$. These results suggested that the women used six-letter words significantly more than the men. We then conducted the one-way ANOVA test to see if there were significant differences within and among groups. The one-way ANOVA test revealed that there was a significant difference on the six-letter words usage at the $p<.05$ level regarding gender and party [$F(3, 391) = 4.62$, $p = .003$]. The post hoc comparisons using the Tukey HSD test indicated that the mean score for the women Democrats ($M = 24.84$, $SD = 2.98$) was significantly different ($p = .010$) than the men Republicans ($M = 23.33$, $SD = 3.83$). Other groups did not differ significantly.

Additionally, we performed another one-way ANOVA test on gender and chamber groups. The test showed that there was a significant difference in the six-letter words usage at the $p<.05$ level regarding gender and chamber [$F(3, 391) = 7.76$, $p = .000$]. The post hoc Tukey HSD test showed that the mean score for the women Representatives ($M = 24.87$, $SD = 3.27$) was significantly different ($p = .001$) than the men Representatives ($M = 23.11$, $SD = 3.76$). Furthermore, the women Senators ($M = 25.23$, $SD = 2.72$) significantly differed ($p =$

.047) from the men Representatives (M = 23.11, SD = 3.76). There was also a significant difference among a group of men. Precisely, the men Representatives (M = 23.11, SD = 3.76) significantly differed (p = .009) from the men Senators (M = 24.58, SD = 2.93). The women Representatives (M = 24.87, SD = 3.27) and the women Senators (M = 25.23, SD = 2.72) did not significantly differ on the six-letter words usage.

We were especially interested in testing the gender differences and education level. The one-way ANOVA signified a statistical difference at the $p<.05$ level [$F(7, 387) = 3$, $p = .007$]. The post hoc Tukey HSD test showed only one significant difference. The mean score for the women who have a graduate level of education (M = 25.14, SD = 3.15) was significantly different (p = .008) than the men with the same level of education (M = 23.46, SD = 3.55). It shows that the female politicians use more complex words than their male counterparts.

To summarize, these results are not in accord with the previous studies claiming that men use long words significantly more than women. Our results suggested that the women affiliated to the Republican party, who served in the Senate, used six-letter words the most thus employing what is traditionally considered as a masculine style feature. The women using complex and less frequent words may be explained by them being more appreciative of the setting formality. Furthermore, by using more complicated and scholarly vocabulary, the women might have tried to establish themselves as valuable contributors to political society. The women serving in the Senate, whose serving position is more competitive than in the House of Representatives, may have wanted to prove their election integrity.

4.1.4 Dictionary words

As we already explained in Section 3.5.1.2, the heart of the LIWC analysis is the LIWC dictionary composed of 4,500 words and word stems. The dictionary consists of subdictionaries containing words tapping a particular domain or a category. All the words included in all of the LIWC categories (except the punctuation category) are jointly presented in the dictionary words category. LIWC calculates the degree to which a subject's vocabulary matches the internal LIWC dictionary. The nature of this category word collection made us

draw the conclusion about the uselessness of studying the category. We decided to skip a detailed analysis of this category and focus on subgroups which are more informative.

4.1.5 Function words

After getting LIWC to start working, Pennebaker (2003) realized that words can be categorized in either the content or function words category depending on psychometric and psychological properties words have. Content words include nouns, verbs, adjectives and adverbs and their function is to convey the content of communication. Function or style words, on the other hand, are pronouns, prepositions, articles, auxiliary verbs, conjunctions, negations and quantifiers. They connect and organize content words. They account for less than one-tenth of 1% of a person's vocabulary but make up almost 60% of words used (Pennebaker, 2003: 8).

There are relatively few studies on differences in the usage of the whole function words category. Biber et al. (1998) found that women used the function words category more thus concluding that women's language was more inclusive compared to men's informative language. Hypothesizing that the largest differences between men and women's language would be in function words, Newman et al. (2008) proved that in a heterogeneous sample of written and spoken texts women tended to use more function words. As a part of a project, Krenn and Schreitter (2015) included 91 Polish participants (33 female and 58 male) aged between 18 and 52 in an experiment. All participants interacted with an artificial communication partner in a positive, negative and neutral mood. The LIWC analysis revealed that female participants used function words significantly more than male participants in a neutral mood system.

Recent attention has focused on profiling authors based on the type of linguistic features they use. Koppel et al. (2002) were able to empirically discriminate male and female authors simply on the function words usage, which, as a general category, were more used by female authors. Houvardas and Stamatatos (2006) used a subset of the Reuters corpus consisting of the same topic texts by 50 different authors. They proved that n-grams were a successful approach in authors' gender identification. The same year, Schler et al. created

Blog Authorship Corpus which consisted of tens of thousands of blogs incorporating almost 300 million words. They reported significant differences in both the content and function words usage by male and female bloggers with the latter using the function words category more. Similarly, Argamon et al. (2009) successfully profiled authors with function words and individual parts of speech being the most consistently effective features. A recent research by Miculicich Werlen (2015) on profiling authors by studying tweets confirmed the previous findings proving for LIWC to be an effective tool which can record better performance than the average state of the art tools in profiling author's gender based on the usage of function words.

The public opinion surveys on politicians' traits done by Rosenwasser and Dean (1989) and Huddy and Terkildsen (1993) revealed that American voters value aggression, competence and toughness. American voters' appreciation of the masculine traits might be the reason of female politicians adopting these traits and rhetoric. The studies by Johnson and White (1994) and Bystrom and DeRosa (1999) on communication styles used by women confirmed that women employed linguistic strategies to emphasize strength rather than warmth, i.e. women used function words at very low rates. In 2013, Meyner's thesis on the language of Japanese female politicians reported a limited usage of female speech characteristics. To put it differently, Japanese female politicians adopted the same speech types as their male counterparts. Lockhart's (2013) analysis on Sarah Palin and Geraldine Ferraro's presidential campaigns speeches confirmed that women were more assertive and direct using function words at low rates.

However, one of the major drawbacks of the previous research was not studying trends of men's speech. Since women were granted the right to participate in the US political system in 1920, feminine style rhetoric might have played an important role in political speeches. Larner (2009) performed a content analysis of the winning presidential candidates from 1932 to 2008. She was especially interested in Nomination Acceptance Speeches and Inaugural Addresses. Her study proved that feminine rhetoric (function words, hedges, modal verbs, and intensifiers) was used in Inaugural Addresses, while masculine rhetoric (directives, exclusive pronouns, quantity, certainty) was used in Nomination Acceptance Speeches. The usage of different rhetoric was explained by the different nature of presidential speeches genre: Inaugural Addresses unify citizenry and foster speaker-audience collaboration which can be

accomplished by using more feminine speech style compared to Nomination Acceptance Speeches whose goal is to represent a speaker as a leader and an expert achieved by using a masculine speech style.

To test the gender differences in the usage of function words, we conducted the independent sample t-test. The results showed a significant difference ($p = .000$) in the usage of function words by the men ($M = 51.67$, $SD = 4.45$) and the women ($M = 49.92$, $SD = 3.38$); $t (230) = 4.14$, i.e. the men used function words significantly more than the women. Following the t-test, the one-way ANOVA recorded a significant difference at the $p<.05$ level regarding gender and education [$F (7, 387) = 3.15$, $p = .005$]. A significant difference ($p = .006$) was recorded between the women with graduate degrees ($M = 49.78$, $SD = 3.26$) and the men with graduate degrees ($M = 51.84$, $SD = 4.32$). Furthermore, a significant difference was in terms of gender and the party affiliation [$F (3, 391) = 5.05$, $p = .002$]. The men Republicans ($M = 51.93$, $SD = 4.72$) used function words significantly more ($p = .039$) than the women Republicans ($M = 49.37$, $SD = 3.94$) and the women Democrats ($M = 50.06$, $SD = 3.22$) ($p = .007$). Gender and chamber also differed significantly [$F (3, 391) = 7.56$, $p = .000$]. The women Representatives ($M = 50$, $SD = 3.46$) and the women Senators ($M = 49.54$, $SD = 3.12$) differed from the men Representatives ($M = 52.11$, $SD = 4.65$) with $p = .001$ and $p = .043$ respectively, while the men Representatives differed from the men Senators ($M = 50.44$, $SD = 3.6$) with $p = .015$. To put it in simpler terms, the men Republicans serving in the House of Representatives used function words the most.

Function words reveal a speaker's style; how something is said. According to Chung and Pennebaker (2007), function words are more closely linked to people's psychological and social worlds. Moreover, they are by the nature very social (Pennebaker, 2003) tying the personal relationship between a speaker and a listener because both actors of communication are required to have basic social skills and shared knowledge to interpret the meaning. Our findings showed that the female politicians used the social category of function words at lower frequency than the male politicians, which is in agreement with Johnson and White's (1994), Bystrom and DeRosa's (1999), Meyner's (2013) and Lockhart's (2013) studies. Furthermore, it supports Larner's hypothesis (2009) that men use some features of the feminine speech style.

There are several possible explanations for this. The women might have been employing a masculine speech style because the public appreciate and value masculine characteristics due to thousand years of male dominance in the field of politics. The men, on the other hand, might have changed their speech style because of a possible perspective change. They might have realized that their function is to serve the community, which is done by addressing issues people are worried about and relating to people in general.

4.1.6 Pronouns

A serious analysis of the relationship between gender and the usage of pronouns emerged in the late 1950s. Interested in a relationship between, as they put it, sex and intelligence, Gleser and the colleagues (1959) recruited 90 Caucasian employed subjects, aged 20 to 50. Their analysis showed that women used pronouns more frequently than men. Later research of transcripts of photographs oral descriptions (Mullac & Lundell, 1986), corpus linguistics (Biber et al., 1998), machine learning approach in identifying authorship (Koppel et al., 2002; Werlen, 2015), formal writing style (Argamon et al., 2003) and heterogeneous written samples (Newman et al., 2008) confirmed the findings.

In the *Handbook of Transsexuality*, Heath (2006) reported results consistent with the previous ones - female authors tended to involve readers in their discourse by using more pronouns than male authors. Furthermore, in a recent study on gender variation in written dating advertisements, Schultz (2013) analyzed a corpus of more than 18,000 online dating ads with 1.4 million words. Sexual orientation of the writer and the gender of addressee were also taken into consideration. Some of the results did not confirm previous findings; however, significantly greater usage of pronouns was confirmed as a feminine writing style. Criticizing previous studies on written texts produced in uncontrolled conditions, Ishikawa (2015) analyzed argumentative essays written by female and male university students on designated topics. Since the study was conducted under controlled conditions, female students using more pronouns was attributed exclusively to gender.

The contrastive studies on Spanish and English and/or Portuguese language by Otheguy and Zentella (2012) and Carvalho and Child (2011) confirmed that immigrant

Colombian, Ecuadorian and Mexican women produced pronouns at higher rates than men. However, they recorded differences within the group of women. Precisely, 93 Latin-born women used pronouns more than men compared to 23 American-born women who used them equally as men, which might indicate a relationship between a culture one was raised in and a linguistic style he or she uses.

More recent studies reported changes in the pronouns usage. In 2012, Andersson examined personal pronouns in editors' letters. The method of corpus linguistic was applied to study 40 editors' letters; 20 from the male-target magazine *Gentlemen's Quarterly* and 20 from the female-targeted magazine *Harper's Bazaar*. The analysis showed that male editors used pronouns more than female editors, which did not support Litosseliti's (2006) findings, supporting the idea that men showed more involvement with their readers. Moreover, Congresswomen using fewer pronouns than Congressmen was reported as the most significant gender difference by Yu (2013) who attributed it to a formal setting. Jones (2015) was especially interested in Hillary Clinton's pronouns change over 21 years in politics. She chose 5 significant career periods: 1) pre-candidate years (1992-1999); 2) first campaign (2000); 3) Senate years (2001-2007); 4) campaign for the Democratic president nomination (2008) and 5) Secretary of State years (2009-2013). The LIWC analysis showed a decline in the usage of pronouns: 18.63%, 17.40%, 17.36%, 18.16% and 16.32% respectively indicating a linguistic shift and a trend of employing a masculine style over time.

The prior studies that have noted the gender differences in pronouns usage reported women's higher usage of the category. The more recent studies have found a significant change with men using pronouns more, i.e. there is a tendency of decline in women using them. Our Mann-Whitney results provided a statistical confirmation on the different usage. The results ($U = 12661.5$, $Z = -2.297$, $p = .022$, two-tailed) confirmed that the men ($M = 205.79$) used pronouns significantly more than the women ($M = 175.63$). Further post hoc Kruskal-Wallis analysis showed no significant difference regarding the education level ($p = .250$) and party affiliation ($p = .052$). However, the post hoc Kruskal-Wallis test revealed a significant difference ($p = .002$) between the men Representatives ($M = 220.08$) and the men Senators ($M = 165.69$).

According to Pennebaker (2003: 169), pronouns, by their very nature, track the relationship between a speaker and a listener/audience. Most pronouns are very social telling us that a speaker is referring to and is aware of other human beings. Our results are consistent with the ones found by Andersson (2012), Yu (2013) and Jones (2015). The usage of pronouns is naturally related to function words because the LIWC category of function words is mostly composed of pronouns. They further support the idea of the women adjusting their speaking style to the formality of a political setting by using a more masculine speech style and the men changing their style by employing feminine features, thus being more socially oriented.

4.1.7 Personal pronouns

Early examples of research (Gleser et al., 1959; Mulac & Lundell, 1986) into the gender differences of the use of personal pronouns indicated women as higher users of the entire category. Mulac and his colleagues (1988) carried on an experiment on 96 university students (48 females and 48 males) who were grouped into same-sex and mixed-sex groups. 20-minute problem solving interactions were videotaped and transcribed. The discriminant analysis results confirmed personal pronouns usage as an indicator of the women's speech style. Sociolinguistic universal of women using linguistics devices to stress the solidarity between a speaker and a listener proposed by Holmes in 1993 was confirmed by Argamon et al. (2003) who proved that female writers used personal pronouns when referring to a listener/reader, while male writers had a tendency of using generic pronouns, i.e. female writers' language pointed to a greater personalization of the text.

Some researchers decided to analyze other languages for differences in the use of personal pronouns. Built on Bodine's (1975) and Ide's (1991) research on marked gender of all three pronoun persons in Japanese, Coates (2003) found significant differences on personal pronouns used by men and women in Japanese. Precisely, *boku* as a first person pronoun and *kimi* as a second person pronoun were exclusive to the men's use. Compared to *ore*, *omae* and *kisama* (exclusive to men), women had no deprecatory pronouns available to them. Coates, therefore, concluded that the lexical choice of pronouns make Japanese women's speech sounded more polite than men's. The results were confirmed by Tanaka the following year.

As a new mode of communication, weblogs have become popular research studies. Herring and Paolillo (2006) investigated language and gender relationship in weblogs with sub-genres of a diary and filter. Their studies showed that filter entries were characterized by masculine style features, regardless of author's gender, while diary entries were characterized by feminine stylistics, with higher personal pronouns usage being one of the distinguishing characteristics. The same result in a study of the language of male and female call operators was reported by Friginal (2009) who explained the female higher personal pronouns usage with their preference for more active involvement and participation during interaction.

Personal pronouns are crucial for the analysis of political speeches because they give a sense of whom a speaker identifies himself/herself with. The results obtained from the Mann-Whitney test (U = 13536.5, Z = -1.416, p = .157, two-tailed) showed no significant difference in the usage of personal pronouns between the men and women. The results do not support the aforementioned findings in various discourses; however, they do support Yu's (2013) finding on female legislators who used fewer personal pronouns. In addition, they support our hypothesis of the female legislators' tendency to use masculine speech characteristics in the predominately male field.

4.1.7.1 Pronoun *I*

A great deal of previous research into political speeches has focused on the use of a specific pronoun. In 1960, Brown and Gilman's pioneering work demonstrated that the choice of pronouns established a relationship of power and inequality or solidarity and equality between participants of communication. According to Karapetjana (2011), the way politicians speak was part of their personality; therefore, pronominal choices are crucial to study.

Several research studies have investigated the usage of the personal pronoun *I* from the perspective of gender. Brownlow et al.'s (2003) research of linguistic behavior of men and women in unscripted televised interviews found than women used the pronoun *I* more than men, which makes them appear more self-focused. The same result was reported by Mehl and Pennebaker (2003) in a study involving 52 undergraduate students in natural conversations

and social environments, Newman et al. (2008) and Andersson (2012). This was at odds with Bell et al.'s (2006) findings on 54 transcribed counselling texts on a variety of topics (sex, infidelity, children, illness, stepfamily, etc.), Larner (2009) and more recent findings of Andersson (2012), Mulac et al. (2013) and Ahmad and Mehmood (2015) who found that men used the pronoun *I* at a higher rate than women in order to exhibit dominance.

Some researchers (Chung & Pennebaker, 2007; Kacewitz et al., 2013, Pennebaker, 2013) hypothesized that the use of the first-person singular pronoun correlated with a status. They believed that when speakers had the power to control the situation, they would be task-focused and consequently less self-oriented. To paraphrase, the higher the status a person has, the less the pronoun *I* use would be recorded. The idea was further extended by an alternative idea of the first-person pronoun use correlating with self-centeredness (Davis & Brock, 1975; Biesen et al., 2015) – a quality not so desirable in politicians; at least not by their voters.

When studying linguistic practices in the traditionally male field, other variables have to be taken into consideration while interpreting results. Arustamyan's (2014) study on Hillary Clinton's language found the frequent use of the pronoun *I* which was interpreted as Clinton's attempt to separate herself from others and present herself as an independent and accomplished politician. This idea was applied to all female politicians. Hakansson (2012) did a comparative study of eight annual speeches on the State of the Union. Four speeches were given by George Bush and four by Barack Obama during their presidency. The two politicians were chosen because of having completely different opinions on political issues, with the aim of studying their pronominal choices. The qualitative research results on the pronoun *I* illustrated that Bush and Obama uses completely different rhetorical strategies. More precisely, Bush used the first-person singular pronoun to express his strong opinion of an issue showing no care for other people's opinion on the subject. Furthermore, taking credit for things he did not or will not do personally, he created an image of a powerful and decisive politician. In comparison, Obama expressed his personal wishes, feelings and compassion for the nation making his speeches more intimate. To paraphrase, Bush's use of the pronoun *I* was more masculine, while Obama's was more feminine.

We conducted the Mann-Whitney test to check if there were any differences between the men and the women on the pronoun *I* usage. The results (U = 14677, Z = -.268, p = .789, two-tailed) showed no significant difference in the usage.

Further post hoc Kruskal-Wallis test on gender and the education level, party and chamber (p = .913, p = .195 and p = .506 respectively) did not show any significant difference. Since the most frequent hedge phrases contain the pronoun *I* (*I think*, *I believe*, *I mean*), we decided to search for them (Control + F function) and calculate their frequency. As presented in Table 9, the women and men equally used the hedge phrases.

Table 9: *Frequency of hedge phrases use*

Token	**Number of the token occurrences in the women's speeches**	**Percentage of the token with respect to the total number of *I* occurrences in the women's speeches**	**Number of the token occurrences in the men's speeches**	**Percentage of the token with respect to the total number of *I* occurrences in the men's speeches**
Total number of *I* tokens	**4,433**		**23,570**	
I think	132	2.98	837	3.55
I believe	88	2	415	1.76
I mean	10	0.22	46	0.2
Total % of hedge *I* phrases		**5.19**		**5.51**

Furthermore, while we were compiling and cleaning the corpus, we realized that the politicians used structured phrases for addressing and greeting Congress. Table 10 lists the most frequent formulaic phrases.

Table 10: *Formulaic phrases of addressing Congress*

Token	Number of the token occurrences in the women's speeches	Percentage of the token with respect to the total number of *I* occurrences in the women's speeches	Number of the token occurrences in the men's speeches	Percentage of the token with respect to the total number of *I* occurrences in the men's speeches
Total number of *I* tokens	**4,433**		**23,570**	
I thank	88	2	233	2
I come to the floor	20	0.45	56	0.24
I wish to honor	28	0.63	96	0.41
I am proud to recognize	58	1.31	212	0.9
I rise	235	5.30	572	2.42
I recognize	7	0.16	179	0.76
I wish	162	3.65	801	3.4
I commend	19	0.43	125	0.53
I yield	58	1.31	406	1.72
I ask unanimous consent	11	0.25	44	0.19
I urge	81	1.83	273	1.16
Total % of using formal addressing phrases		**17.30**		**12.72**

The results showed that both men and women used formulaic expressions to address Congress. However, 17.30% of women's *I* occurrences were phrases of address compared to men's 12.72%. Therefore, the women used the pronoun *I* more formally than the men.

Furthermore, we were inspired by Hakansson's (2012) research which proved that despite having the same overall frequency, the two politicians used the pronoun *I* to convey different messages and create different political profiles. So, we decided to take a look at the sample of speeches given by the women (1-5) and the men (6-10) and selected the most frequent types of examples for comparison. The sample was done in two steps; once we

detected every *I* pronoun using a Word search function, we investigated every third page the pronoun was found on.

1) ***I*** have visited Guantanamo, which is a secure detention facility where people are treated humanely, kept very securely, but not on USA soil. (Kelly Ayotte, Senate - March 12, 2013; Terrorists trials)

2) In the multiple times that ***I*** have been to Israel and the multiple times that ***I*** have been to Ramallah, to the Palestinian Authority, it is a building bonanza going on in the Palestinian Authority. And if it is their land, more power to them. Let them go ahead and build. (Michelle Bachman, House of Representatives, April 28, 2014; Israeli-Palestinian conflict)

3) As the daughter of a 25-year veteran of the Armed Forces, ***I*** recognize the sacrifices our young men and women have made in Iraq and continue to make in Afghanistan. ***I*** am deeply concerned with the widespread incidences of PTSD and the alarming suicide rates among our returning veterans. (Barbara Lee, House of Representatives - March 19, 2013; 10-year anniversary of Iraq war)

4) In the past year, ***I*** have had the opportunity to not only see firsthand what our joint efforts have produced with the Iron Dome antimissile defense system, but also why this is such an important venture. Last August, ***I*** led a congressional delegation trip to Israel to discuss our bilateral relationship with Prime Minister Netanyahu and other top-ranking Israeli officials. (Ileana Ros-Lehtinen, House of Representatives - November 19, 2014; Increase of violent attacks in Israel)

5) ***I*** still remember getting into college. ***I*** still remember back then - and ***I*** graduated from high school in 1978 - that it was 10,000 dollars a year to go to the college ***I*** went to. ***I*** remember my dad thinking: “I can't afford this.” (Amy Klobuchar, Senate – May 7, 2014; Student loan debt)

As demonstrated in examples (1-4), the women reported their personal experience; however, the experience was strictly job-related. Example (5) was different because it reported personal experience. However, one needs to read the entire speech not to draw a wrong conclusion. The reason why this politician shared her personal experience served a purpose of introducing a new bill.

6) Finally, ***I*** want to say that it has been a great pleasure to work with Michael. ***I*** am a pretty good Republican, he's a pretty good Democrat, but that does not make any difference. (Lamar Alexander, Senate – July 9, 2014; Financial aid simplification and transparency Act)

7) My wife and ***I*** drive a Ford Fusion Hybrid, 36 miles a gallon, and we can beat that with other cars, but we are pretty happy with our little Ford. Nobody put a gun to my head and said buy it. My wife and ***I*** thought it was the responsible thing to do. Ford made a great product and we bought it. (Richard Durbin, Senate - June 3, 2014; Global warming)

8) Growing up, ***I*** loved and admired my brother Frank, who was deaf. But ***I*** was deeply disturbed by the discrimination and obstacles he faced every day. That is why ***I*** have always been a passionate advocate for full equality for people with disabilities. (Tom Harkin, Senate - April 30, 2014; Shelby County, Iowa)

9) In all candor, Mr. Speaker, ***I*** play a small part in this film, and ***I*** am pleased the filmmakers allowed me to give my thoughts on the problem of hunger in America in ways that we can address it. (James P. McGovern, House of Representatives - February 26, 2013; A place at the table)

10) When ***I*** first came to Washington, ***I*** was absolutely amazed by the number of academicians, researchers, thinkers, and intellectuals that work and reside in our Nation’s Capital. (Mike Kelly, House of Representatives – July 23, 2014; Another example of American exceptionalism)

The men in examples (6) and (7) share their professional experience; though differently than the women. The women stated that they had visited a place or met with someone putting further emphasis on the event or a reason for a visit. In comparison, the men emphasized their role or importance in the event. Furthermore, in examples (8-10), the male politicians shared their personal experience letting their colleagues and the audience to meet them privately.

In addition to the personal pronoun *I*, the LIWC *I* category consists of the possessive pronoun *my*. We searched for the most frequent *my + X* occurrences and calculated their frequency as shown in Table 11.

Table 11: *List of My + X occurrences*

Token	**Number of the token occurrences in the women's speeches**	**Percentage of the token with respect to the total number of *my* occurrences in the women's speeches**	**Number of the token occurrences in the men's speeches**	**Percentage of the token with respect to the total number of *my* occurrences in the men's speeches**
Total My + X	**1,774**		**8,674**	
My colleagues	368	20.74	978	11.27
My family	24	1.35	67	0.77
My son	3	0.17	35	0.40
My daughter	5	0.28	29	0.33
My husband/wife	7	0.39	78	0.90
My father/dad	17	0.95	75	0.86
My mother/mom	9	0.51	32	0.37
My parents	2	0.11	16	0.18
My friend	46	2.59	265	3.06
My dear + X	5	0.28	19	0.22
My fellow + X	9	0.51	49	0.56
Total My + colleagues		20.74		11.27
Total My + family members		3.76		3.82
Total My + friends		2.59		3.06

The total numbers of the occurrences showed that while there were no greater differences in mentioning family members (women 3.76%; men 3.82%) and friends (women 2.59%; men 3.06%), the women mentioned their colleagues (20.74%) more than the men did (11.27%). In addition to being more formal, the women were more supportive of their colleagues building an atmosphere of collegiality and cooperation. However, these results are not in accord with the ones reported by Schwartz et al. (2013) who analyzed 700 million words collected from volunteers' Facebook messages. Using the open-vocabulary technique, they found that men used possessive adjective *my*, usually accompanied by *wife* or *girlfriend*, more than women used *my husband* and *boyfriend*.

Generally, in political speeches, the pronoun *I* can be used by a speaker to convey his/her opinion, show authority, compassion with the audience and to narrate a story (Bramley, 2001). Another function is to create a relationship because using *I* personalizes the speech. The disadvantage is the issue of subjectivity, which makes some speakers avoid using the pronoun *I* (Pennycook, 1993). Based on the quantitative results of our analysis, we believe a more critical, context-based approach should be used. A more critical approach should be applied by contextualizing the pronoun. The sample we analyzed applying a Critical Discourse Analysis approach clearly demonstrated the gender differences in the pronoun usage. Sharing public rather than private experiences, the female politicians used the pronoun *I* to separate themselves from the audience and group/party affiliation and establish themselves as independent and accomplished politicians, which is consistent with Arustamyan's (2014) findings. The male politicians, on the other hand, did not feel the need to establish themselves because they might have already done it or they possibly believed that their right for establishment had been granted with the election. Therefore, by sharing their personal experiences, the male politicians used the pronoun *I* to create relationships and build a rapport.

Furthermore, the results proved that both men and women used formulaic expressions to address Congress, but the women did it more often. The majority of speeches given by the women started with an explanation or motivation for the speech, while the men occasionally skipped that part and started a speech using the *in medias res* technique, hence the difference. Abiding by the pre-established setting rules, the women might have shown a more serious

approach in performing the job. This idea is further supported by the results of *My + X* occurrences where we showed that the women made significantly more references to colleagues than to their family members or friends. It seems possible than the women saw serving in Congress merely as performing a job and if that was the case, following the rules, not exceedingly mentioning family members and friends or sharing private experience sounded like a reasonable choice.

4.1.7.2 Pronoun *we*

Study of the first-person plural pronoun has been of great interest to researchers who deal with the analysis of political speeches. The previous research on the usage of the pronoun *we* in political interviews coined an expression of institutional identity (Goffman, 1974, 1981; Wilson, 1990; Sacks, 1992). The researchers found that politicians, regardless of gender, used the pronoun *we* to identify themselves with the party they represented. The research was further extended by Janet Holmes (1993), who found that women used inclusive pronouns (we, us, our) more than men with the aim to invite addressees into conversation. Her research supported the findings of Harness Goodwin (1980) who believed that feminine language incorporating more inclusive pronouns than masculine language stemmed from women's views of collaboration and leveling the status of all participants in communication thus forging a common identity. The idea was supported by Skarpol Kaml's research (2000) on Ann Richards' rhetoric. To compare, Pennebaker and Lay's (2002) findings on mayor Rudolph Giuliani's language during crises revealed that, in addition to a marker of a group identity, politicians tended to use the pronoun *we* as a sign of emotional distancing.

Contrary to the previous findings reporting women use the pronoun *we* at a higher rate than men, our Mann-Whitney analysis showed no significant difference (U = 13980, Z = -.970, p = .332, two-tailed) in the usage of the pronoun *we*. However, the post hoc Kruskal-Wallis test registered the significant difference in gender and chamber use. Specifically, the men Senators (M = 162.84) used the pronoun *we* significantly less than the men Representatives (M = 206.08, p = .026) or the women Representatives (M = 14.77, p = .025). Other within or among group differences were not recorded. A closer calculation showed that the subject pronoun *we* made 0.98% of both the men's and women's total vocabulary, the

object pronoun *us* made 0.59% of the men's and 0.70% of the women's, while the possessive adjective *our* made 0.17% of the men's and 0.8% of the women's vocabulary. According to the statistical data, we can conclude that the men and the women used the first-person plural pronoun the same rate-wise.

Additionally, we were interested in the ways the politicians used the pronoun, hence we applied the Critical Discourse Analysis method on a sample of texts. The sample was selected using the same steps as with the pronoun *I*; first we detected every *we* pronoun using the Word search function and then analyzed every third page the pronoun was found on. We were able to identify six contexts in which the pronoun *we* was used.

11) It is not only that ***we*** are providing coverage; ***we*** are providing access to care, so ***we*** can reduce low birth weight babies in our community. Look at the numbers of infant survival. Look at the numbers of low birth weight babies. ***We*** are improving those numbers daily because of the Affordable Care Act. (Benjamin Cardin, Senate – May 12, 2014; Affordable Care Act)

12) ***We*** need to fix our broken tax system, and what better time than Tax Day to highlight this need? (Deb Fisher, Senate – April 15, 2013; Tax day)

Examples (11) (man) and (12) (woman) served the politicians to create an institutional identity speaking on behalf of Congress. However, there is a subtle gender difference. While the men usually stressed what Congress had done or is doing currently, the women were more future-oriented, i.e. what Congress has to do. To put it differently, the men constructed an image of an active Congress, while the women focused on limitations and emphasized the need for further actions.

13) That is why ***we*** Republicans are going to maintain our focus where it belongs - on the people ***we*** represent and on the issues that truly matter to them because our constituents understand that ObamaCare is about so much more than a Web site. (Mitch McConnell, Senate – December 16, 2013; Health care reform)

14) ***We*** do not miss our deadlines, and this year, ***we*** did it. I know that the White House did their Sweet 16 bracket before they did their budget, but ***we*** were still pleased to see that they were willing to participate in that process. (Marsh Blackburn, House of Representatives, September 28, 2013; Protecting the financial solvency of the United States)

Another type of the institutional identity is when expressing party affiliation illustrated in examples (13) (man) and (14) (woman). The men occasionally (eight times in total; four by each party) wanted to stress their affiliation by naming the party explicitly. The women, on the other hand, never mentioned their party in the *we + party* construction. Simultaneously with the party unity, the women even more frequently wanted to stress politicians as a unity as exemplified in example (15).

15) This ought not to be a Republican issue or a Democratic issue. It ought not be a woman's issue. It is an issue that should bother all of us when ***we*** cannot stand together and help those who have been victims of domestic violence. If ***we*** can't do that as a minimum, ***we*** really aren't doing our job, ***we*** really aren't doing service to people. (Lisa Murkowski, Senate – February 28, 2013; Violence against women reauthorization Act)

The third type of the institutional identity is presenting a state as in examples (16) and (17).

16) My State of South Dakota is a good example. ***We*** have balanced our budget every year since 1889. ***We*** have zero personal income tax, zero corporate income tax, and ***we*** have a very well-trained, hard-working, educated workforce. (John Thune, Senate – March 27, 2014; Midterm elections)

17) The people of my State have a disagreement. ***We*** are very fearful about climate change. So ***we*** are also worried about the health impact of the tar sands. (Barbara Boxer, Senate – September 18, 2014; Keystone pipeline)

As illustrated in the examples, both the men (16) and women (17) identified themselves with the people they represented. However, when presenting an issue or an activity done in a particular state, using the pronoun *we*, the women kept their collective identity more than the men who frequently used passive constructions thus isolating themselves as in example (18).

18) As I conclude, let me just say for the 1.7 million children served nationally by CCDBG and the 80,000 served in my State of North Carolina, safe and quality childcare will now be a priority, ensuring working parents trying to better their lives and those of their children will feel safe using their Federal vouchers. (Richard Burr, Senate – November 7, 2014; CCDBG program)

The fourth type is identifying with various committees they served on, such as in examples (19) and (20).

19) What ***we*** found out, through our committee hearings of the committee I am privileged to chair, the Committee on Health, Education, Labor, and Pensions, is that a lot of employers in this country are not abiding by some of the provisions of the Equal Pay Act. (Tom Harkin, Senate – April 8, 2014; Equal pay day)

20) ***We*** will continue in the Armed Services Committee to make sure the reforms that have been passed are implemented, that commanders are held accountable for a climate of zero tolerance within their units, and that victims of sexual assault are treated with dignity and respect and know they will be supported if they come forward to report. (Kelly Ayotte, Senate – March 10, 2014; Victims protection Act)

Both the male (19) and female (20) politicians reported findings and future intentions on behalf of committees they were a part of. In those kind of cases, they rarely expressed their personal opinion on a subject. In other examples not reported in this thesis, if in a chair positions, both the male and female politicians stressed it thus emphasizing their importance or, as they often humbly put it, their privilege.

21) ***We*** all acknowledge the progress that our great country has made on civil rights and voting rights issues. Over time, ***we*** as a Nation have indeed grown to be more perfect - and more inclusive in some ways - than just a few generations ago. (Richard Durbin, Senate, June 27, 2013; Voting rights Act)

22) Those who have suffered discrimination have paid the greatest price for this lack of legal protection. But ultimately ***we*** all pay a price. If our coworkers cannot be themselves in the workplace, they certainly cannot be their best selves. (Barbara Boxer, Senate – November 6, 2013; Employment non-discrimination Act)

Another type, though relatively rarely used, was the politicians identifying themselves with the entire nation as in (21) and (22). Even though the tone in two examples (positive and approving by a man in (21) and worried by a woman in (22) differed, we cannot claim this applied to male and female politicians in general since we found only a couple of such examples.

23) And why have ***we*** abandoned our goal to stop uranium enrichment? Because the Iranian negotiating team has told us they would never tolerate an end to their long, expensive path to an enrichment industry. (Daniel Coats, Senate - February 27, 2014; Iran)

24) Our agreement should absolutely make sure ***we*** are given access to their military facilities so ***we*** can stop them from their programs where they are working on weaponization of nuclear materials. (Kelly Ayotte, Senate – February 26, 2014; Iran)

Finally, the last type of collective identity stemmed from the *we* versus *they* dichotomy exemplified in (23) (man) and (24) (woman). In establishing the *we and they* dichotomy, the politicians created favorable pictures of the unity they presented and simultaneously attributed negative characteristics to the groups they opposed in some way. The dichotomy was usually established to justify the groups' previous or future actions.

To summarize, the statistical analysis pointed to no significant difference between the men and the women on using the pronoun *we* and its variants, whose usage always invokes a

collective identity. However, critically analyzing our corpus, we managed to identify six types of contexts the pronoun *we* was used by the politicians with subtle gender differences. The most prominent use of *we* was for the politicians to construct institutional identity (Goffman, 1974, 1981; Wilson, 1990; Sacks, 1992). The institutional identity was realized as the entire nation, politicians serving in Congress in general, states they represented, parties they were affiliated to and committees they served on. In creating a collective identity, the men emphasized their previous and current successful actions, while the women were more future-oriented. Furthermore, the men were more party-oriented explicitly mentioning it, while the women identified themselves with the state they represented. According to these examples, we propose an idea that the women identified themselves with people who had chosen them and whom they represented, while the men valued the political programs of their party, hence their identification choice. Such interpretation supports somewhat similar findings provided by Ndambuki and Janks (2010) who found that women constructed their identity through the community of people they belonged to.

We were unable to confirm Cassell et al.'s (2006) and Kacewitz et al.'s (2013) correlational results of the pronoun *we* use and high status. This inconsistency may be due to the settings of the two studies. Cassell et al. examined the junior summit online community composed of 3,062 adolescents from 139 countries. The analysis of their messages showed a positive correlation between the pronoun *we* and future leaders. In five studies, Kacewitz et al. included a small number of participants who worked on decision-making tasks, informal chats in get-to-know-you sessions, nine participants' e-mails and letters written by the soldiers during Saddam Hussein's regime. Their result was that high status people used the pronoun *we* at a high rate. We believe this finding could not be replicated in our study because both the Representatives and Senators, or politicians in general, have electoral power and enjoy a high status in society.

One type of institutional identity subtly differed from the others and that was the *we and them* dichotomy. Using *we* acknowledges the existence of another group according to Hirschman (1973) and Pennycook (1993). The dichotomy was used with the purpose of sharing responsibility and collectivity preceded or followed by a controversial decision-making, which supports Jones and Stilwell Peccei's (2004), Karapetjana's (2011) and Al-Faki's (2014) ideas.

Greenwald and Breckler (1986) and Triandis (1989) distinguished between public, private and collective facets of oneself. Given the public setting of politics, the politicians created both the public facet of themselves, reflected in relationships and interactions with the audience, and the collective facet using group reference and identification. They worked on creating a picture of self as part of a positive collective identity. Our analysis has proved that a mere counting of *we* occurrences is irrelevant and not informative. Politicians' artful navigation through wanted identities by using the pronoun *we* with the aim of achieving a range of effects is to be studied in the future.

4.1.7.3 Pronoun *you*

Several studies have so far dealt with the gender differences in the use of the second-person pronoun *you*. They have mostly directed their attention to analyzing women's language in predominately male fields. Analyzing two incidents of domestic violence in Pittsburgh, McElhinny (1998) found that female police officers employed typically masculine strategies trying to appear less emotional. Kuo (2003) videotaped televised sports in Taiwan and analyzed the gender differences in the female and male sports reporters use. The study showed that male sports reporters, regardless of the speaker's role, used second-person pronoun more than female reporters. It also found that men used the pronoun in a more varied way (to refer to a specific and non-specific athlete and the audience), while women used it only to refer to a non-specific athlete. The results have recorded a change in the men's speech style, which, given the frequent use of the pronoun *you*, was described as more informal and conversational. The higher rate usage of the pronoun *you* was also recorded in Friginal's (2009) call centers study; however, it was interpreted with the men's directness and more specific requests supporting the hypotheses of Harness Goodwin (1980), who believed that the pronoun *you* creates hierarchy and enforces authority, and Mulac et al. (1988), who believed it subordinates the audience.

In comparison, a recent newspaper columns study by Ahmad and Mehmood (2015) recorded that women used the pronoun *you* more than men verifying Ruijuan's (2010) conclusion that the pronoun *you*, together with the pronoun *we*, reduces the distance between

a writer and masses. Shifting the perspective from studying the gender differences in the men and women's speech, Larner (2009) studied presidential speech genres and found that male politicians used masculine rhetoric and exclusive pronouns (*I* and *you*) in Acceptance Speeches, whose main purpose is for a politician to prove his expertise and leadership competences, while the feminine rhetoric and inclusive pronouns (*we* and *us*) were used in Inaugural Addresses aimed to unify people.

Similar to the aforementioned studies, the Mann-Whitney test showed a significant difference (U = 12524, Z = -2.446, p = .014, two-tailed) in the use of the pronoun *you*. The mean ranks of the men (M = 206.26) and women's use (M = 174.28) pointed to the men using the pronoun more than the women. Furthermore, the post hoc Kruskal-Wallis test revealed a significant difference in gender and party use. The men Republicans (M = 204.02) used the pronoun *you* significantly more (p = .017) than the women Democrats (M = 167.59). Additionally, the men Representatives (M = 217.33) used the pronoun more than the women Representatives (M = 176.12, p = .031) and the men Senators (M = 175.20, p = .032). Overall, the men Republicans who served in the House of Representatives (M = 221.54) used the pronoun significantly more (p = .047) that the women Democrats from the House of Representatives (M = 166.35). The statistical data may lead to a conclusion that the male politicians were more direct than their female counterparts.

Furthermore, a simple *Word* search function registered 4,380 occurrences in the men's and 762 in the women's speeches. Once we were provided with the specific speeches and contexts of the pronoun, we critically examined every third page the pronoun was recorded on. Note that we did not include those occurrences which were parts of quotes. We were interested in the purposes of the pronoun use both between and within the groups of the male and the female politicians and selected only a few examples for illustration.

The most recorded type of the pronoun use was a generic one presented in (25) and (26).

25) Mr. Speaker, the policy was ***you*** get paid for killing and/or scalping Native American Indians. And if ***you*** kill an Indian boy, ***you*** get paid 50 pounds. If ***you*** get a scalp of an

Indian, ***you*** also get paid 40 pounds. (Eni F.H. Faleomavaega, House of Representatives – June 27, 2013; "Redskin" offensive to Native Americans)

26) Economic principles don't care if ***you*** are a family, a business, or a country. If ***you*** borrow more than ***you*** can pay back, ***you*** go bankrupt. (Mo Brooks, House of Representatives – October 23, 2013; Financial responsibility: the battle resumes in 2014)

Both the female (25) and male (26) politicians used the generic meaning of the pronoun *you* constructing experiences shared by everyone and invoking a sense of what is typical thus exemplifying the theories formulated by Laberge and Sankoff (1980), Sacks (1992) and Malone (1997). In such contexts, the speakers talked impersonally giving the audience a choice of recognizing themselves in the situations. They, however, did not exclude themselves either. According to Sacks (1992), regardless of the type of *you* (generic, singular or plural), a listener is always included in conversation, unlike with the pronoun *we* which might exclude the listener depending on his membership.

The second and third most frequent types of *you* differed between the two groups. The second most frequent type used by the women was the "critical you", while the men's was the "intimate you". First, we will elaborate on the "critical you" exemplified in (27-30).

27) Why do the Republicans keep objecting to this bill? ***You*** cannot, with a straight face, tell me ***you*** truly care about our foreign personnel when ***you*** stand in the way of S. 1386, a bill to provide for enhanced security, a bill that is bipartisan, a bill that came out of the committee on which I serve, Foreign Relations. (Barbara Boxer, Senate – May 20, 2014; Benghazi)

28) I know what ***you*** are doing, and the American people know what ***you*** are doing. ***You*** are using this legislation in your constant effort to discredit President Obama and set the stage for a despicable impeachment proceeding should ***you*** hold the majority in the House and gain the majority in the Senate. (G.K. Butterfield, House of Representatives – July 29, 2014; Institutional litigation is unprecedented)

In the examples, the female (27) and the male (28) politicians criticized the actions of the opposing party related to the image of collective identity and the *we and you* dichotomy with the latter one necessarily being given negative attributes. The next subgroup of the "critical you" is exemplified in (29) (woman) and (30) (man).

29) ***You*** know what, Mr. President? It has everything to do with the budget because of the amount of growth that is taking place in this program. (Marsha Blackburn, House of Representatives – October 2, 2013; Government shutdown)

30) At one time, Mr. President, ***you*** had the White House and ***you*** had the House and ***you*** had the Senate, and yet ***you*** did not even try to get this stuff done. (James Inhofe, Senate – January 22, 2013; Climate change)

When disagreeing with the presiding officer in Congress, neither the female nor the male politicians hesitated in speaking their mind. Hence, they showed they were equal in spite of the presiding officer's higher congressional status and power. Further, the presiding officer's gender played no role in being criticized. Moreover, in none of the examples a criticized presiding officer was a woman.

In terms of the number of examples, the difference between the male and the female politicians was even greater in the usage of the "intimate you" exemplified in (31) (woman) and (32) (man).

31) HR 1797 provides commonsense protections for unborn children who feel pain just as ***you*** and I do. (Virginia Foxx, House of Representatives – June 18, 2014; The pain capable unborn child protection Act)

32) As many of ***you*** know, I own a small business. I understand what it's like to work hard in trying to build a business from the ground up. (Kerry Bentivolio, House of Representatives – February 27, 2013; Protecting small businesses)

While the female politician presupposed the audience's general feelings of pain and used it as shared experience, the male politician went a step further and shared a fact from his

personal/business life. Moreover, he emphasized the existence of a more personal relationship with his colleagues, partially removing the barriers between private and public life.

Additionally, we wanted to support our findings of the women using "critical you" more than the men with the analysis of the *You + modal verb* occurrences listed in Table 12. In the English language, should/ought to, may, might, can, have to, need to and must are modal verbs expressing advice, possibility, necessity and requirement imposed by some source of authority (Lyons, 1977; Palmer, 1986).

Table 12: *List of You + modal verb occurrences*

Token	**Number of the token occurrences in the women's speeches**	**Percentage of the token with respect to the total number of *you* occurrences in the women's speeches**	**Number of the token occurrences in the men's speeches**	**Percentage of the token with respect to the total number of *you* occurrences in the women's speeches**
You should/ought to	5	0.66	36	0.82
You should not	3	0.39	7	0.16
You may/might	10	1.31	62	1.42
You can	52	6.82	296	6.76
You cannot	35	4.59	104	2.37
You have to	15	1.97	88	2
You need to	9	1.18	27	0.62
You must	13	1.71	14	0.32
Total number of *you* occurrences	**762**		**4,380**	

As shown in Table 12, the women and the men did not differ in giving advice or expressing possibility. However, when expressing necessity or requirement of something by using *cannot*, *need to* and *must*, the data showed that the women expressed their stronger opinion more than the men, which supports our previous findings of the women being more critical in an attempt to strengthen their authority.

To sum up, the statistical analysis pointed to a significant difference in the usage of the pronoun *you* with the male politicians using it at a higher rate than their female counterparts, such as in Kuo's (2003) and Yu's (2013) studies. Further Critical Discourse Analysis showed the differences in the purposes of the pronoun usage. While both the male and the female politicians used "generic you" at the same rate to construct universal typical experience, they significantly differed in "critical you" and "intimate you". The research recorded that the women used more "critical you" thus enforcing authority and subordinating both the opposing party and a male presiding officer, which confirmed the hypothesis of Harness Goodwin (1980) and Mulac et al. (1988). The men, on the other hand, used more "intimate you" thereby confirming Chafe's (1982) and Tannen's (1983; 1989) correlation between the pronoun *you*, emotional involvement and connectedness. The follow-up analysis of the modal verbs use supported our CDA findings with the women being more critical than the men. It further supports our idea of the women used a more masculine speech style to prove themselves and assert their authority, while the men used strategies to create an image of the people-oriented politicians.

4.1.7.4 Pronoun *he/she*

Recent studies have investigated gender differences in the usage of the third-person singular pronoun *he* and *she*. Despite studying different discourses, they have reported the same results. In the studies of gender writer's profiling by Koppel et al. (2002) and Argamon et al. (2003), the female writers' use of the pronoun *he/she* was recorded as a very strong indicator of a feminine writing style. Argamon et al. (2003) extended their research to different genres and found that in both fiction and non-fiction works, female writers used the pronoun significantly more than male writers. Herring and Paolillo's (2006) weblogs analysis on the female and male preferential linguistic features confirmed the third-person singular pronoun as a female linguistic feature. Friginal's (2009) call centers study confirmed the previous findings interpreting them with women's preference for more involvement and active participation during an interaction. Yu's (2013) recent study of political speeches supported the findings of the pronoun *he/she*. The only study which did not confirm these results is Ahmad and Mehmood's (2015). They found that the occurrences of the pronoun *he*

was higher in male column writers, whereas the pronoun *she* had a higher occurrence in female column authors.

Unlike the previous studies which reported gender differences in the usage of the pronoun *he/she*, our Mann-Whitney results pointed to no significant difference (U = 13245.5, Z = -1.712, p = .087, two-tailed) between the male and the female politicians in the pronoun use. However, post hoc Kruskal-Wallis recorded that the men Republicans who served in the Senate (M = 262.27) used the pronoun the most, while the women Republicans from the House of Representatives (M = 155.86) used it the least. Interestingly, both the male and the female politicians talked about male persons more. Precisely, the male politicians used the pronoun *he* 9,561 and the pronoun *she* 2,956 times. In comparison, we recorded 1,391 occurrences of the pronoun *he* and 780 of *she* in the female politicians' speeches.

The third-personal singular pronoun *he/she* is a marker to suggest that a speaker is socially engaged and other-oriented. We were unable to replicate the previous studies (Koppel et al., 2002; Argamon et al., 2003; Herring & Paolillo, 2006; Friginal, 2009 and Yu, 2013) findings that women, who are stereotypically more socially aware, refer to other people more than men. However, a more detailed analysis showed that the men Republican Senators talked about other people the most, while the women Republican Representatives the least. Further, when referring to other people, both the men and the women concentrated on men giving them more public attention. From a functional perspective suggested by Halliday (1994), no significant difference in other-orientation implied that both groups presented things in a relational way with the subgroups, violating Biber et al.'s (1998) idea of women's involved and men's informational style. The findings are in line with our hypothesis of the women used more masculine rhetoric and vice versa.

4.1.7.5 Pronoun *they*

The gender differences in the use of the pronoun *they* have not received a lot of attention by researchers. One recent study that dealt with the pronoun was Bell et al.'s (2012). For the purpose of the study, the researchers compared the biological and sociological theories among which they studied Maltz and Borker's (2009) and Gilligan's (1982; 1987)

model. Maltz and Borker's model (2009), based on a biological theory, claims that men and women's speeches have utterly different purposes. Men use language to assert their dominance and position when other speakers have the floor, hence their speech is adversarial and competitive. To compare, women use language to create and preserve relationships and support others, hence their speech is collaborative and affiliative. According to Gilligan (1982; 1987), men's conflict style has a competitive orientation. In order to resolve a conflict, men use logic and rules separating themselves from others. Women's conflict style, on the other hand, has a caring orientation that is focused on establishing and maintaining relationships. They are likely to make exceptions to the rules for the purpose of resolving the conflict. Both models agree that men's style reflects dominance and competition, while women's reflects cooperation, submission and care for others. In the light of both models, the researchers predicted for women to use more social words (among which they studied the pronoun *they*) because women are nurturing and preserve social connections. However, their corpus analysis showed no significant differences in the social words (pronoun they) use. Therefore, their results did not support the social constructionist theories.

Our own corpus analysis confirmed Bell et al.'s (2012) findings. The Mann-Whitney test recorded no significant gender difference in the use of the pronoun *they* ($U = 14283.5$, $Z = -.664$, $p = .507$, two-tailed). The post hoc Kruskal-Wallis test recorded no difference neither between nor among groups of politicians. Therefore, we can conclude that the politicians referred to other people using the pronoun *they* at the same rate. Further, we recorded 11,812 occurrences of the pronoun *they* by the men and 1,994 by the women. Selecting every third page the pronoun was found on, we critically examined the contexts of the usage and found several types which we will elaborate on.

Based on the *we* and *they* dichotomy, the politicians used the pronoun *they* to distinguish themselves, both as individuals and members of a group, from others. Contrary to expectations, others were not necessarily given negative connotations. Moreover, we found that speakers mentioned others in positive, negative and neutral contexts.

33) I think the Founders were right. The Founders in the Constitution outlined the duties of our respective branches of government. ***They*** enumerated them. People will talk about enumerated powers. ***They*** made those powers very few for the Federal

Government. ***They*** emphasized that with the 10th Amendment. (Virginia Foxx, House of Representatives – October 2, 2013; 10th Amendment of the Constitution)

34) I am grateful to Senators Wyden and Murkowski for the bipartisan energy ***they*** have crafted on the committee and for the positive tone ***they*** have set. (Christopher Coons, Senate – September 12, 2013; Energy savings)

Examples (33) (woman) and (34) (man) showed that others were mentioned in a positive context. If it was not for the introductory sentence in example (33), we would not be able to interpret the sentence. By expressing her opinion, the female politician gave the Founders credit for their actions. The Senators from example (34) were also recognized for their actions. One would suggest to check the speaker's party affiliation because if he was in the same party as the Senators, that would be an explanation for the positive comments. The only reason why we did not check it was the fact that the Senators Wyden and Murkowski affiliated with different parties; therefore, while recognizing his party colleague, the speaker also recognized the party opponent.

More frequent were examples of the oppositional relationship between a speaker and other group of people. Here we also found two subtypes, both of which were more frequent in the women's speeches.

35) We were very disappointed, quite frankly, when ***they*** said ***they*** would not move to the conference table with us until we agreed to a tax increase. That is what ***they*** want - an agreement to a tax increase in this kind of economy and with about 8 percent unemployment and with 20 million Americans either un or underemployed? ***They*** want more taxes - more control over people's lives? (Marsha Blackburn, House of Representatives – September 28, 2013; Protecting the financial solvency of the United States)

36) Our friends on the other side of the aisle say ***they*** want to vote on a so-called clean CR. ***They*** insist that we ignore the voices of millions of our constituents who are flooding our offices with calls asking for protection from ObamaCare. (Andy Barr, House of Representatives – October 3, 2013; Government shutdown)

In examples (35) and (36), *they* referred to the opposing party, whose actions were criticized by the speakers. The female politician in (35) started with reported speech and continued with strong accusations of the opposing party's wishes the same as the male politician in (36) who introduced the pronoun *they* by euphemizing opponents and using a spatial metaphor.

37) We need to make sure ***they*** stop enrichment and put a stop on the Arak plutonium reactor and weaponization program. (Kelly Ayotte, Senate – February 26, 2014; Iran)

38) There is nothing more essential than stopping Iran's nuclear program. In order to do that, we need more sanctions. Why? Because every day ***they*** develop ways to get around the existing sanctions program. That is why we need to do a bit more as ***they*** are undoing what we already have in place. (Bras Sherman, House of Representatives – October 4, 2013; Keeping tough Iranian sanctions in place)

Another set of oppositional relationship examples is in (37) and (38), where both the female and the male politician invited their political colleagues and opponents to even stronger unity in order to fight their communal enemy – someone who is not them. Additionally, the *we* and *they* dichotomy was further highlighted by the repetition of *we* which made *them* even more distant. Needless to say, both subtypes were based on a collective identity.

Occasionally, the politicians referred to others neutrally; they did not favorably or negatively evaluate them. Such were the cases when *they* did not belong to the same group as a speaker but differed from the general public.

39) The health care law wasn't about substitute teachers, but ***they*** are the ones feeling the negative side effects and ***they*** are the ones seeing smaller paychecks. (John Barrasso, Senate – April 30, 2014; Health care)

40) Journalists are in prison solely because ***they*** were doing investigative journalism, human rights activists are in prison just because ***they*** felt it was necessary to speak out

about the injustice in their society, and there are people who have been arrested, harassed or tortured because ***they*** disagree with the government and the judicial system of that country is unable to deal with those types of issues. (Benjamin Cardin, Senate – December 8, 2014; Profiling)

In such cases, both the female and the male politicians expressed concerns about the underprivileged or discriminated groups, thus showing their concern for others.

Finally, the last contextual type of the pronoun *they* was a generic one, i.e. the politicians used it to make generalizations about categories of people exemplified in (41) and (42).

41) Millions of Americans' insurance plans have been canceled; ***they*** have lost access to ***their*** doctors and hospitals. (Diane Black, House of Representatives – November 13, 2014; Obamacare architect)

42) Republicans are going to continue to talk good patient-centered reforms, reforms that get patients across the country the care ***they*** need from a doctor ***they*** choose and at a lower cost. (John Barrasso, Senate – September 16, 2014; Health care)

Unlike with the neutral subtype, where the politicians expressed concerns about the specific subgroups, the female (41) and the male (42) politicians used the generic *they* pronoun to show that they cared for the entire nation.

In conclusion, we confirmed Bell et al.'s (2012) findings of no statistically significant gender difference in the usage of the pronoun *they*. The politicians used the pronoun *they* as a resource to identify a group they did not belong to. The generic and neutral contexts were used to show emotions of concern for the selected subgroups or the entire nation. The positive context was used as an agreement or approval of someone else's actions even if those referred to an opponent. As a continuum of the *we* and *they* dichotomy, the negative context emphasized the collective identity in criticizing both the opponent party and foreign nations' political actions the examples of which were recorded at a higher rate in the women's speeches, thus further supporting our "critical you" findings.

4.1.8 Impersonal pronouns

It should be noted that the LIWC category of impersonal pronouns consists of the third-person pronoun *it*, indefinite and relative pronouns. Unlike personal pronouns, very few studies have so far dealt with the use of the impersonal pronouns probably because researchers found personal pronouns more interesting and revealing. Raumolin-Brunberg (1998) studied pronominal changes in the seventeenth century and found significant gender differences caused by the Civil War. In the first observed period (1620-1639), men and women used the third-person possessive adjective *its* at the same rate; however, in the third observed period (1660-1681), women used it even more significantly than men. An even more significant difference was recorded with the relative pronoun *who*, whose acceleration curve had a clear upwards trend after the Civil War. The significant difference from the first period, regarding the use of the compound pronouns in *–body*, when women used the pronouns more, was reduced in the third period with men using them more than women. Finally, men used the compound pronouns *–one* more in the first period, though the difference was reduced in the third period when both groups used them at almost the same rate. Raumolin-Brunberg (1998) concluded that by accepting new alternatives, women were leading the changes in language.

Furthermore, Brownlow and her colleagues (2003) found that in televised interviews men used more impersonal pronouns and passive constructions, which resulted in a depersonalized speech style. Using the validated and recognized LIWC software, Nagarajan and Hearst (2009) examined language use in online dating profiles. They used a popular dating site *Yahoo Personals* and randomly selected 500 female and 500 male profiles between the ages of 18 and 60. They were interested in the *Me and my Partner* section where members described themselves. Their results confirmed that men used impersonal pronouns significantly more than women.

Our independent sample t-test recorded a significant difference ($p = .005$) in the men's ($M = 5.21$, $SD = 1.71$) and the women's ($M = 4.75$, $SD = 1.27$) use of impersonal pronouns, with the men using them more than the women. Further, the one-way ANOVA showed a

significant difference at the p<.05 level regarding gender and chamber [F (3, 391) = 7.7, p = .000]. The post hoc Tukey test showed that the men Representatives (M = 5.42, SD = 1.82) used the category significantly more than the men Senators (M = 4.61, SD = 1.13, p = .000), the women Representatives (M = 4.85, SD = 1.28, p = .027) and the women Senators (M = 4.35, SD = 1.19, p = .020). Our analysis confirmed the trend of men using more impersonal pronouns than women as previously reported by Brownlow et al. (2003) and Nagarajan and Hearst (2009).

In addition to the entire LIWC category, we were interested in examining potential intra-category differences. Precisely, we wanted to see if there were any gender differences in the usage of indefinite pronouns. The pronouns we searched for are provided in Table 13.

Table 13: *List of indefinite pronouns*

Indefinite pronoun	**Number of occurrences in the men's speeches**	**Number of occurrences in the women's speeches**
Anybody	121	15
Anyone	246	34
Anything	302	53
Anywhere	92	17
Nobody	118	4
No one	0	45
Nothing	515	68
Nowhere	54	9
Somebody	123	14
Someone	348	70
Something	946	142
Somewhere	40	3
Everybody	239	29
Everyone	405	79
Everything	431	58
Everywhere	102	24
Total frequency	18%	16%

The number of specific occurrences is accompanied by their total frequency with respect to the total word number in the men's and the women's speeches. As illustrated in the

table, the male politicians used the indefinite pronouns more than the female politicians. The men depersonalized their speech by using hypothetical examples and generalizations making their statements less direct and straight-forward. Furthermore, we found some intra-group differences. With an exception of the pronoun *no one* in the men's speeches, both the male and the female politicians used the indefinite pronouns formed with the suffix *–body* more than the pronouns with the suffix *–one*. According to the online Cambridge dictionary, the pronouns with the suffix *–body* are less formal than the ones with *–one,* hence, both the male and the female politicians used more formal indefinite pronouns probably to adjust to the formality of the setting.

4.1.9 Articles

One might think that a study of articles is useless because they are governed by the rules, which makes their use obligatory. Indeed, articles themselves are not so revealing and thought-provoking. The only reason why one would want to study articles distribution is because they are used with nouns whose usage displays speakers' ability of categorizing things.

Numerous researchers have studied the gender differences in articles distribution in various discourses. Written texts and oral speech studies were equally appealing to researchers. Using a machine-learning algorithm in authors' profiling, Argamon et al. (2003) found that male writers used articles more than female writers. The same result was recorded by Schler et al. (2006) with male bloggers, Newman et al. (2008) in a study on 14,000 text samples, Kapidžić and Herring (2011) on self-presentation in the teen chatrooms and Flekova and Gurevych (2013) on the author profiling in different social media. Furthermore, researchers studying speech had to use different research techniques in recording articles usage. Nevertheless, men using articles at a higher rate than women was reported by Gleser et al. (1959) in the 5-minute verbal samples, Mulac and Lundell (1986) in the photographs description, Mehl and Pennebaker (2003) in the daily conversations of 52 college students, Brownlow et al. (2003) in the television interviews and Yu (2013) in the political speeches. The only study which did not confirm the previous finding was Ludu's (2014) research on the

Twitter users profiling which was unable to demonstrate the association of the articles usage and gender.

Similar to Ludu (2014), we were unable to confirm the gender differences in the article usage. The Mann-Whitney test (U = 13959.5, Z = -.990, p = .322, two-tailed) pointed to the men (M = 201.36) and the women (M = 188.36) using articles equally rate-wise. The post hoc Kruskal-Wallis test did not record neither within nor among subgroups differences. However, we recorded differences within the group of articles. The men used the definite article *the* more than the indefinite *a/an*, while the women used the indefinite article significantly more than the definite one. We can conclude that the politicians used articles equally regardless of their gender; yet, the men made more specific references than the women.

We have already hinted that the usage of articles is related to the usage of nouns. Since articles are used with nouns, especially highly specific and concrete nouns, a conclusion that a person who uses articles is speaking about a particular thing or an object can be made. It has been claimed that the reason why men think and talk about objects in a highly clear-cut way is because of their natural categorization of things (Pennebaker, 2013). Therefore, men naturally use more nouns and consequently more articles. However, articles might be used differently by men depending on their type of thinking. In a comparative study of the speeches by John McCain and Barack Obama during 2008 presidential campaign, Pennebaker (2013) studied the article usage by the two candidates. In the explanation on the American educational system, McCain used articles at very high rates compared to the extremely low usage by Obama. Additionally, McCain broke down the problem into its components, while Obama offered abstract explanations relying on broader and ever-changing principles. This has led Pennebaker (2013) to the conclusion that McCain had a mere categorical and Obama a mere dynamic thinking style. The categorical thinking style includes the categorization in (sub)components and requires the usage of concrete or specific nouns and consequently articles. In comparison, the dynamic thinking style involves evaluating a problem from a developmental perspective and includes more verbs and less nouns and articles.

There are several explanations for our contradictory results. Firstly, the transcripts of speeches given by our politicians may have been prepared in advance, which could have

contributed to the more frequent use of articles. Secondly, the speeches were made in a formal setting where people usually use more formal and distant language (characterized by concrete nouns and more articles), i.e. in such settings, according to Pennebaker (2013), speakers tend to speak like a prototypical man. Finally, highly-formal or low-immediacy thinking is realized with high rates of articles (Pennebaker, 2013), which corresponds to the congressional speech setting. Therefore, we can conclude that the male and the female politicians adjusted their language to the formality of the setting they were delivered in, with both groups equally referring to specific objects or events. This is another example of the women using the masculine linguistic style to adapt to the predominately male field as proved by Lovenduski (2005), Dodson (2006) and Pennebaker (2013).

4.1.10 Verbs and auxiliary verbs

Since all the previous research studied both the verbs and auxiliary verbs categories and found the same results on the gender differences for both of them, we will elaborate on them in the same section.

One of the first research studying these two categories was Gleser et al.'s (1959). When asked to describe a dramatic event in life, women used verbs and auxiliary verbs more than men. Aiming to test Lakoff's hypothesis of women expressing uncertainty, McMillan and colleagues (1977) videotaped women and men in the same-sex and mixed-sex problem solving groups and confirmed that women used verbs and auxiliary verbs more. Similarly, Biber et al. (1998), Mulac et al. (2001) and Mehl and Pennebaker (2003) recorded that women used verbs and auxiliary verbs (especially the modal auxiliary verbs like *could*) at higher rate than men claiming that women's language was thus more tentative than men's. Furthermore, the written texts studies such as 14,000 text files by Newman et al. (2008), authors profiling by Argamon et al. (2009), Iranian EFT students' letters by Hamdi and Dabaghi (2012), social media by Schwartz et al. (2013) and expressing wishes in quotations from English and Dutch sports news articles by Abbas (2014) confirmed that women used verbs and auxiliary verbs more than men. In contrast, only Yu's (2013) study reported that congressmen used more verbs than congresswomen.

We conducted the independent sample t-test which showed a significant difference (p = .013) in the men's (M = 10.67, SD = 2.67) and women's (M = 10, SD = 2.2); t (211) = 2.5 verbs usage. More precisely, the men were found to use verbs significantly more than the women. The test was followed by the one-way ANOVA test for within and among groups' differences. The only significant difference was in terms of gender and chamber [F (3, 391) = 4.36, p = .005]. The post hoc Tukey HSD test showed a significant difference (p = .034) between the men Representatives (M = 10.91, SD = 2.84) and the men Senators (M = 9.90, SD = 1.97) indicating that the men Representatives used verbs more.

The same tests were done on the auxiliary verbs category. The independent sample t-test found a significant difference (p = .013) in the men's (M = 7.12, SD = 1.78) and the women's (M = 6.64, SD = 1.5); t (207) = 2.6 auxiliary verbs usage with the men's higher rate usage. Identically, the one-way ANOVA test found a significant difference in gender and chamber [F (3, 391) = 4.31, p = .005]. The post hoc Tukey HSD test again showed a significant difference (p = .049) between the men Representatives (M = 7.27, SD = 1.88) and the men Senators (M = 6.69, SD = 1.43), with the men Representatives using auxiliary verbs more.

Furthermore, LIWC offers the negated verbs category which was included in our study. No significant difference between the men (M = 202.92) and the women (M = 183.85) using negated verbs was shown by the Mann-Whitney test (U = 13500, Z = -1.453, p = .146, two-tailed). We were not able to find a connection between this category and gender, hence we believe that negated verbs are not gender-specific.

The use of verbs and auxiliary verbs has proven to be related to power and status. In *The Secret Life of Pronouns*, Pennebaker (2013) cited a study carried out by Adam Galinsky - a researcher from the Kellogg School of Management at Northwestern University – who carried on a number of studies on power and language. His results showed that when people have or believe they have power in a group, they are task-oriented, which is reflected in the higher use of articles and nouns. On the other hand, those who do not have power pay attention to others and themselves reflected in the higher usage of pronouns and verbs. Therefore, those higher in status and power are drawn to noun clusters compared to those lower in status who use verb clusters.

Our verbs and auxiliary verbs results confirmed the most recent results in political discourse found by Yu (2013). Further, they are in line with the previous categories results, with the men, especially those serving in the House of Representatives, being more socially and other-oriented and the women being more task-focused.

4.1.11 Tenses

The issue of tenses has not received much attention in sociolinguistic research; yet, it has been studied from a psychological perspective. Berry and the colleagues (1997) videotaped 73 male and 68 female students while they talked about themselves, their classes and activities. The videotapes were then rated by the social perception judges who were unacquainted with the subjects. The category of verb tenses was included to examine whether the subjects who focused more on the present were viewed more favorably than those focused on the past or future. The results showed that the subjects who used more present-tense verbs were seen as warmer people. Additionally, a positive correlation was recorded with perceived competence and tenses use, i.e. those subjects who spoke in the present tense were perceived as more competent individuals. In Gunsch et al.'s (2000) study on political advertising, it was found that the ads written in the present and future tenses were perceived as positive, while those in the past tense invoked negative feelings because the negative ads were seen as focusing on opponents' past actions and positive focused on present and future candidates' actions, which subjects were more interested in. Furthermore, Pasupathi (2007) found that people used the past tense when referring to disclosed and present tense with undisclosed events indicating a psychological distance and a higher degree of resolution for disclosed events. The same year, Mairesse et al. reported a positive correlation with openness to experience and the present tense usage.

The studies dealing with the gender differences in the tenses use in different settings reported contradictory results. Newman et al. (2008) found that men used more present, while women used more past-tense verbs. They believed that women were likely to discuss other people's actions, hence the tense choice. In the experimental setting chat environment, Krenn and Schreitter (2015) found that female participants used present tense verbs significantly

more than male participants. Studying the suicide notes in Spain, Fernandez-Cabana et al. (2015) discovered that women used more verbs in the past and future tenses thus showing more complexity and expressing more interest in transferring information to others. Finally, focusing only on past-tense verbs in the online reviews, Popova (2015) found no gender difference in the usage, acknowledging a small sample size as the main limitation of the study.

Given the contradictory results, we were very open-minded about the tense usage gender distribution. The Mann-Whitney test showed a significant difference in the past tense use ($U = 11658.5$, $Z = -3.307$, $p = .001$, two-tailed). The mean ranks showed that the men ($M = 209.21$) used the past tense more than the women ($M = 165.80$). Furthermore, post hoc Kruskal-Wallis revealed that the women Representatives ($M = 154.43$) used the past tense significantly less than the men Representatives ($M = 205.27$, $p = .004$) and the men Senators ($M = 220.27$, $p = .002$). It was also discovered that the men Republicans ($M = 216.87$) used the past tense significantly more ($p = .008$) than the women Democrats ($M = 166.47$). Overall, the highest use of the past tense was recorded with the men Republican Senators ($M = 225.69$) and the lowest with the women Republicans from the House of Representatives (M 146.36).

Secondly, the independent sample t-test did not find a significant difference in the men's ($M = 6.44$, $SD = 2.32$) and women's ($M = 6.21$, $SD = 1.69$) present tense use [$t (240) = 1.09$, $p = .277$]. Neither within nor among subgroup differences were found by the one-way ANOVA test. Finally, no significant difference was found in the future tense use ($U = 14787$, $Z = -.157$, $p = .875$, two-tailed) by the Mann-Whitney test ($M = 198.53$ for the men; $M = 196.47$ for the women). Similarly, either within or among subgroups differences were not found by the post hoc Kruskal-Wallis test.

Since language tracks our focus of attention, the study of tenses can reveal which dimension subjects are occupied with. Our results showed that the male and the female politicians were equally present and future oriented, which might indicate their openness to new experience and future proactive plans as suggested by Mairesse et al. (2007). Additionally, as demonstrated by Gunsch et al. (2000), present and future orientation are preferable in political discourse. Unlike in Newman et al.'s (2008) and Krenn and Schreitter's

(2015) studies, we found that the men discussed and reported both theirs and other people's actions. Therefore, while being focused on present events and future actions, the men also emphasized their importance and accomplishments thus reminding the audience that in addition to promises and future actions, they have already acted.

4.1.12 Adverbs

Lakoff's (1975) hypothesis of tentative and powerless women's language reflected in exceeding use of intensifying adverbs has received renewed interest recently. The studies successfully mirroring Lakoff's results were Biber et al. (1998), McMillan et al. (1977), Mehl and Pennebaker (2003), Mulac et al. (2000), Heath (2006), Newman et al. (2008) and Jieun and Jae-Woong (2009). Somewhat different Larner's study (2009) on the type of speeches illustrated that adverbs were used by male politicians in Inaugural speeches whose purpose encouraged and favored feminine linguistic characteristics. Furthermore, Vuorinen (2002) studied Queen Elizabeth I's language. Despite her sex, the queen's role in society was perceived as masculine so the study aimed to see whether her language contained masculine characteristics as well. The study had some methodological issues, such as the reason why the queen's linguistic features were compared to personal letters by selected informants and not predefined gender characteristics. Nevertheless, the study showed that the queen's language in general resembled more women's characteristics. However, as far as adverbs were concerned, some adverbs were used as a prototypical woman would, whereas some were not. So her linguistic style was a combination of feminine and masculine linguistic characteristics probably due to her political and leadership role. Another methodologically questionable study was Eliason's (2007), which focused on advertisements in wedding magazines. Even though the study reported that women used intensifying adverbs more than men, the usage could not be attributed solely to gender due to an uneven comparison of advertisements for men.

2012 research by Zaini et al. indicated possible changes in gendered adverbs use. Their study on teen bloggers' language found a very small difference in the number of adverbs used, which led them to conclude that adverbs as a language feature were not gender specific. Furthermore, Xiufang's (2013) research proved that adverbs were used by both men

and women, yet differently. Namely, women tended to used adverbs such as *awfully*, *pretty*, *vastly*, *terribly*, *so* and *quite*, while men used *utterly*, *very* and *really*. Hanafiyeh and Afgari (2014) study involving 120 students confirmed no statistical difference in the usage of adverbs and Yu's (2013) study recorded that congressmen used adverbs more than congresswomen.

Our study confirmed the most recent Yu's results. The results from the Mann-Whitney test (U = 12965, Z = -1.992, p = .046, two-tailed) showed that the men (M = 204.75) used adverbs significantly more than the women (M = 178.61). We were also interested in specific adverbs used by the men and the women so we searched for the adverbs suggested as gender specific by Jieun and Jae-Woong (2009) and Xiufang (2013) listed in Table 14.

Table 14: *List of gender specific adverbs*

Adverb	Number of occurrences in the women's speeches	Number of occurrences in the men's speeches
Utterly	0	19
Most	382	2,899
Terribly	6	21
Pretty	22	228
Very	444	2,075
Sort of	5	109
Quite	31	190
Really	140	577
Much	183	1,326
More	877	5,982
So	943	4,989
Simply	99	679
Seriously	13	107
Totally	6	67

Despite finding only 19 occurrences, we confirmed Xiufang's (2013) findings for the adverb *utterly* being used exclusively by the men. The adverb *most** was used by the male politicians significantly more than by the female politicians. The adverbs *terribly* and *pretty* could not be confirmed as gender (women) specific. Even though these numbers might point

to the significant difference in the usage, the numbers have to be calculated with respect to the total number of words in the men's and women's corpus, which did not point to the significant difference in the usage. However, even though the difference was not significant, *pretty* was used by the men more. Further, the adverb *very* as a significant indicator of men's language, according to Jieun and Jae-Woong (2009) and Xiufang (2013), was in our corpus used equally by the men (9.42%) and the women (9.98%). Also, the adverbs *sort of, quite, really*, *much* and *more* were more used by the male politicians; yet, the difference was not statistically significant. The word *so* can be both an adverb and a conjunction and since LIWC cannot distinguish between those, we checked the data in more detail and found that the frequency data were highly skewed. Namely, the majority of the men's *so* occurrences were conjunctions. To paraphrase, we confirmed Jespersen's (1922) and Lakoff's (1975) findings on the women using *so* as an intensifier more than the men. Finally, we found the adverbs *simply, seriously* and *totally* to be used significantly more by the men.

4.1.13 Prepositions

Several attempts have been made to study gender differences in the usage of prepositions. Shofwan and colleagues (2013) studied gender differences from the perspective of writing errors. They found that female students made fewer prepositional errors, hence they were claimed as more confident users of the category. The following year, Bamman et al. analyzed Twitter messages and even though they found that women used prepositions more than men, the difference was almost unnoticeable.

Several other researchers reported contradictory results. In a blog study by Argamon et al. (2007), prepositions were reported as a strong indicator of the male writers' style confirmed by Newman et al. the following year. Even though Koppel et al. (2003) found that men used prepositions at greater frequency than women, they found anomalies within the category. Namely, while other prepositions were predominately used by men, *for, of* and *with* were used by women significantly more. Moreover, men used the set of all other prepositions with the same frequency as women used the preposition *of*. The previous studies inspired Saeed and colleagues (2015) to examine preposition error writings by 26 ESL students. They

hypothesized that male students would make more errors; however, the research proved for women to make more errors, hence the hypothesis was rejected.

Finally, the most recent Baumann et al.'s (2015) study on blogosphere, including two authors who independently reviewed all resultant materials, pointed to no gender difference in the usage of prepositions.

We performed the independent sample t-test to check distribution differences. The test [t (393) = -.195, p = .845] illustrated that the men (M = 14.56, SD = 1.33) used the category at the same frequency as the women (M = 14.58, SD = 1.22). However, we found the differences in gender and chamber [F (3, 391) = 3.77, p = .011] with the one-way ANOVA; precisely within the group of the men. The post hoc Tukey HSD test showed that the men Senators (M = 14.94, SD = .68) used prepositions significantly more than the men Representatives (M = 14.41, SD = 1.47). Furthermore, we calculated the percentage of the prepositions *for*, *of* and *with* to test Koppel et al.'s (2007) idea of them being gender (women) specific. Our results could not confirm the hypothesis since they were used equally rate-wise (*for* – men 1.08%, women 1.11%; *of* – men 2.9%, women 3.01%; *with* – men 0.57%, women 0.58%).

Generally, the usage of prepositions signals that categorization is done in spatial and hierarchical ways. Our study confirmed Baumann et al.'s (2015) results of no gender differences between the prepositions used. To put it differently, we could not confirm the previous sociolinguistic studies or psychologists' (Pennebaker, 2013) ideas that men use prepositions at a higher rate because they naturally categorize things and assign objects to spatial locations. The participants from our study situated their ideas in time and place equally; therefore, prepositions were rejected as a gender specific category.

4.1.14 Conjunctions

Conjunctions play an essential part in both written and spoken discourse, making sentences more complex and signaling text structure. Therefore, the study of conjunctions can reveal the complexity of a speaker's style. The studies by McMillan et al. (1977), Biber et al.

(1998), Mehl and Pennebaker (2003) and Mulac et al. (2001) reported that women used more conjunctions, especially the conjunction *but*, relating that finding to women being tentative. Ling and Baron's (2007) study on text messaging among American college students confirmed that women used more conjunctions, particularly subordinating conjunctions. Transitional phrases made their sentences longer and text in general easier to follow. In comparison, Mulac et al.'s (1998) study proved that the higher usage of conjunctions was a strong male indicator in spoken discourse.

More recent studies have found different results. Vali and Kianiparsa (2010) research on 24 MA Persian EFL students indicated that conjunctions were used at the same rate in writing. Shofwan et al. (2013) calculated that female and male students made the same number of errors in conjunction use thus claiming that they used them equally. Finally, as a part of her doctoral thesis, Nicolau (2013) found that female students used more conjunctions than male; however, the difference was not statistically significant.

Our findings were in line with the ones done by Mulac et al. (1998). A significant difference ($U = 12676$, $Z = -2.283$, $p = .022$, two-tailed) was found in the conjunctions use by the Mann-Whitney test between the men ($M = 205.74$) and the women ($M = 175.77$) with the men using them more than the women. Further post hoc Kruskal-Wallis illustrated that the women Republicans ($M = 117.93$) used conjunctions significantly less than the women Democrats ($M = 191.68$, $p = .044$), the men Republicans ($M = 207.37$, $p = .003$) and the men Democrats ($M = 203.77$, $p = .007$). Overall, the women Republicans from the House ($M = 113.97$) used conjunctions the least, whereas the men Democrats from the Senate ($M = 242.33$) used them the most. While we could not find significant differences in a specific conjunction usage, we may draw a conclusion that the men's speaking style was more complex and consequently more formal.

4.1.15 Quantifiers and numbers

We are going to elaborate on quantifiers and numbers in the same section because we believe they are somewhat dichotomous concepts. The categories of quantifiers and numbers have not received much attention in gender research. From the corpus linguistic perspective,

the categories were attributed to a men's linguistic style (Argamon et al., 2003; Koppel et al., 2003; Newman et al., 2008; Manjavacas, 2015) whose usage, especially numbers, led to men being perceived as more credible speakers. Yet, they were used by women as well. However, the categories' usage has changed over time. In *Great Speeches of the 20th Century*, universally known speeches by both male and female politicians were compiled. The most striking difference between the 20th century and contemporary speeches is the almost total absence of numbers in the former ones. Hence, the nature of political communication has changed from relying on imagery to using factual statements.

To answer the question of whether contemporary men and women's speech differed in terms of quantifiers, we performed the independent sample t-test which demonstrated that the men (M = 2.44, SD = .72) and the women (M = 2.43, SD = .72) used quantifiers at the same frequency [t (393) = .133, p = .894]. Similar results were recorded by the Mann-Whitney test on numbers. Namely, even though the mean ranks pointed that the men (M = 203.77) used numbers more than the women (M = 181.43), the difference was not statistically significant (U = 13253, Z = -1.702, p = .089, two-tailed). Therefore, we were unable to confirm Argamon et al.'s (2003), Koppel et al.'s, (2003), Newman et al.'s (2008) and Manjavacas' (2015) results.

Furthermore, we also checked for specific occurrences of quantifiers and numbers and calculated their frequency. The four most frequently used quantifiers with almost identical percentage in the men and women's speeches were *all*, *many/much*, *some* and *every*. The major gender difference was in the fifth most used quantifier. Precisely, the men (1,884 occurrences) used statistical data more than the women (268 occurrences). Regarding numbers, there were no in category gender differences. However, when comparing quantifiers and numbers occurrences, we found that both the men and women behaved equally. Namely, in the men's speeches there were 25,173 (1.14%) quantifiers and 3,808 (0.17%) numbers occurrences. Similarly, 4,423 (1.07%) of quantifiers and 813 (0.2%) of numbers occurrences were recorded in the women's speeches. Therefore, we may draw a conclusion that both the male and the female politicians were more comfortable with using vague quantifiers than real numbers.

Using numbers in political speeches is important for several reasons. Firstly, talking in abstract terms is less effective than providing numerical proof for one's arguments. Secondly, a speaker can amplify the emotional response from the audience if using numbers. For example, if a speaker provides a specific number of rape victims, it will have a stronger effect than the quantifier *many*. Using statistical data shows that a speaker did research for the speech, which will be appreciated by the audience and it will simultaneously boost one's credibility. Hence, politicians tend to use statistical data by combining *ethos*, *pathos* and *logos* – the three pillars of public speaking. Our research showed that both the men and women used them equally.

In addition to numbers, the politicians also use quantifiers; in our case, they preferred them more than numerical data. There are several possible explanations for this. Firstly, remembering specific numbers can be problematic and may lead to mistakes. This problem, obviously, can be overcome by preparing written notes. Secondly, unlike quantifiers, numbers are definite and their usage is verifiable so politicians pay special attention when using them. Numerical data are probably used differently before and after elections. If they cannot keep their promises in numerical terms, politicians will be punished on next elections, whereas they cannot be punished when using indefinite quantifiers. We, therefore, believe that both the women and men from our study preferred quantifiers over numbers because, in spite of losing some credibility, the vagueness of their use could do less political harm than lying to voters would.

4.1.16 Swear words

Traditionally in the folklinguistic belief, men use taboo words and swear more than women. One of the first researchers to initialize such belief was Jespersen (1922: 246) who claimed that "women have instinctive shrinking from coarse and gross expressions and a preference for refined and (in certain spheres) veiled and indirect expressions". The belief was confirmed in Gleser et al.'s (1959), Foote and Woodward's (1973), Selnow's (1985), Limbrick's (1991), De Kirk's (1991), Jay's (1992), Mehl and Pennebaker's (2003), Berger's (2003), Newman et al.'s (2008) and Kryeziu's (2015) studies and challenged by Staley (1978), Mulac and Lundell (1986), Risch (1987), Hughes (1992) and Jay and Janschewitz

(2007). Flexner and Wentworth (1975) believed that most American slang vocabulary was both created and used by men. Lakoff's (1975) impressionistic view that men used much stronger expletives than women was confirmed in self-reported studies by Oliver and Rubin (1975) and Bailey and Timm (1976) and refuted in observation studies by Anshen (1973), Gomm (1981), Jay (1986) and Limbrick (1991). Kramer's (1974) study on cartoons in which students were asked to identify captions used by female and male characters proved that swearing was perceived as the men's speech thus confirming cultural stereotypes. Jay (1992) suggested that the gender differences in the offensive language usage existed because of different views on the world – while women operate in the world dealing with social acceptance, security and intimacy, men are concerned with power, sex and physical attraction.

Furthermore, some researchers (Gomm, 1981; Wells, 1989; Jay, 1992; Coates, 2003) investigated vulgar language in the same-sex and mixed-sex groups and found that both men and women swore more in the same-sex groups showing more restraint in mixed-sex groups. Gauthier (2012) confirmed that men swore more in the all-male groups; however, women generally tended to swear equally regardless of the type of a group. In addition, Thelwall's (2008) study on the MySpace pages did not find significant gender difference among British teenagers in comparison to American male adolescents who swore more than their female counterparts, which pointed to inequalities among cultures. Further, Broadbridge's (2003) analysis confirmed the previous findings on men swearing more. However, when asked to evaluate their speech regarding profanity, female speakers believed they swore a lot and in some cases even more than men. It was interesting that the person who believed she swore the most in the group did not swear once. This example pointed to the differences in self-evaluation and possibly profanity categorization. Another perception study by Cavazza and Guidetti (2014) on the politicians' weblogs proved that voters did not change their opinion about a politician (especially male) when he used informal and vulgar language.

Our study partially confirmed Thelwall's (2008) results. The Mann-Whitney results demonstrated no significant gender differences on using vulgar vocabulary ($U = 14864.5$, $Z = -0.149$, $p = .881$) with the mean ranks $M = 197.73$ for the men and $M = 198.77$ for the women. More specifically, the majority of the politicians' speeches did not contain vulgar language, i.e. 30 male and 11 female speeches contained some profanities. The only within group difference, yet not statistically significant, was that the vulgar language was more

recorded in the House of Representatives' speeches (20 men and 7 women) than in the Senate (10 men and 4 women).

Additionally, Bailey and Timm's (1976) and Wierzbicka's (1991) studies revealed an assumption of swearing as an expression of negative emotions so we were interested whether swear words correlate with positive and negative feelings. In order to test that, we performed the Spearman's rank-order correlation test. While there was no correlation between the swear words and positive ($p = .175$) and negative ($p = .978$) emotions, we found a positive correlation between the swear words and anxiety [r_s (393) = .137, $p = .006$, two-tailed]. To put it differently, the more a speaker was anxious, the more he/she swore.

The stereotype of tough-talking men on the one and very polite and never-vulgar women on the other hand was false. Even though swearing has been claimed to be an integral part of masculine traits, research showed that feminine language has been changing and becoming more vulgar (Thewall, 2008). However, when conducting research, one needs to bear in mind the context of the speech because it might be crucial for the register usage. Our swear words findings were rather expected. The formality of setting contributed to the formality of language used by speakers of both gender. Both the men and women respected the institution of Congress by almost never using any profanities, and in those rare cases when they were used, the rate was extremely low (the highest was 0.8% of the total words in the corpus) with no gender differences.

4.2 Psychological processes category

The psychological processes category consists of *social processes* which includes the subcategories dealing with family members, friends and humans in general, *affective processes* containing positive and negative emotions, anxiety, anger and sadness, *cognitive processes* including insight, causation, discrepancy, tentativeness, certainty, inhibition, inclusive and exclusive words, *perceptual processes* dealing with auditory, visual and sensation and *biological processes* composed of the words related to body, health, sex, ingestion, relativity, motion, space and time, all of which will be elaborated on in the following subsections.

4.2.1 Social processes

The category of social processes includes everything related to family members, friends and human beings in general. Traditionally, caring and talking of other people has been ascribed to women because they are, by nature, more people-oriented. Several research projects have confirmed the stereotype. For example, Brownlow et al.'s (2003) study proved that women used language that focused on social processes; yet, women were more self-referent and did not show more emotions than men, which was contradictory. Further, female participants from Newman et al.'s (2008) corpus study made more references to social processes especially mentioning family members. In Stepney's (2014) MA thesis, female students had a significantly greater percentage of social processes words in their essays. To compare, Bell et al.'s (2006) study found no significant gender difference in social processes references.

Since personal pronouns are highly social, the results of the two categories might be related. In order to check that, we performed the Spearman correlation test which showed a positive correlation [r_s (393) = .794, p = .000, two-tailed] between the personal pronouns and social processes use. Taking the results into account, and bearing in mind that the personal pronouns were equally used by the politicians, we hypothesized that the gender differences would not be found in the social processes category usage. The Mann-Whitney test results (U = 14534, Z = -.412, p = .680) and the mean ranks M = 199.40 for the men and M = 193.99 for the women pointed to acceptance of our hypothesis and confirmation of Bell et al.'s (2006) findings. However, the post hoc Kruskal-Wallis test showed within group differences. Namely, the men Senators (M = 165.60) used the social processes references significantly less (p = .014) than the men Representatives (M = 211.44). Even more specifically, the men Democrat Senators (M = 155.99) used the category significantly less (p = .039) than the men Republican Representatives (M = 225.20).

Furthermore, we performed some additional Spearman two-tailed correlation tests which showed that the politicians who used the social processes category tended to use shorter [r_s (393) = .-246, p = .000] and less complex words [r_s (393) = .-457, p = .000], i.e. the

level of formality decreased when speaking about other people. There was a positive correlation between the social processes category and positive emotions [r_s (393) = .150, p = .003] as well as the content categories of home [r_s (393) = .147, p = .003] and death [r_s (393) = .110, p = .029], which might be explained by the politicians commemorating other people. Finally, a negative correlation was recorded between the social processes category and the content categories of work [r_s (393) = .-159, p = .001], achievement [r_s (393) = .-105, p = .037] and money [r_s (393) = .-264, p = .000], which illustrated that the politicians clearly distinguished between the formal work-related matters and social other-orientation topics.

4.2.1.1 Family, friends, humans

The stereotype of women being more other-oriented, reflected in them talking and writing about family, friends and humans in general, has been confirmed by Schler et al. (2006) and Newman et al. (2008). In a recent study on Google+ posts, Cunha et al. (2014) confirmed the stereotype providing evidence that women used the social networks to talk about social and family relations, while men used them to discuss technical topics and their professional achievements. Starting from the same hypothesis, Xu et al. (2014) found no evidence of women using more other people's reference, hence they rejected it. Interested in the language of leaders, Cassell et al. (2005) hypothesized that elected politicians would employ more powerful language in their messages. Contrary to their beliefs, leaders have been proven to use the social processes references of friends, family and other humans.

Since the overall category of social processes was equally used by the male and the female politicians, we believed that the same result would be replicated in the family, friends and humans subcategories. With the subcategory of friends, the Mann-Whitney test (U = 12179, Z = -2.785, p = .005, two-tailed) recorded a significant difference. The mean ranks pointed that the women (M = 225.10) used the family references more than the men (M = 188.57). Additional differences were recorded by the post hoc Kruskall-Wallis test. The men Representatives (M = 183.74) used fewer family references than the women Representatives (M = 223.87) with a significance of p = .040. As far as the party and gender were concerned, the women Democrats (M = 238.92) used the family references significantly more than the men Democrats (M = 190.38, p = .016) and the men Republicans (M = 187.06, p = .005).

Furthermore, all politicians referred to other people's family members more than their own. We recorded 47 personal and 671 other people family members in the women's and 347 personal and 2,871 others in the men's speeches. The results did not come as surprise because it would be perceived as inappropriate to refer to their family more than voters'. Interestingly, both the women and men referred to other men more than to women, which is in accordance with our results of the pronoun *he*/*she*. When mentioning their family members, the women more referred to male (29) than female (18) family members, whereas when mentioning other people's family members they referred to both male (336) and female members (335) equally. The difference was more obvious in the men's speeches. Their male family members were mentioned 206 and other people's 1,534 times, while their female family members were mentioned 140 and others 1,304 times. This pointed to men receiving more attention in both gender's speeches.

To compare, despite the mean ranks (M = 192.51 for the men and M = 213.78 for the women) recording higher occurrences in the women's speeches, the Mann-Whitney test on the friends references (U = 13333, Z = -1.623, p = .105) found no gender differences. A more detailed post hoc Kruskal-Wallis test showed that the men Republicans (M = 176.50) referred to friends significantly less (p = .005) than the women Democrats (M = 228.38). Similarly, the women (M = 210.71) and the men (M = 193.58) did not differ in mentioning human beings in general as proved by the Mann-Whitney test (U = 13647, Z = -1.305, p = .192). Yet, the men Senators (M = 140.66) spoke about other people significantly less (p = .000 in both cases) than the men Representatives (M = 212.44) and the women Representatives (M = 224.20).

Additionally, some Spearman two-tailed correlation tests showed that when speaking about family and friends, the politicians expressed positive emotions [r_s (393) = .239, p = .000 and r_s (393) = .244, p = .000]. Interestingly, we also found a negative correlation between the family and friends categories and tentativeness [r_s (393) = .-173, p = .001; r_s (393) = .-221, p = .000]. A possible explanation for this is that when the politicians became more emotional, they did not read a prepared speech and used more informal varieties, which made them appear more tentative.

We partially confirmed Schler et al.'s (2006), Newman et al.'s (2008) and Cunha et al.'s (2014) results by proving that the women did refer to family members more than the men; however, with friends and humans, there were no gender differences as in Xu et al.'s (2014) study. As expected, when referring to both their and other people's loved ones, the politicians expressed positive feelings and showed more insecurities in their speeches.

4.2.2 Affective processes

Numerous sociological and psychological research projects have inspired linguists to examine gender differences in expressing emotions. According to Trudgill (2000), gender differences in language are a reflection of different social roles and consequently behavior patterns. Humans are indeed emotional creatures; but, they use different strategies to cope with emotional situations (Cameron, 2005). A considerable amount of literature has so far been published on the gender differences in expressing emotions. Despite the stereotype of women being more emotional than men, research findings were conflicted. Studies by Gilligan (1982), Mulac et al., (1990), Zahn-Waxler et al., (1991), Jordan et al. (1991), Thomas and Murachyer, (2001) and Fivush and Buckner, (2000) demonstrated that women were more emotionally expressive and responsive, while men restrained from expressing their feelings. A recent study on gender and emotion differences in online collaboration by Iosub and colleagues (2014) pointed to the status as an important variable in expressing emotions. They found that female regular editors were the most relationship-oriented promoting social affiliation and emotional connections, while male administrators were the least emotional. To paraphrase, the higher status a person had, the less emotional they were.

In comparison, in a study of managers giving criticisms, Mulac et al. (2000) found that male managers were more emotional than female thus not confirming the stereotype. In the newest study by Ahmadi-Azad and Azad (2015), 103 female and 82 male narrative writings in EFL did not find a statistically significant difference in emotional linguistic content. The data showed a slight tendency of females using more emotional vocabulary. The results pointed to changing trends of equality among academically educated population.

To test gender differences in affective processes, we conducted the Mann-Whitney test whose results (U = 12728.5, Z = 55799.5, p = .026) showed a statistically significant difference, with the women (M = 219.1) using the category more than the men (M = 190.44). Therefore, we confirmed the idea of women being more emotionally expressive than men, as suggested by Gilligan (1982), Mulac et al., (1990), Zahn-Waxler et al., (1991), Jordan et al. (1991), Thomas and Murachyer, (2001), Fivush and Buckner, (2000) and Cameron (2005).

We also conducted the Spearman two-tailed correlation test and found a positive correlation of the politicians' expressing emotions and speaking about family and friends [r_s (393) = .243, p = .000 and r_s (393) = .175, p = .000]. Furthermore, a positive correlation was recorded with the death category [r_s (393) = .173, p = .001] and a negative correlation with the money category [r_s (393) = -.117, p = .020]. A possible explanation for these correlations may be in formality – when speaking about the serious money-related issues, feelings should not be and were not shown, i.e. when reporting another serious issue of someone's death, the politicians showed their human side by expressing emotions and not just reporting it as numbers.

As we were interested in the gender differences, we split our data file according to the gender groups and examined the same correlations. The male politicians' feelings positively correlated with friends, family and death [r_s (291) = .234, p = .000; r_s (291) = .205, p = .000 and r_s (291) = .240, p = .000] and negatively with money [r_s (291) = .121, p = .039]. Surprisingly, the female politicians' feelings positively correlated only with family [r_s (100) = .251, p = .011], while there was no correlation with friends, death or money [r_s (100) = .044, p = .663; r_s (100) = .048, p = .631 and r_s (100) = -.145, p = .145]. We may conclude that even though the women expressed more feelings in general, they oriented them towards someone's family members, while the men expressed fewer feelings but in relation to more diverse topics.

4.2.2.1 Positive and negative emotions

Studying linguistic expression of emotions in general was extended to studying gender differences in expressing positive and negative emotions. Studies which examined the gender

differences in expressing different types of positive feelings found identical results. A recent Iosub et al.'s (2014) study found that women expressed more positive feelings in the online context, which was in line with the previous studies by Mehl and Pennebaker (2003), Newman et al. (2008), Kivran-Swaine et al. (2012), Kucuktunc et al. (2012) and Schwartz et al. (2013). To put it another way, women were shown to express more positive feelings than men in both online and offline contexts.

While the positive emotions results were similarly reported by the researchers, the negative emotions results were contradictory. Mehl and Pennebaker (2003) and Newman et al. (2008) discovered that men used more negative emotions, while Mulac et al. (1990) and Thomson and Murachver (2001) proved that women used more negative emotions. The researchers (Gross & John, 1998; Seidlitz & Diener, 1998; O'Kearney & Dadds, 2004; Wolfson, 2005 and Chraif & Anitei, 2013) who studied and compared both categories found that women used greater intensity of both positive and negative emotions than men.

Finally, even though female's emotion words were more active, Fischer (1995), Bell et al. (2006) and Cuming (2013) did not find a statistically significant gender difference, i.e. Bell et al. (2006) believed that those results empirically supported biological but did not support social constructionist theories.

The positive emotions mean ranks (M = 193.47 for the men and M = 211 for the women) indicated that the women indeed expressed positive feelings more than the men; however, the Mann-Whitney test results (U = 13617, Z = 56688, p = .182, two-tailed) showed that the difference was not statistically significant, which is in accordance with Fischer's (1995), Bell et al.'s (2006) and Cuming's (2013) results. Additionally, the post hoc Kruskal-Wallis test revealed that the men Representatives (M = 179.39) expressed positive emotions significantly less than the men Senators (M = 232.99, p = .002) and the women Representatives (M = 252.82, p = .036). The Spearman two-tailed correlation tests revealed that the men expressed positive feelings when speaking about their family [r_s (291) = .262, p = .000], friends [r_s (291) = .290, p = .000] and achievements [r_s (291) = .412, p = .000] with a strong positive correlation with the *I* category [r_s (291) = .167, p = .004] and a negative correlation with the *we* category [r_s (291) = -.232, p = .000]. The women, on the other hand, expressed positive feelings related to their achievements [r_s (100) = .422, p = .000] with a

significant positive correlation with the *I* category [r_s (100) = .213, p = .032]. According to the results, we can draw a conclusion that even though the men and the women used positive feelings at the same frequency, the male politicians focused on the categories of their family, friends and accomplishments, while the female politicians expressed positive feelings in more categories within which they were significantly more focused on their achievements, which further supported our hypothesis of the women acting more formal in Congress.

In comparison to positive emotions, the Man-Whitney test results (U = 12955, Z = 56026, p = .045, two-tailed) pointed to a statistically significant difference on negative emotions usage, with the men (M = 191.22) using them less than the women (M = 217.49), which supported the results by Mulac et al. (1990), Thomson and Murachver (2001), Gross and John (1998), Seidlitz and Diener (1998), O'Kearney and Dadds (2004), Wolfson (2005) and Chraif and Anitei (2013). Furthermore, the post hoc Kruskal-Wallis test demonstrated that the men Senators (M = 158.68) expressed negative emotions significantly less than the men Representatives (M = 202.81, p = .022) and the women Representatives (M = 231.88, p = .000). In addition, the Spearman two-tailed correlation tests illustrated that the men used negative feelings when speaking about death [r_s (291) = .377, p = .000]. In the men's speeches, negative emotions negatively correlated with the pronoun *I* [r_s (291) = -.213, p = .000] and work topics [r_s (291) = -.201, p = .001], hence when expressing their negative attitude towards something, the male politicians exclude themselves from the context. To compare, the women used negative emotions to speak about someone's family members [r_s (100) = .222, p = .025], which might have happened in situations of commemorations and tributes. Similarly to the men' speeches, the women excluded themselves [r_s (100) = -.207, p = .037] from the context when expressing negative feelings. Based on the results, the female politicians expressed more negative feelings than the male politicians. However, the results suggested that the men focused on the results (death), while the women were more concerned about family members faced with traumatic situations.

4.2.2.2 Anxiety

Even though anxiety has been more studied by psychologists, linguists have recently shown interest in examining its realization in language. Regarding the gender differences,

Mulac et al. (1990), Thomas and Murachver (2001), Yu (2013) and Manjavacas (2015) found evidence of women linguistically expressing anxiety more than men. In a somewhat different study, Fehlman (2015) studied dialogue differences in American novels. Even though the authors themselves did not use anxiety-related language, they wrote novels in such way that female characters used anxiety-related linguistic expressions. In comparison, Ireland's (2008) study on American political candidates' speeches revealed that the female politician Palin was an extremely cheerful candidate using fewer anxiety words that the male politicians McCain, Obama and Biden. Finally, Xu et al. (2014) rejected the hypothesis that women use more anxiety words than men due to their results of no significant gender difference on the anxiety usage.

Our results obtained from the Mann-Whitney test (U = 13390, Z = -1.569, p = .117, two-tailed) showed that there were no statistically significant differences between the male (M = 192.70) and the female (M = 213.23) politicians on using anxiety-related vocabulary thus confirming recent Xu et al.'s (2014) results. Further, the two-tailed Spearman correlation tests showed that the politicians in general expressed anxiety when speaking about present [r_s (393) = .156, p = .002] and future [r_s (393) = .196, p = .000] money issues [r_s (393) = .142, p = .005], which is in line with Jordan and Pennebaker's (2015) research. However, here we found some gender differences. Namely, while the male politicians expressed their concern about present [r_s (291) = .185, p = .001] and future [r_s (291) = .225, p = .000] money issues [r_s (291) = .159, p = .007], the female politicians talked about someone's family members [r_s (100) = .252, p = .011] using shorter sentences [r_s (100) = -.274, p = .005], which is a stereotypical male-female behavior.

4.2.2.3 Anger

Central to the study of emotions is the concept of anger. Early research (Rudman, 1998; Heilman, 2001) on anger related it to men's speech indicating that societal expectations discouraged women from expressing feelings of anger. Moreover, when women expressed anger, its power was denied like "You're cute when you're mad" thus intensifying male power and female powerlessness (Lakoff, 2003: 163). Indeed, some researchers (Ashby Plant et al., 2000; Mehl & Pennebaker, 2003; Wang & Hseih, 2007; Ireland, 2008 and Schwartz et

al., 2013) confirmed the previous findings of men expressing their emotions by using anger-related references more than women.

By contrast, Chentsova-Dutton and Tsai (2007) found that women expressed all emotions, including anger, more than men. Further, in a study on Hillary Clinton's language, which was hypothesized to have changed over the years, Jones (2015) discovered a linguistic shift during Clinton's political career, i.e. she used a masculine index of anger the most in her Secretary of State years (2009-2013) confirming the notion of her language becoming more masculine over time. The only study we found that did not replicate the gender differences findings in expressing anger was Newman et al.'s (2008).

The Mann-Whitney results from our study (U = 14891, Z = -.052, p = .958, two-tailed) demonstrated that there were no significant differences in the men (M = 197.82) and the women (M = 198.51) expressing the feelings of anger, which confirmed Newman et al.'s (2008) findings. Further statistical analysis revealed that the men who completed secondary education (M = 326.57) felt more freely than any other politicians in expressing their anger feelings. Furthermore, we wanted to see which contexts the politicians expressed their anger in and found some subtle gender differences obtained by the two-tailed Spearman correlation test. While the female politicians expressed anger only when speaking about death [r_s (100) = .361, p = .000], the male politicians also talked about death [r_s (292) = .526, p = .000] using longer words [r_s (292) = .120, p = .040], thus increasing the level of their speech formality. They further expressed concern about future events and potential consequences [r_s (292) = .118, p = .044]. A possible explanation for the anger and death correlation may be related to the politicians expressing their opinion on the issue of American soldiers sent to wars overseas, with the male politicians trying to find solutions and expressing concern about future consequences should the issue not be resolved.

4.2.2.4 Sadness

Traditionally, men are not supposed to cry or express sadness since that is perceived as an expression of powerlessness and helplessness (Lakoff, 2003: 163). Therefore, the findings of women expressing the emotion of sadness more than men did not come as surprise to the

researchers Mulac et al. (1990), Thomas and Murachver (2001), Wolfson (2005) and Bamman et al. (2014). However, several researchers reexamined the stereotype and found contrary results. Wang and Hsieh (2007) discovered that boys expressed sadness more than girls who were more focused on harmonious relationships. The following year, Ireland revealed that the male politicians (Biden, McCain and Obama), referred to sad emotions more than Sarah Palin, which is contrary to the running stereotype. The same year, Craig examined the gender stereotypes in the children's picture books and was unable to find a significant gender difference, i.e. both male and female characters used the emotion of sadness equally.

The Mann-Whitney results from our research (U = 11982, Z = -2.983, p = .003, two-tailed) pointed to a statistically significant difference between the male (M = 187.89) and the female (M = 227.03) politicians expressing sad emotions, supporting the idea of women expressing sadness more than men as proved by Mulac et al. (1990), Thomas and Murachver (2001), Wolfson (2005) and Bamman et al. (2014). A further analysis by the post hoc Kruskal-Wallis test illustrated both within and among groups differences. Namely, the women Representatives (M = 247.06) the most freely expressed their sadness with a recorded significant difference compared to the women Senators (M = 144.90, p = .002), the men Representatives (M = 196.86, p = .004) and the men Senators (M = 162.74, p = .000). Furthermore, the gender differences were recorded between different parties. The women Democrats (M = 237.14) more readily talked about sadness in comparison to the men Democrats (M = 194.22, p = .047) and the men Republicans (M = 182.63, p = .003). To summarize, the women Democrats who served in the House of Representatives (M = 262.17) expressed the feelings of sadness the most, while the women Democrats from the Senate (M = 137, p = .003) and the men Democrats from the Senate (M = 157.36, p = .000) did so the least.

The Spearman correlation test (two-tailed) revealed some additional gender differences. While the female politicians expressed sadness talking about health issues [r_s (100) = .291, p = .003] and people's family members [r_s (100) = .201, p = .043], the male politicians talked about death [r_s (292) = .118, p = .044], health [r_s (292) = .118, p = .044], family members [r_s (292) = .136, p = .020] and humans [r_s (292) = .141, p = .016]. These findings are in line with our positive emotions results with the male politicians expressing

fewer emotions than the females but when they did express sadness, they used it in wider range, whereas the women expressed more sadness emotions focusing on fewer topics.

4.2.3 Cognitive processes

Cognitive processes include mechanisms that reveal different models of one's thinking such as causality, tentativeness, inclusion, self-reflection, etc. Cognitive words typically suggest that people are trying to comprehend what is happening in their lives. Research dealing with the gender differences in using the category have reported contradictory results. Mulac and Lundell (1994) found that women used more cognitive verbs in written language and Hartman (1976), Poole (1979) and Ireland (2008) found the same results for spoken language. Pennebaker (2013) found that women used more cognitive words than men. Believing that cognitive words reflect different ways of insight, causal thinking and related dimensions, Pennebaker's (2013) result that shows women use more cognitive words than men, in Pennebaker's opinion, refuted Aristotle's belief that women are incapable of philosophical thought and rational thinking like men. In a recent study on politicians' language, Pennebaker and Jordan (2015) found that Hillary Clinton scored the highest and Bernie Sanders the lowest on cognitive processes words. The authors believed that it happened because Clinton worked through the issues as they came up, while Sanders had already made up his mind.

In a study on the people's biography on Wikipedia with 84% male contributors, Graells-Garrido and colleagues (2015) discovered that men were described with words related to cognitive processes and work concerns as these two aspects are believed to be more important in men's lives. No strong evidence of the gender differences was found by Kapidžić and Herring (2011) in chat messages, which were generally very low in cognitive expressions.

In order to examine the gender distribution of cognitive processes words, we conducted the independent sample t-test whose results [t (393) = .601, p = .548, two-tailed] pointed to the men (M = 14.6, SD = 2.22) and the women (M = 14.44, SD = 1.99) using cognitive words equally, thus confirming Kapidžić and Herrings' (2011) results. Taking Pennebaker and Jordan's (2015) interpretation into account, we may conclude that our

politicians equally thought about and made up their minds on issues. However, the cognitive processes subcategories might be more revealing and informative on the politicians' type of processing.

4.2.3.1 Insight

According to Pennebaker (2013), insight is the category which often suggests that a speaker is more self-referent and focused on the underlying meaning of a subject. Several researchers dealt with the gender differences in using insight-related words. In a study of suicide notes written by Australian men and women, Lester and colleagues (2010) found that the notes written by women contained more insight words thus being self-referent. Brónnimann et al. (2013) took a different approach in the analysis of witness narratives. They were interested in whether witnesses, regardless of their gender, composed their testimonies differently depending on an interviewer's gender. Their results showed that witnesses used more insight-related words when being interviewed by females. A clear difference was found by Epstein et al. (2005) who reported that men used more insight words in their writing believing for them to be more problem-focused.

It is apparent from our Mann-Whitney test results ($U = 13457$, $Z = -1.496$, $p = .135$, two-tailed) that the male politicians ($M = 203.07$) used the category of insight-related words at a higher rate than the female ($M = 183.43$) politicians; however, the difference was not statistically significant, which is not in line with any of the previous results. Bearing in mind Pennebaker's (2013) ideas of insight vocabulary being related to rationalization and analytical thinking, we can conclude that both the male and the female politicians equally processed and linguistically expressed these issues.

4.2.3.2 Causation

Maltz and Borker's (1982) and Gilligan's (1982, 1987) presumptions that women were more nurturing and other-oriented which was realized in the use of more causal words was

challenged by Epstein et al. (2005) and Graells-Garrido et al. (2015) who found evidence of the men's greater usage of causality words believing for them to be more problem-oriented. Furthermore, the presumption was rejected by Newman et al. (2008), Bell et al. (2012) and Manjavacas (2015) who did not find any gender difference on causal words use.

Our findings, done by the Mann-Whitney test (U = 12669.5, Z = -2.289, p = .022, two-tailed) pointed to the women (M = 220.29) using causality more than the men (M = 190.24). These results seem to be consistent with the ones done by Maltz and Borker and Gilligan. However, we believe that these findings further support the idea of the female politicians being analytical thinkers previously proved by the insight-related word results. We also believe that it is possible to explain these findings by the women's lack of security and the need to explain and support their ideas. For example, when proposing a bill or explaining their actions, the women might have felt the need to elaborate on the causality of the topic, while the men, feeling more secure, might have proposed a bill or reported actions without more detailed elaboration.

4.2.3.3 Discrepancy

The discrepancy words category is mostly composed of modal verbs which are accompanied by words related to wishes, desires, hopes, ideals, regrets, etc. The category is of interest because it suggests discrepancy between how the world looks like and how it should, could or must be. Hence, by studying discrepancy words we may draw conclusions on a speaker's dissatisfaction and aspirations for changes. According to McMillan et al. (1977), Biber et al. (1998), Mulac et al. (2001) and Mehl and Pennebaker (2003), the usage of discrepancy words, especially the modal verb *could*, is the feature of women's language. The authors believed that by using discrepancy words, women tended to express their tentativeness.

Conducting the Mann-Whitney two-tailed test (U = 13611.5, Z = -1.341, p = .180), we were unable to find the differences between the men (M = 202.54) and the women (M = 184.95) on using discrepancy words. Hence, we could not confirm the previous findings. A possible explanation for this is the setting since discrepancy words have proven not to be used

in formal settings (Pennebaker, 2013). Therefore, we may conclude that the politicians were equally (dis)satisfied with the current situation and equally hypothesized about past, present and future situations.

4.2.3.4 Tentativeness

One of the central issues to the entire discipline of gender differences is the concept of tentativeness. 1975 was a breakthrough year because Lakoff published her *Language and Woman's Place* in which she described women's language, due to the usage of hedges, tag questions, intensifiers and disclaimers/qualifiers, as tentative. Since then, numerous research have examined the issue. O'Barr and Atkins (1980) disagreed with Lakoff in characterization. They believed that tentative language should not be considered as women's but as powerless language since it is used by people in powerless positions and by uneducated people (O'Barr, 1982). Simultaneously, tentative language is a reflection on the American society in which women have subordinate position, which is the reason for the equalization. Similarly, Harris (1984) disagreed with tentative language (especially question tags) expressing uncertainty; rather it is a request for confirmation. Hence, people who use questions tags are to be considered as powerful and authoritative.

Furthermore, on the sample of 3,502 participants, in spite of a small effect size, Leaper and Robnett (2011) confirmed that women were more likely than men to speak tentatively, thus supporting Basow and Rubenfield's (2003) findings. Generally, people who speak tentatively are evaluated as less knowledgeable than those speaking assertively (Erikson et al. 1978; Wiley & Eskilson, 1985). However, later research have proven that men and women were not perceived equally even when they spoke identically. Bradley (1981) found that tentative women were perceived as less intelligent and insightful than tentative men. Carli (1990) did an experiment on the perception of men and women using both persuasive and tentative speech. The experiment showed that women who used assertive speech were perceived as more influential than tentative women by women listeners. The result was reverse with men listeners who perceived tentative women as more influential than assertive women. Interestingly, men were perceived equally knowledgeable and likeable regardless of the assertiveness or tentativeness of their speech. Reid et al. (2003) showed that under certain

conditions, tentative women were more persuasive with men listeners, yet, judged less favorably.

Built on the previous research (Reid et al. 2003; Palomares, 2004, 2008 and Palomares et al. 2004), Palorames (2009) examined tentative speech in different topics. The results demonstrated that with masculine topics, women were more tentative in intergroup than in intragroup conversations. Counter-stereotypically, with feminine topics, men were more tentative in intergroup contexts and no differences emerged with gender-neutral topics. Not taking topic-related contexts into account, Zimmerman and West (1975), Crosby and Nyquist (1977), Brouwer et al. (1979), Schmader et al. (2007) and Newman et al. (2008) did not find the gender differences in tentativeness.

Aiming to examine the gender differences in tentative speech, we carried on the Mann-Whitney test whose results ($U = 11076.5$, $Z = -3.894$, $p = .000$, two-tailed) clearly demonstrated the existence of gender differences. Interestingly, the mean ranks pointed to the male politicians ($M = 211.20$) as more tentative than the female politicians ($M = 160.09$). Therefore, our findings did not confirm any of the previous ones claiming that women are more tentative than men. Further analysis showed that the men Representatives ($M = 213.39$, $p = .001$) and the men Senators ($M = 205.03$, $p = .036$) were significantly more tentative than the women Representatives ($M = 155.30$). Additionally, we also found the gender differences in party affiliations. Namely, the women Democrats ($M = 160.25$) were significantly less tentative than the men Democrats ($M = 209.17$, $p = .015$) and the men Republicans ($M = 212.88$, $p = .005$). Overall, the women Democrats from the House of Representatives ($M = 153.61$) were the least, whereas the men Republicans from the Senate ($M = 224.35$) the most tentative speakers.

We also tested O'Barr's (1982) hypothesis that tentative language is used by uneducated speakers and we could support his ideas as far as the women's language was concerned. The women holding a PhD degree ($M = 92.72$) were the least and the women who had secondary education ($M = 238.50$) the most tentative. On the other hand, the men with the undergraduate level of education were the least ($M = 168.56$) and the men with the graduate level of education ($M = 220.13$) the most tentative speakers.

Furthermore, we conducted the two-tailed Spearman correlation tests and found a negative correlation between tentativeness, words per sentence and long words. To clarify, in both the men's [r_s (291) = -.242, p = .000; r_s (291) = -.417, p = .000] and the women's [r_s (100) = -.282, p = .004; r_s (100) = -.245, p = .013] speeches, tentativeness was realized with the usage of shorter sentences and simpler vocabulary. While the men used fewer numbers [r_s (291) = -.162, p = .005] and more vulgar language [r_s (291) = .116, p = .046] when speaking tentatively, the women did not [r_s (100) = -.042, p = .678; r_s (100) = .163, p = .101]. In addition, we were unable to confirm Reid et al.'s (2003), Palomares' (2004, 2008, 2009) and Palomares et al.'s (2004) findings on the correlation of tentativeness and topics since both the men [r_s (291) = -.319, p = .000] and the women [r_s (100) = -.380, p = .000] were extremely assertive when speaking about work and achievements and tentative when speaking about money-related issues [r_s (291) = .183, p = .002 men; r_s (100) = .257, p = .000 women].

According to our findings, we may conclude that a generally subordinate women's position in society inspired them to employ means to become more visible. The men using more tentative speech than the women in Congress may be explained by the women being more work-oriented and having prepared for their speeches more thoroughly, which was reflected in their more assertive speech. In other words, the women successfully adapted to the workplace which respects masculine traits. Yet, since there is no place for submissive behavior in Congress, the assertive women might have challenged tentative men's status position.

4.2.3.5 Certainty

The certainty category mostly consists of intensive adverbs (absolutely, definitely, undoubtedly, extremely, etc.), adverbials of frequency (always, never), some modal verbs (must*, have*) etc. all of which have received some attention by researchers. The majority of research (McMillan et al., 1977; Mulac & Lundell, 1986; Jaffe et al., 1995; Biber et al., 1998; Mulac et al., 2000; Mondorf, 2003; Mehl & Pennebaker, 2003; Pennebaker et al., 2003 and Newman et al., 2008) in various both written and spoken contexts found that women consistently used certainty words more than men believing for it to be a typical feminine

language feature. The only study so far which did not report the gender difference in the usage of certainty words was Schmader et al.'s (2007).

Even though the mean ranks pointed to the women (M = 205.66) using certainty words more than the men (M = 195.33), the Mann-Whitney results (U = 14162, Z = -.786, p = .432) showed that the difference was not statistically significant. Hence, our findings are in agreement with those obtained by Schmader et al. (2007). Since the usage of certainty words points to analytical thinking (Pennebaker, 2013), a conclusion of the male and the female politicians thinking analytically at the same rate can be made. This result is in line with our insight results, which are also related to the analytical thinking process.

We also tried to find the gender differences in the contexts specific certainty words were used in but could not find them. The only certainty word which was used differently by the male and the female politicians was the word *undoubtedly*. The word *undoubtedly* was used 4 times by the female and 34 times by the male politicians. In all 4 cases, the women used the word *undoubtedly* when speaking about other people's contributions, exemplified in (39), while the men more frequently used it to express their strong point of view on a certain issue as in example (40).

39) Cameron's decision truly demonstrates the strength of his character, but perhaps most importantly, his selfless act will ***undoubtedly*** never be forgotten by the man who received another chance at life. (Jeanne Shaheen, Senate - May 9, 2013; Tribute to Cameron Lyle)

40) The situation in Syria is ***undoubtedly*** grim and Egypt faces a prolonged period of instability, but the news is not uniformly bad. (Adam Schiff, House of Representatives - October 23, 2013; Sustaining the Arab Spring)

4.2.3.6 Inhibition

Since the issue of inhibition involves a psychological perspective and is more oriented on studying individual's emotions and perceptions from a psychological standpoint, receiving

almost no attention by linguists, let alone sociolinguists, comes as no surprise. The LIWC category of inhibition consists of numerous words related to restrain and suppress such as *avoid, ban, block, constrain, neglect, prohibit, stop,* etc. To put it differently, LIWC calculates the frequency of using inhibition related vocabulary and not how restrained participants are in their speeches. The two studies which dealt with inhibition (McClelland, 1979 and Ireland, 2008) found a positive correlation between inhibition and negative emotions claiming that they are markers of self-restraint. Our Spearman correlation test [r_s (395) = .323, p = .000] confirmed McClelland's (1979) and Ireland's (2008) results of a positive correlation of inhibition and negative emotions. We can conclude that when using inhibition related vocabulary, the politicians from our study expressed negative emotions.

Apart from the correlation, Ireland (2008) studied gender differences in individual political speeches and found that Sarah Palin used the least, while John Biden the most inhibition words concluding that Palin was the least, while Biden the most restrained politician. The Mann-Whitney results from our study (U = 13744.5, Z = -1.207, p = .227, two-tailed) showed that the male (M = 193.91) and the female (M = 209.75) politicians as groups were equally self-restraining. They probably believed that they were elected to actively speak their mind trying to make positive changes and did not feel the need for self-control.

Even though the Mann-Whitney results suggested there were no gender differences in the usage of the inhibition category, we examined the contexts words were used in aiming to find potential gender differences in the usage of specific words. The only inhibition word the politicians used differently was the word *neglect* as shown in examples (40) and (41).

40) She left foster care at age 4 only to return at age 15 because of ongoing ***neglect*** and abuse. (Karen Bass, House of Representatives - May 21, 2013; Congressional Foster Youth Shadow Day)

41) Meanwhile, we ***are neglecting*** other urgent national priorities. How about the jobs deficit, the deficit in our investment in our infrastructure, the deficit in our investment in a strong, growing, middle class? (Tom Harkin, Senate - February 28, 2013; Sequestration)

We recorded 11 occurrences of the word *neglect* in the women's and 43 in the men's speeches. While the male politicians used the word in different contexts, the female politicians used it only when referring to child neglect and abuse. Such usage was rather expected as women have always been seen as more nurturing and caring. The usage of the word *neglect* by the female politicians from our study pointed to a stereotypical female behavior of caring mothers.

4.2.3.7 Inclusive and exclusive words

Inclusive and exclusive words have widely been studied in terms of individual's personality, i.e. sociolinguists have rarely dealt with the issue. Yet, some attention has recently been directed to gender differences in the use of inclusive and exclusive words. In the studies of gender differences in workplaces, Eagly et al. (2003) and Ng and Leung (2015) found that women in leadership positions used more inclusive words than men. Gorbatai and Nelson (2015) confirmed the findings that women used inclusive words in business context. In Ireland's (2008) politicians' study, Palin frequently used inclusive words, which might have indicated rambling.

Conducting a statistical analysis by the Mann-Whitney test, the results (U = 14464, Z = -.482, p = .630, two-tailed) inspired us to conclude that the male (M = 199.63) and the female (M = 193.30) politicians used inclusive words equally. However, we did find some correlation gender similarities and differences by conducting the two-tailed Spearman correlation test. Both the men's [r_s (291) = .447, p = .000] and women's [r_s (100) = .372, p = .000] inclusive word positive correlated with the pronoun *we*, which was rather expected. While there was a positive correlation between family [r_s (100) = .242, p = .014], friends [r_s (100) = .223, p = .024] and inclusive words in the women's speeches, no such correlation [rs (291) = .000, p = .995 for family and rs (291) = .039, p = .502 for friends] was detected in the men's speeches. However, we recorded a negative correlation of inclusive words and negative emotions [r_s (291) = -.182, p = .002] in the men's speeches meaning that whenever they included themselves in the context, they expressed very few negative emotion and criticism words.

Our results are not in line with any of the previous ones. We may draw a conclusion that both the male and the female politicians were equally self-conscientious and interdependent (Oberlander & Gill, 2006) with a distinction of the female politicians being more focused to include other people in the context whereas the male politicians concentrated on their positive self-presentation.

Regarding exclusive words, the findings so far have been contradictory. McGregor (2010) found that women used more exclusive words, while Ireland (2008) recorded few exclusive words in women's and a lot more in men's political speeches believing that few exclusive words might indicate dishonesty. Newman et al. (2008) did not record gender differences in the category usage.

We conducted the independent sample t-test whose results [t (214) = 2.9, p = .004, two-tailed] showed that the men (M = 1.6, SD = .73) used more exclusive words that the women (M = 1.39, SD = .60), which is in the agreement with Ireland's (2008) results. A one-way ANOVA revealed some additional intergroup and intragroup differences. Namely, the men Representatives (M = 1.68, SD .78) used exclusive words significantly more than the men Senators (M = 1.37, SD = .53) and the women Representatives (M = 1.40, SD = .63). More precisely, the women Republicans from the Senate (M = 1.21, SD = .28) used them the least, whereas the men Republicans from the House (M = 1.68, SD = .78) the most. Since exclusive words are used when one wants to make a distinction between the concepts that lie within or outside of a given domain (Pennebaker, 2013), we may draw a conclusion that the male politicians were more direct by specifically excluding something from a category whereas the female politicians did not make a clear distinction.

4.2.4 Perceptual processes

The perceptual processes category includes words related to perceptual and sensory concept divided in three subcategories of visual, auditory and tactile sensory concepts. Several researchers have examined the gender differences in perceptual processes distribution. Perceptual processes words have been found in women's writing and speech by Hartman (1976), Poole (1979), Mulac and Lundell (1994), Yale (2007), Newman et al. (2008) and

Gorbatai and Nelson (2015), whereas Brónnimann et al. (2013) did not find any gender difference in the usage of the category.

The Mann-Whitney results from our study (U = 14602, Z = -.343, p = .731, two-tailed) confirmed Brónnimann et al.'s (2013) findings of no differences between the men (M = 199.16) and the women (M = 194.66) in the usage of perceptual processes words. However, we did record some differences by the post hoc Kruskal-Wallis tests. Namely, the men Republicans (M = 217.03, p = .020) were significantly more perceptual than the men Democrats (M = 177.67). Likewise, the women Representatives (M = 213.17) who were more perceptual than the women Senators (M = 118.75, p = .005), the men Representatives (M = 222.48) were more perceptual than the men Senators (M = 133.75, p = .000). Overall, the men Democrat Senators (M = 131.69) were the least and the men Republican Representatives (M = 242.95) the most perceptual. Similarly, the women Republican Senators (M = 115.62) scored the lowest and the women Republican Representatives (M = 227.86) the highest on the perceptual words use. According to the results, we may draw a conclusion that the female and the male politicians were equally perceptual. However, gender differences may occur in specific sensory concepts which we will elaborate on in the following subsections.

4.2.4.1 Visualy-related words

Even though sensory concepts of all three types have been more studied from a psychological perspective, we believe psychological research might be beneficial in our case as well because they might be positively correlated with a particular sensory concept and its linguistic expression. The LIWC category of visual words calculates the use frequency of all words related to vision (such as colors and the manners in which one sees something). The majority of research (Sherman, 1978; Kirk, 1992; Eisenman, 1997; Reiman; Cameron, 2007, Ardila et al., 2011), with an exception of Merten and Beal (1999) who did not find gender differences, recorded that men, regardless of their age, outperform women in visual tasks, hence men are more visually dominant than women.

In our study, however, when it comes to the linguistic expression of what one had seen, the men (M = 201.66) did not score higher that the women (M = 187.48) proved by the

Mann-Whitney test (U = 13870, Z = -1.082, p = .279, two-tailed). Yet, as recorded by the post hoc Kruskal-Wallis test, the men Republicans (M = 218.77) used visually-related words significantly more than the men Democrats (M = 181.08, p = .029) and the women Democrats (M = 177.06, p = .045). Furthermore, while the men's visual words positively correlated with the pronouns *I* and *we* [r_s (292) = .147, p = .012; r_s (292) = .148, p = .011], the women's did not [r_s (100) = -.146, p = .143; r_s (100) = -.070, p = .482]. Therefore, our findings confirmed Merten and Beal's (1999) of no gender differences in the use of visually-related words. However, the subgroup results did point to the men using more visual words especially when reporting their own and the group they affiliate themselves to visual experience.

4.2.4.2 Auditory-related words

Yale (2007) found that women used hearing-related words more than men. In comparison, the same year Wehrwein and colleagues studied learning preference styles and found that male students preferred auditory style, while female students liked it only in a combination with other styles. With this LIWC category, one can analyze the use frequency of auditory-related vocabulary (hear, deaf, music, scream, speech, loud, yell, etc.).

Our Mann-Whitney results (U = 14484, Z = -.462, p = .644) demonstrated no differences between the men (M = 199.57) and the women (M = 193.50) on using auditory-related vocabulary. Further post hoc Kruskal-Wallis showed intragroup differences. Namely, the women Representatives (M = 218.32) used hearing-related words more than the women Senators (M = 91.75) the same as the men Representatives (M = 237.61) used them more than the men Senators (M = 92.28), which is in accordance with our visual-related words results. Generally, the women Republican Senators (M = 53.62) and the men Democrat Senators (M = 75.89) used the category the least, while the women Republican Representatives (M = 226.17) and the men Republican Representatives (M = 249.69) the most. Additionally, the Spearman correlation tests showed that both the male and the female politicians reported what they had heard [r_s (292) = .197, p = .001; r_s (100) = .219, p = .027] about other people [r_s (292) = .349, p = .000; r_s (100) = .324, p = .000]. However, the difference was that the male politicians were rather tentative [r_s (292) = .253, p = .000] reporting that, whereas the females were not

[r_s (100) = .055, p = .584], which might indicate that the male politicians did not find reported stories as reliable as the women did.

4.2.4.3 Tactile-related words

The previous research (Kirk, 1992; Moir & Jessel, 1992) that dealt with the gender differences in tactile sensory concepts found that women outperformed men in tactile skills. Yale's (2007) research proved that women's outperformance was reflected in their higher usage of tactile words. In comparison, a more recent research by Manson (2014) proved otherwise – that men were more tactile-oriented.

In our study, the Mann-Whitney test results (U = 14878, Z = -.066, p = .948, two-tailed) pointed to no gender differences, i.e. the men (M = 197.78) used tactile-related words equally as the women (M = 198.64). However, consistent with our visually and auditory related vocabulary results, intragroup differences on gender and chamber were recorded by the post hoc Kruskal-Wallis test. The men Senators (M = 238.25) used more tactile-related words than the men Representatives (M = 183.35), which is inconsistent with the previous results when the Representatives were more sensory-oriented than the Senators. Since there was a negative correlation between the category and the pronoun *I* [r_s (100) = -.227, p = .022] and a positive correlation with the pronoun *they* [r_s (100) = .312, p = .001] in the women's speeches recorded by the Spearman correlation test, we may conclude that when describing tactile sensory feelings, the female politicians referred to other people and not themselves, which may be due to the women caring for other people or refusing to share their own feelings, which supports our idea of the women perceiving Congress as their job where there is no place for sharing personal stories and feelings.

4.2.5 Biological processes

The biological processes category includes words related to body parts, health, intercourse and eating. Neither the general category nor the subcategories have received much attention by linguists so far. Our results might be valuable for other researchers should they

decide to examine gender linguistic differences in the usage of the biological processes category and its relevant subcategories.

Aiming to examine the existence of gender differences in the biological process category, we conducted the Mann-Whitney test whose results (U = 14595.5, Z = -.359, p = .720, two-tailed) demonstrated that the male (M = 196.81) and the female (M = 201.41) politicians used the category equally. No further intragroup or intergroup differences with respect to other categorical variables (education, party and chamber) were recorded. These results are not consistent with the data obtained by Nagarajan and Hearst (2009) who found that men used the category significantly more than women. The lack of other researchers' interest in the category and our results of no gender differences might lead to a conclusion that biological processes are not a gender-specific category. Yet, the subcategories might be more revealing.

4.2.5.1 Body parts

As the name suggests, the body parts subcategory is composed of all body parts; both formal and vulgar forms. Following the Mann-Whitney results (U = 14349.5, Z = -.599, p = .549, two-tailed) which proved no differences between the men (M = 200.03) and the women (M = 192.18), we decided to examine if the female and the male politicians used female and male body parts respectively. Before reporting the results, we need to say that the politicians very rarely referred to any body parts whatsoever.

There were only two body parts that were used differently by the men and the women. With only 4 occurrences, the male politicians used the word *prostate* in the contexts of raising awareness of both timely examinations and prostate cancer. In comparison, the women never mentioned the word. Another difference was in the usage of the word *breast* which was mentioned 34 times in the men's and 52 times in the women's speeches. Given the significantly lower number of the total words in the women's speeches, the women used the word *breast* at higher frequency than the men. However, a closer look of the word contextual usage revealed another difference. While the women used the word solely when speaking about cancer trying to raise awareness, during which they 8 times referred to either

themselves or other ill women, the men, who 7 times referred to ill women, also used the word when speaking about the issue of breastfeeding. Since breast cancer can happen to anyone (even men, though rarely) compared to breastfeeding which only women can do, the female politicians only speaking about the illness and not breastfeeding might be interpreted as them perceiving the latter as taboo.

In addition, the Spearman correlation test showed that when speaking about body parts, among which the word *breast* is the most frequently used by both the men and the women, the men frequently talked about death consequences [r_s (292) = .232, p = .000, two-tailed], while the women did not [r_s (100) = .189, p = .059, two-tailed]. Since some of the female politicians who talked about the issue of breast cancer were patients themselves, the women not referring to death consequences may possibly be interpreted as them not wanting to relate death consequences with the illness they suffer from.

4.2.5.2 Health

Several studies have examined the issue of the gender differences in relation to health. Umberson (1992) and O'Brien et al. (2005) demonstrated women's involvement in monitoring and ensuring family health. Men tended to rely on their female partners to recognize symptoms and persuade them to seek medical help. Seale (2006) extended the research to the online cancer support groups and confirmed women's greater involvement even in the prostate cancer forums concluding that women expressed more concern in health matters in general. Furthermore, in the author profiling study, Soler (2013) extracted health related words as those mainly used by female authors. The findings indicating that women deal with health issues more than men were proven by Cunha and colleagues (2014) in a study on social media statuses. Finally, according to Jones (2015), the fact that Hillary Clinton was in charge of the health reform further supported the idea of women being more health conscious.

Even though the mean scores from our analysis demonstrated a slight tendency for the women (M = 211.60) to use health-related words more than the men (M = 193.27), the Mann-Whitney results (U = 13556, Z = -1.397, p = .162, two-tailed) showed that the difference was

not statistically significant. Additionally, we were interested in possible gender differences in specific health issues. With 18 in the men's and 0 occurrences in the women's speeches, the results showed that the men dealt with the dentistry topic whereas the women did not. On the other hand, with 0 occurrences for anorexia and bulimia in the men's and 5 and 4 respectively in the women's speeches, it was clear that the men were not interested in eating disorders. This result was not very surprising because eating disorders are the problems mostly women are faced with; therefore, this may have happened because the women expressed their concern and support for all girls and women who face the problem. Almost no differences were found in alcohol and cancer related words; yet, a slightly higher frequency was recorded in the women's speeches. Finally, even though the difference was not significant, rehabilitation issues, especially war veterans, were discussed by the men more, which was expected. Women can more easily identify themselves with the problems of anorexia and bulimia, while men can more readily identify themselves with other men who served in wars.

In addition, interesting findings were recorded with the Spearman correlation tests. When speaking about health issues, the men frequently alluded to death consequences [r_s (292) = .215, p = .000, two-tailed], while the women did not [r_s (100) = .099, p = .322, two-tailed]. Also, health words from the men's speeches positively correlated with money-related words [r_s (292) = .116, p = .048, two-tailed], whereas the women's did not [r_s (100) = .110, p = .271, two-tailed]. These correlations may be explained by the men being more concerned about treatment costs of patients who will eventually die. This pessimistic men's view is in line with our body words results.

4.2.5.3 Sexual

Very little is known about gender distribution of sex-related words. A recent study by Graells-Garrido et al. (2015) found that sexual processes words dominated in women's speeches. Despite the mean ranks from our study pointing to the women's (M = 213) higher usage of the subcategory compared to the men's (M = 192.78), the Mann-Whitney results (U = 13413.5, Z = -1.601, p = .109) recorded no statistically significant gender difference. However, the post hoc Kruskal-Wallis test showed that the men Representatives (M = 174.16) talked about sexual processes significantly less than the men Senators (M = 245.02, p = .000)

or the women Senators (M = 264.30, p = .003). A more detailed analysis revealed additional differences, though not statistically significant. Namely, the women dealt with the problems of incest, abortion and rape more than the men, which is rather expected because all three problems are more related to girls or women, hence the female politicians could identify with victims and their problems more easily.

4.2.5.4 Ingestion

Data about the gender differences in the use of ingestions words are limited to Abbar and colleagues' study (2015) who found that women tweeted about food, especially low-fat, more than men. However, since the LIWC ingestion subcategory is restricted to types of drinks and meals in more general sense, we could not test the findings. The results from our Mann-Whitney study (U = 14595.5, Z = -.359, p = .720) did not show statistically significant difference between the male (M = 196.81) and the female (M = 201.41) politicians on the use of ingestion words. Though, the post hoc Kruskal-Wallis tests pointed to intragroup differences with respect to the party affiliation and chamber. Precisely, likewise the men Republicans (M = 172.40) who talked about ingestion significantly less than the men Democrats (M = 226.18, p = .000), the men Representatives (M = 183.44) did not refer to the subcategory as much as the men Senators (M = 234.32, p = .003). The gender equality on the subcategory usage may be explained by the nature of the setting and consequently the lack of contextual encouragement for the use.

4.2.6 Relativity

To date, very little is known about the relativity words distribution among men and women. Mirroring Nardi et al.'s (2000) findings on instant messaging in the workplace, Yale (2007) found the gender differences in the relativity category. Namely, he proved that men used words referencing the past, while women focused on present and future believing that was related to women using language for creating and maintain relationships. Studying the category in general, Nagarajan and Hearst (2009) confirmed the females' usage of relativity words at a higher frequency than men in their profiles study.

We were unable to support the previous findings since our Mann-Whitney results (U = 14205.5, Z = -.743, p = .458, two-tailed) showed that the male (M = 200.52) and the female (M = 190.77) politicians used the relativity words at an equal rate. A possible explanation for this gender equality can be in the relativity words subcategories which show the opposite results which will be elaborated on in the following subsections. To paraphrase, one subcategory was claimed to be dominantly used by the men and other by the women, which might have equalized the effect for the general category of relativity words.

4.2.6.1 Motion

The study of motion from a linguistic perspective is restricted to Isaac et al.'s (2011) research on 227 male and 70 female medical students who applied for a residency program. The authors examined the medical student performance evaluation letters. Their results showed that female students used significantly more motion-related words than male students. We tested their hypothesis by carrying on the Mann-Whitney test. The results (U = 14198, Z = -.750, p = .453, two-tailed) showed that the men (M = 195.46) and the women (M = 205.30) identically used motion references, hence, we did not confirm Isaac et al.'s (2011) results. Furthermore, in a study on the American politicians' linguistic expression, Ireland (2008) found that Democrats, regardless of their gender, used more motion verbs thus being more concrete and restrained. With the mean ranks of 206.69 for the Democrats and 187.83 for the Republicans, the Mann-Whitney results (U = 17532, Z = -1.637, p = .102, two-tailed) did not point to a significant difference, i.e. we could not support Ireland's (2008) findings.

The only gender differences we found were recorded by the Spearman correlation tests. When using motion words, the male politicians used shorter sentences and less complex words [r_s (292) = -.127, p = .030; r_s (292) = -.172, p = .003, two-tailed] compared to the female politicians in whose speeches this correlation was not recorded [r_s (100) = -.030, p = .761; r_s (100) = .070, p = .484, two-tailed]. In addition, both the male and the female politicians' motion words positively correlated with money references [r_s (292) = .224, p = .000; r_s (100) = .229, p = .021, two-tailed]. The motion and money words correlation may be due to the politicians using motion words when metaphorically talking about the country's

progress rather than using it in physical sense with all politicians being equally concerned about money issues in progressing or setting back. The male politicians' less complex vocabulary and sentences pointed to a lower level of formality, which was not expected given the importance of the contexts. Furthermore, a positive correlation was also found between motion words and the pronouns *we* and *they* in the men's speeches [r_s (292) = .326, p = .000; r_s (292) = .178, p = .002, two-tailed], though not in the women's speeches [r_s (100) = .081, p = .417; r_s (100) = .014, p = .889, two-tailed]. It seems possible that motion words and the pronouns *we* and *they* correlation are due to the men specifically indicating what they (their party) did for the country progress, i.e. what the opposition failed to do, or did it to the country disadvantage.

4.2.6.2 Space

Numerous cognitive and psychological researchers (Delgado & Prieto, 1996; Dabs et al., 1998; Rilea et al., 2004; Parsons et al., 2004; Driscoll et al., 2005; Iachini et al., 2005) have found the gender difference in spatial abilities, with men outperforming women in navigational and orientation spatial tasks. Such research have inspired linguists to examine if spatial abilities are reflected in linguistic choices of men and women. Yale (2007), Ardila et al. (2011) and Isaac et al. (2011) confirmed that men used more spatial references than women thus indicating a relationship between cognitive abilities and linguistic expressions.

Even though the mean ranks from our analysis indicated that the men (M = 202.82) did use spatial words more than the women (M = 184.16), the Mann-Whitney results (U = 13531, Z = -1.422, p = .155) showed that the difference was not statistically significant. However, we found a gender difference regarding chamber by the post hoc Kruskal-Wallis test. Precisely, the women Representatives (M = 174.26) used spatial relations significantly less than the men Senators (M = 226.19, p = .025). We also calculated the percentage of specific words use (above, below, inside, outside, near, farther, etc.) but were unable to detect any significant differences. The only difference we found was in the measuring units of lengths (meters, kilometers, etc.) which were used more by the male politicians. This can be interpreted as the men being more precise in their speeches.

Further Spearman correlation tests revealed additional gender differences. Spatial words in the men's speeches negatively correlated with the pronouns *I*, *we*, *you* and *they* [r_s (292) = -.178, p = .002; r_s (292) = -.123, p = .036, r_s (292) = -.124, p = .034, two-tailed], whereas this was not the case in the women's speeches [r_s (100) = -.136, p = .260; r_s (100) = -.102, p = .310, r_s (100) = -.183, p = .066, r_s (100) = -.123, p = .218, two-tailed]. Additionally, there was a positive correlation of space words with long sentences and six-letter words in the men's speeches [r_s (292) = .244, p = .000; r_s (292) = .120, p = .041, two-tailed] and no correlation with the groups in the women's speeches [r_s (100) = .202, p = .061; r_s (100) = .095, p = .342, two-tailed], which further supports the idea that men are more formal and informative when using spatial words.

4.2.6.3 Time

While the concept of time has received attention in studies related to deception, there is little published data on the gender differences in the use of time-related words. One of the few studies that examined a distribution of temporal words among genders were Pennebaker et al.'s (2003), whose findings showed that men used more temporal words and are more precise and less emotional in their linguistic expression than women. These findings were challenged in a recent study by Iosub and colleagues (2014) who found that female editors used time-related words more than men. Moreover, they extended their research on time dimension and demonstrated that women were more concerned with past and present events in the article talk pages.

As illustrated in our Mann-Whitney results (U = 14904.5, Z = -.039, p = .969), the male (M = 197.87) and the female (M = 198.38) politicians used temporal references identically. Hence, we were unable to support any of the previous results. Further, we did not find any gender differences regarding the education level (p = .745), party affiliation (p = .765) or chamber (p = .103). However, we did record some correlations by using the Spearman correlation two-tailed tests. Firstly, temporal words in the men's speeches positively correlated with the past tense [r_s (292) = .233, p = .000]. In the women's speeches such correlation was not recorded [r_s (100) = .053, p = .564]. Temporal words and the past tense correlation, in addition to Iosub et al.'s (2014) findings inspired us to examine the

gender differences in the specific temporal words usage. Even though the percentage of the words use was very low, calculated with respect to a total number of words, the words *before* and *old/older/oldest* were more used by the men, while the words *recent/recently* and *lately* by the women, which is not in line with Iosub et al.'s (2014) findings. Hence, the male politicians were more focused on events and actions that are certain and unchangeable, while the female politicians focused on future actions, which might be a reflection of their higher task-orientation and not comments and acknowledgements of what had been done. Secondly, a high positive correlation with numbers was found in the men's [r_s (292) = .312, p = .000], but not in the women's [r_s (100) = .049, p = .623] speeches. Finally, in both the men's and the women's speeches, we found a negative correlation with tentativeness [r_s (292) = -.123, p = .035; r_s (100) = -.236, p = .017]. Even though both groups of the politicians expressed a high level of certainty when they used temporal references, a positive correlation with numbers may allude that the men are more formal and precise, which is consistent with our spatial words findings.

4.3 Current concerns

The current concerns category consists of words related to the subcategories of work, achievement, leisure, home, money, religion and death which represent different conversational topics. Gender differences in conversational topics have been extensively and systematically studied since 1922 when Henry Moore carried on a field observation study and found the gender differences in topic choices. His findings inspired numerous researchers to conduct similar studies in various subfields. Komarovsky's (1962), Klein's (1965), Harding's (1975), Reiter's (1975), Aries' (1976), Caldwell and Peplau's (1982), Haas and Sherman's (1982), Aries and Johnson's (1983), Johnson and Aries' (1983), Bishoping's (1993), Freed and Greenwood's (1996), Eggins and Slade's (1997) and Martin Rojo and Gomez Esteban's (2005) are just some of the studies that recorded the gender differences in conversational topics by using various approaches such as ethnographic descriptions, controlled setting group conversations and self-reports on topics. In comparison, the studies by Freed and Greenwood (1996) and Dolgin and Minowa (1997) recorded as strong, if not stronger, gender similarities in conversational contents. This fruitfulness evidently shows the importance of studying conversational topics among genders. Since the current concerns as a general category is not

analyzed by LIWC, we cannot provide statistical evidence of its use. However, we expect gender differences in the use of the subcategories.

4.3.1 Work

One of the most important years in examining gender differences across conversational topics was 1922, when Henry Moore theorized that gender differences in a topic choice were timeless since they were biologically oriented, i.e. they were manifestations of men and women's primal nature. Almost 70 years later, Bishoping (1993) decided to test his ideas by conducting a replication of Moore's study and providing a comparison of similar studies carried on from 1922 till 1990. Specifically, she compared Moore's (1922), Landis and Burtt's (1924), Landis' (1927), Sleeper's (1930), Stoke and West's (1930), Carlson et al.'s (1936), Watson et al.'s (1948), Meil's (1984), Kipers' (1987) and her results (1990) (all cited in Bishoping, 1993). Bishoping's (1993) study challenged Moore's ideas because she proved that conversational topics have changed over the years. Namely, work-related topics have decreased from 1922 till 1990 in men's and simultaneously dramatically risen in women's speeches, hence, a topic choice was proven not to be biologically determined. A comparative overview of other mentioned studies also disapproved Moore's ideas. Even though work topics have continually been dropping over the years in men's speeches, with the lowest recorded result in 1948, they again rose in 1987. In comparison, the lowest level of work topics in women's speeches was found in 1936, whereas the highest was in 1984. To paraphrase, the largest difference in women's speeches were in 1936 and 1948, which might have its basis in women started working outside their homes after the Second World War. However, despite the reduction of gender differences in work-related topics, men still prevailed in their usage, which was confirmed by Fehr (1996) who believed the reason for this was in work being a non-personal topic, hence the men's choice.

As can be seen from the mean ranks from our analysis, the male politicians (M = 191.83) talked about work less than the female (M = 215.71); however, the Mann-Whitney test (U = 13136.5, Z = -1.819, p = .069) showed that the difference was not statistically significant. Nevertheless, our results confirmed Bishoping's (1993) hypothesis of the increase of work-related topics in women's speeches. Furthermore, we conducted the post hoc

Kruskal-Wallis test to examine the gender differences in the party affiliation, education level and chamber. While there were no differences in the education level and party affiliation, we did find differences in chamber. Precisely, the men Representatives (M = 171.49) talked about work significantly less than the men Senators (M = 248.92, p = .000) and the women Senators (M = 278.10, p = .000). The intragroup difference was recorded among the female politicians, i.e. the women Senators (M = 278.10) used more work references that the women Representatives (M = 200.49, p = .039). Overall, the men Republican Senators (M = 254.13) and the women Democrat Senators (M = 296.94) talked about work the most, while the men Republican Representatives (M = 174.00) and the women Republican Representatives (M = 162.72) the least. Therefore, we may conclude that a serious issue of work was discussed more extensively in the Senate.

Furthermore, correlation gender differences were recorded with the two-tailed Spearman correlation tests. When speaking about work, the male politicians tended to use long sentences [r_s (292) = .418, p = .000] and more complex words [r_s (292) = .134, p = .022], which was not the case in the female politicians' speeches [r_s (100) = .347, p = .000; r_s (100) = -.055, p = .583]. Additionally, work-related words negatively correlated with pronouns [r_s (292) = -.168, p = .004] and social processes words [r_s (292) = -.282, p = .000] in the men's speeches in comparison to the women's where no such correlations were found [r_s (100) = -.112, p = .264; r_s (100) = .047, p = .636]. This could point to a higher level of formality and objectivity when discussing work topics in the men's speeches. Since no such correlations were found in the women's speeches, we could not draw a similar conclusion; yet, it would be incorrect to claim that the women were not formal and objective. Rather, there was no statistical evidence to support that.

4.3.2 Achievement

Even though there might be a natural connection between the concepts of work and achievement, researchers studied them separately so we will do the same. The first recognized researchers who examined the distribution of achievement words were Thorne and Henley (1975). They found that men preferred topics of work and achievement more than women. In 1991, in a study on children's beliefs and responses to failure and success in mathematics,

Stipek and Galinsky confirmed the findings that boys reported pride and achievement more than girls. However, this might have happened because boys outperform girls in mathematics in general. The traditional view that men use more achievement words was partially confirmed in Ireland's (2008) study. Comparing the American politicians' speaking styles, she found that McCain used achievement words the most (with 4% of his words related to the need for achievement), Biden moderately and Obama and Palin the least. Ireland claimed that McCain was the most ambitious and success oriented. Further, starting from a hypothesis that a recommendation letter written for females would contain less achievement and more communication skills references, Schmader et al. (2007) rejected it since they did not find statistically significant differences on the usage of achievement words. In a recent study, Adler (2013) proved that women were more likely to report pride of their achievements thus indicating possible changes in linguistic choices.

It is clear from our Mann-Whitney results ($U = 12733.5$, $Z = -2.225$, $p = .026$) that significant gender differences existed. The mean ranks showed that the women ($M = 219.66$) were achievement oriented more than the men ($M = 190.46$). Hence, our findings are consistent with Adler's (2013). Additionally, we recorded the gender differences with respect to the chamber. Namely, the men Representatives ($M = 169.30$) used achievement words significantly less than the men Senators ($M= 249.81$, $p = .000$) or the women Senators ($M = 268.8$, $p = .001$). To put it another way, the men Republican Senators ($M = 254.67$) and the women Republican Senators ($M = 270.75$) used achievement references the most, whereas the men Democrat Representatives ($M = 164.64$) and the women Republican Representatives ($M = 202.67$) the least. Consistent with our work-related vocabulary results, the Senators were more ambitious and success focused than the Representatives. The obvious change of women using more achievement references than men might be explained by the fact that, in comparison to the first research done in 1975, more women nowadays work. 20^{th} (especially in the first half) century women were mostly housewives who took care of children and did house chores. In spite of it being admirable and valuable job, neither women themselves nor society appreciate it enough. Moreover, rare were those who even consider it to be a job. These attitudes and monotonous routine days might have contributed to women not appreciating themselves and consequently not reporting any achievement they had made. However, with women's greater rights and employment, a situation has changed and is

reflected in their linguistic choices, as demonstrated by Schmader et al.'s (2007), Adler's (2013) and our study.

Additionally, we were interested in the gender differences in the usage of the specific achievement words. Since the percentages of those words were low calculated, with respect to the total number of words, we will only report the number of occurrences written in the brackets. We found that the male politicians used the words *beat* (58), *complete* (490), *control* (619), *win* (286) and *lose* (838). The female politicians, on the other hand, used the words *achieve* (90), *succeed* (176) and *improve* (163). Therefore, we may conclude that the male politicians were more competitively oriented and perceived their success in terms of defeating the other participatory party, while the female politicians perceived achievement as successful task completion which did not include anyone's failure or defeat.

Finally, the Spearman correlation tests confirmed our expectations and demonstrated gender similarities. When speaking about achievement, both the male [r_s (292) = .171, p = .003; r_s (292) = .447, p = .000] and the female [r_s (100) = .251, p = .011; r_s (100) = .352, p = .000] politicians used complex sentences and long words. The reported achievements were strictly related to work [r_s (292) = .584, p = .000 for the men; r_s (100) = .364, p = .000 for the women], i.e. neither the male nor the female politicians shared their personal achievements with their political colleagues. This is rather expected since the definition of achievement can be subjective. One's personal success need not be interpreted as such by someone who has not dealt with it. Business achievements, since they share more or less similar goals, are perceived differently and politicians can more easily identify themselves with the situation and success. Finally, a negative correlation with tentativeness [r_s (292) = -.319, p = .000 for the men; r_s (100) = -.380, p = .000 for the women] showed that both groups of the politicians were extremely certain when reporting their successful actions.

4.3.3 Leisure

Gender differences in speaking about leisure activities have been systematically studied since the first half of the 20th century; precisely, since Moore's (1922) study. In a comparative overview of eight studies carried on from 1922 until 1990, mentioned in

Subsection 4.3.1 and provided by Bishoping (1993), it was clear that men dominated in talking about leisure activities, especially sports. However, gradual rising of leisure references over the years in women's speeches was also evident. Yet, later research showed similar results. Eggins and Slade (1997) found that during coffee breaks at workplace, men tended to talk about sports, whereas women talked about personal experiences. Similarly, Martin Rojo and Gomez Esteban (2005) found the same results believing that men had problems when they talk about personal topics, i.e. they felt more relaxed talking about soccer. The same results that show men talk about sports or leisure activities in general were confirmed even in more recent studies by Yale (2007), Newman et al. (2008), Krenn and Schreitter (2015) and Manjavacas (2015).

In order to test the previous research results, we conducted the two-tailed Mann-Whitney test whose results ($U = 14181$, $Z = -.767$, $p = .443$) showed that the male ($M = 195.40$) and the female ($M = 205.47$) politicians identically used leisure references, which does not support any of the previous findings. No gender differences were recorded with respect to the chamber or party affiliation. Even though they were not statistically significant, we found that the men talked about ball sports (385) and video games (19) more than the women. Despite a tendency of equalization, ball sports are still more played by men and receive more media and fans attention than female ball sports. Also, men are more frequent video game players; hence, higher frequency of these words in the men's speeches was not surprising. Interestingly, the men also used more references to shopping (37) and mall (33) which are traditionally related to women.

Finally, the Spearman correlation tests revealed some gender differences in the use of leisure words. Leisure words positively correlated with achievement references in both the male [r_s (292) = .289, $p = .000$] and the female [r_s (100) = .212, $p = .033$] politicians' speeches, which might have happened when they were reporting and recognizing someone's sport results. However, the gender difference was in the men's expressing positive emotions [r_s (292) = .305, $p = .000$] while doing that, whereas the women did not [r_s (100) = .171, $p = .085$], which might mean that even when talking about casual topics such as leisure activities, the women were more formal and did not express their feelings.

4.3.4 Home

Recently, there has been some interest in examining gender differences in home references. The two studies which examined it, Newman et al.'s (2008) and Manjavacas' (2015), showed that home references were more used in women's speeches. The researchers attributed it to the women's natural mothers and caretakers' roles. Our study and the Mann-Whitney results (U = 12011, Z = -2.953 p = .003) confirmed the existence of the gender differences and the mean ranks showed that indeed the female politicians (M = 226.75) talked about home more than the male politicians (M = 187.99). Hence, our results are in accord with Newman et al.'s (2008) and Manjavacas' (2015). Further post hoc Kruskal-Wallis analysis showed that the women Representatives (M = 232.52) used home references significantly more than the men Senators (M = 159.73, p = .000). Also, the women Democrats (M = 244.08) used home references significantly more than the women Republicans (M = 163.70, p = .021), the men Democrats (M = 183.16, p = .002) and the men Republicans (M = 187.85, p = .003). Calculating the categorical variables of gender, party and house together, the men Republican Senators (M = 138.04) and the women Republican Senators (M = 156.50) referred to home the least, whereas the men Republican Representatives (M = 203.91) and the women Democrat Representatives (M = 251.43) the most. The women using more home references may be influenced by their social roles of mothers, wives and caretakers with home and their family playing a central role.

Furthermore, the two most used words from the home category by both the men and the women were *family* and *domestic* which was in the majority of cases followed by *violence*. The words pointed to the politicians being concerned about families in general specifically paying attention to the issue of violence. In addition, the Spearman correlation test revealed that the male politicians were concerned about families' health [r_s (292) = .151, p = .010] and expressed their sad feelings [r_s (292) = .118, p = .043], while such correlations were not found in the women's speeches [r_s (100) = .110, p = .270; r_s (100) = .009, p = .932], which again pointed to the women being more formal, i.e. they did not express any feelings even when speaking about home and families.

4.3.5 Money

Money and work have always been intertwined concepts so some researchers studied them as a unit rather than two separate parts. A comparative overview of eight studies from Bishoping's (1993) paper combined the issues of work and money. The results we reported in Subsection 4.3.1 apply for this subsection as well. Bishoping (1993) noticed a trend of decreasing the number of money references in men's and simultaneously increasing in women's speeches. Yet, men still used more money references than women. Since it was reported in 1993, the gender differences in the usage of money references have attracted a lot of interest. However, researchers found the same results. Regardless of examining different settings or written and spoken discourse, money was reported as a characteristic of the men's linguistic style (Lester, 2004; Schler et al., 2006; Yale, 2007; Ottoni et al., 2013; Singh Ludu, 2014; Cunha et al., 2014; Gorbatai & Nelson, 2015). The only subtle difference was found by Ireland (2008) who showed that McCain talked about money nearly three times more than another male politician Biden or Sarah Palin thus pointing to possible intragroup differences.

Conducting the Mann-Whitney two-tailed test, we found that the gender differences in money references were not significant ($U = 13170$, $Z = -1.785$, $p = .074$). However, the mean ranks pointed that the female politicians ($M = 215.38$) referred to money issues more than the male politicians ($M = 191.95$), which does not support the previous findings and point to gradual changes in money topics. Despite non-significant differences, we found that the words *tax* (2,495), *bargain* (587) and *bank* (152) were more used in the men's speeches and the words *debt* (149) and *insurance* (236) in the women's. These results may be interpreted as the female politicians being more focused on ensuring financial security, while the references to *bank*, which is known as a very powerful lobby, and *bargain* in the men's speeches pointed to negotiation and competitiveness – the concepts more associated with men.

We were also interested in correlations with other dependent variables so we conducted the two-tailed Spearman correlation tests. A positive correlation of money references and the concept of tentativeness showed that both the male [r_s (292) = .183, p = .002] and the female [r_s (100) = .257, p = .009] politicians were extremely cautious when they gave promises, suggestions or criticize previous actions because voters can forgive and forget various things but if you jeopardize their wellbeing by wasting money, the forgiveness will be

very difficult. Secondly, a positive correlation was found with present and future tenses in both the men's [r_s (292) = .320, p = .000; r_s (292) = .303, p = .000] and the women's speeches [r_s (100) = .287, p = .003; r_s (100) = .290, p = .003], which might mean they were comparing the current financial situation with possible future improvements, investments or savings. Lastly, the pronouns *we* [r_s (292) = .215, p = .000] and *they* [r_s (292) = .175, p = .003] positively correlated with money words in the men's while not in the women's [r_s (100) = .139, p = .165; r_s (100) = .092, p = .357] speeches. Taking the high number of *bank* and *bargain* references in the men's speeches into account, the pronouns *we* and *they* might stand for people, as users of loans and money in general, and banks as providers. Since the two parties have completely opposite interests, their money relationship has to be negotiated.

4.3.6 Religion

Since we were unable to find any previous research on gender differences in the use of religion references, our results might be beneficial to future researchers. The Mann-Whitney results (U = 12508, Z = -2.468, p = .014, two-tailed) demonstrated that the gender differences with a statistical significance existed. The mean ranks showed that religion references were more found in the men's (M = 206.31) than in the women's (M = 174.13) speeches. The gender differences were also found with respect to the party affiliation and chamber by the post hoc Kruskal-Wallis tests. The men Republicans (M = 210.88) referred to religion significantly more than the women Democrats (M = 166.32, p = .025) and the men Senators (M = 224.92) talked about religion more than the women Representatives (M = 165.90, p = .006). However, since the word *God* (418), accompanied by *Jesus* (38), was used at the highest frequency, the results might have been skewed by using the words in formulaic expressions such as *God bless America*, *Thank God*, *Thank Jesus*, etc. In addition to the catholic religion, which was mentioned the most in both the men's (1,002) and the women's (138) speeches, Islam was the second most mentioned religion with 208 references in the men's and 25 in the women's speeches. The majority of those references were related to the issue of jihad soldiers and ISIL. In the light of recent events, the religion of Islam being used almost exclusively with negative connotations, comes as no surprise.

4.3.7 Death

The last of the current concerns category, death, has received almost no research attention so far. To be more precise, only Graells-Garrido and colleagues (2015) dealt with the gender differences in the death reference use who, according to the results, claimed that men talked about death more than women. The Mann-Whitney test results from our analysis (U = 12814, Z = -2.172, p = .030, two-tailed) demonstrated that the male politicians (M = 205.27) talked about death significantly more than their female counterparts (M = 177.13). We confirmed recent Graells-Garrido et al.'s (2015) results. Based on these results, we were interested in possible gender differences in terms of the party affiliation and chamber seats examined by the post hoc Kruskal-Wallis tests. The significant gender difference was found in the party affiliation. Namely, the women Democrats (M = 166.81) talked about death significantly less than the men Democrats (M = 222.17, p = .003).

We further calculated the number of the use of specific death references and found that the most used word was *war* in both the men's (2,315) and the women's (396) speeches. Yet, it was used by the men at a much higher frequency. Further, while there were small or no differences in general death-related words such as decease, murder, overdose, kill, etc., the words *genocide* and *massacre* were more used in the men's (90, 52) than in the women's (7, 3) speeches. So, it was clear that the female politicians used more generalized death references, while the men, in addition to general ones, also talked about extremely violent crimes involving a lot of casualties.

The Spearman correlation two-tailed tests were also revealing. With a significant positive correlation of death references, the pronoun *he/she* and family, the male politicians [r_s (292) = .246, p = .000; r_s (292) = .165, p = .005] were more focused on victims and their family members, while the correlation was not found in the women's speeches [r_s (100) = .185, p = .062; r_s (100) = .023, p = .818], which indicated that the women focused more on the problem and not people. Both the men [r_s (292) = .280, p = .000; r_s (292) = .116, p = .047] and the women [r_s (100) = .301, p = .002; r_s (100) = .201, p = .040] talked about real events using past tense and reporting the number of victims in them. However, while the male politicians did not hesitate in expressing negative emotions, anger and sadness [r_s (292) = .377, p = .000; r_s (292) = .526, p = .000; r_s (292) = .229, p = .000], the female politicians only expressed

anger because of death victims [r_s (100) = .175, p = .079; r_s (100) = .361, p = .000; r_s (100) = .106, p = .288]. Again, the female politicians were proven to be reluctant to express their emotions even when speaking about the issue which usually unites everyone regardless of any differences.

4.4 Spoken categories

Spoken categories processed by LIWC include three paralinguistic subcategories – assents, nonfluencies and filler words, which will be elaborated on in the following subsections.

4.4.1 Assents

The assent category includes words such as *yeah*, *ok*, *alright*, *agree*, etc., i.e. words that signal a listener's improvement of a speaker's content of speech. Studies from the last decade have shown interest in examining gender differences in the use of assent words. In a study on bloggers' written language, Schler et al. (2005) found that female bloggers used more assents than their male counterparts. Severance (2012) extended the research and, in addition to the gender differences, examined the linguistic behavior in the same and mixed dyads. She found that women used the most assent words in male-dominated and the least in female-dominated groups, which made her conclude that women felt the strongest pressure to agree with men in male-dominated groups. The same findings of female dominance in the use of assent words were reported in the recent studies by Goedert et al. (2013) and Bamman et al. (2014) who believed that low power and status were being compensated by verbalizing their expressions of affirmation.

Since we had mixed-dyads in our study, we could not test, confirm or reject Severance (2012) claims. Therefore, we were just interested in potential gender differences in a formal mixed-group setting. Despite the mean ranks were pointing that the men (M = 200.86) used more assents than the women (M = 189.79), the Mann-Whitney results (U = 14105.5, Z = -.902, p = .367) showed that the difference was not statistically significant. This result may be

interpreted by the nature of the speeches. We believe that assent words are a characteristic of a natural turn-taking conversation and given the nature of our speeches (the politicians asking for a permission to address Congress and speaking on previously prepared or arranged topic without listeners answering to it), assent words were not expected to be used at a high rate. Therefore, in our opinion, examining assent words in speeches like ours which do not contain the element of turn-taking is not of relevance.

4.4.2 Nonfluencies

Similar to assent words, nonfluencies (*um*, *uh*, *er*, *hm*, etc.) have recently started to receive some research attention. The results reported by researchers were contradictory. While Yale (2007) and Fast and Funder (2008) found that women used more nonfluencies trying to get the attention of the conversation partner, McFadyen (1996), Freed and Greenwood (1996) and Mulac (1996) reported gender parity in the use of the category in question. Regardless of our results, we believe that examining nonfluencies on official congressional transcripts may not be very revealing. To put it differently, there is a possibility that transcripts were edited and nonfluencies removed which might highly skew the results. However, the scientific curiosity inspired us to examine the possible gender differences in nonfluencies because if the speech transcripts had been edited, they had probably been edited by the same person or a team editing them systematically, i.e. the editing process might have skewed the results in the overall frequency of nonfluencies and not the gender differences in the usage.

In order to test potential gender differences in the use of nonfluencies, we conducted the Mann-Whitney test. The results ($U = 12106.5$, $Z = -2.912$, $p = .004$, two-tailed) showed that the gender difference existed and that it was statistically significant. The mean ranks revealed that the male politicians ($M = 207.68$) used more nonfluencies than the female politicians ($M = 170.19$) which pointed to the women being more prepared for their speeches consequently using less nonfluencies. Hence, we could not support any of the previous research results. Additionally, the post hoc Kruskal-Wallis tests revealed more gender differences. Namely, the women Democrats ($M = 162.38$) used nonfluencies significantly less than the men Democrats ($M = 211.53$, $p = .012$) and the men Republicans ($M = 204.48$, $p = .036$). There were also relevant findings with the education level. In the same sex group, the

women holding a PhD degree (M = 131.06) used nonfluencies the least, while the women with a secondary education degree (M = 395.01) the most, which was expected because a higher degree might make a speaker more self-confident while speaking. Also, speakers with university degrees may be more exposed to public speaking and defending their views, which contributes to them being more self-asserting. In comparison, while the men with a secondary education degree (M = 161.58) used nonfluencies the least, the graduate degree men (M = 210.62) used them the most, which might be explained by the men with the lowest degree feeling more need to prepare for speeches. Lastly, we found a positive correlation of nonfluencies with one current concerns category by the Spearman correlation test. When speaking about death, the women used a lot of nonfluencies [r_s (100) = .257, p = .009], whereas the men did not [r_s (292) = .063, p = .282]. Since, as it was proved in our analysis, the female politicians rarely expressed their feelings towards anything, using nonfluencies when speaking about the issue of death might be interpreted as the women implicitly expressing their feelings.

4.4.3 Fillers

Since Lakoff's study in 1975 and serious beginnings of gender studies, women's language has been labelled as the weaker one mostly because of their usage of hedges and fillers. Some studies (Fast & Funder, 2008; Iosub et al., 2014; Manjavacas, 2015) demonstrated that women still used more fillers than men claiming that women were more insecure and used linguistic expressions more oriented towards informal discourse. Other researchers (Mulac & Lundell, 1986; Mulac et al., 1988; Christenfeld, 1995; Mehl et al., 2006) found the opposite result – men used fillers at a higher frequency than women. Interestingly, however, in the latter studies, men were not described as insecure or informal in their speeches. Rather, men's fillers were interpreted as a communicational strategy of holding a turn, i.e. preventing others from speaking, thus showing their dominance. Finally, several studies (McFadyen, 1996; Freed & Greenwood, 1996; Mulac, 1998; Hancock & Rubin, 2014) did not find any significant gender differences in the use of fillers.

The Mann-Whitney results from our analysis (U = 12954, Z = -2.024, p = .043, two-tailed) proved the existence of significant gender differences with the women (M = 217.50)

using more fillers than the men (M = 191.21). However, it should be noted that both the male and the female politicians used fillers at very low rates, which was probably due to previously prepared speeches some of which were surely completely read. Those female politicians who had not read their speeches might have used more fillers which need not have signified their insecurity; rather it might have been a way of keeping the talk flowing. Furthermore, using the Kruskal-Wallis test, we found that the women Representatives (M = 230.79) used fillers significantly more than the men Senators (M = 168.49, p = .003). Generally, the women Republican Senators (M = 136.12) and the men Republican Senators (M = 151.31) used fillers the least, whereas the women Democrat Representatives (M = 237.70) and the men Republican Representatives (M = 201.02) the most, which pointed to the Senators being more self-confident or prepared than the Representatives.

Interestingly, unlike with the previous categories, filler words correlated with only a couple of other variables, which was probably due to their very low usage. The Spearman correlation tests showed that when they were uncertain and used filler words to keep the speech flowing, the men's number of complex words decreased [r_s (292) = -.120, p = .039]; yet, this correlation did not apply to the women's speeches [r_s (100) = .109, p = .278]. Secondly, given the positive correlation with the present tense in the men's speeches [r_s (292) = .124, p = .034], it seemed likely that the men were more uncertain when they spoke about current actions and events. Again, this was not the case in the women's speeches [r_s (100) = .065, p = .516].

4.5 Pronunciation category

The pronunciation category is composed of a period, comma, colon, semicolon, question mark, exclamation mark, dash, quote, apostrophe, parenthesis, other punctuation marks and all punctuation marks. We decided to analyze the gender distribution of question and exclamation marks because they are realized in one's speech. Also, while we were preparing our corpus for the computational analysis, we noticed a tendency of the politicians quoting other people's speeches, so we chose to include quotation marks as well. We believe that other punctuation marks are not necessarily a reflection of one's speech. To paraphrase, since the politicians' speeches were transcribed, the punctuation marks we did not include in

our analysis might have been a transcriber's personal stylistic choice rather than a reflection of a pause in a politician's speech.

4.5.1 Question mark

Since Robin Lakoff's 1975 pioneering work in which she claimed that women used questions, especially question tags, more than men, the number of similar studies has rapidly increased. However, most of those studies directed their attention to question tags. Since one of our research questions was whether men and women used questions in general differently, we did not study question tags as a separate subcategory; hence, we will not report on question tags examined in previous studies. One of the earlier research on gender distribution in the use of questions was Fishman's (1980) who taped daily conversations of three young American couples. Based on fifty-two hours of tape recordings, Fishman counted the number of questions and found that the majority of them were asked by women. She concluded that women asked questions to keep the conversation going. Two years later, Maltz and Borker supported Fishman's claims that women viewed questions as conversation maintenance, while men used them to request information. In spite of finding the same quantitative results, some researchers (Holmes, 1988; Coates, 1993) offered an alternative interpretation – by raising questions, women's language was labelled as tentative. Later research (James & Clerke, 1993; Suborn, 2013) disagreed with equalizing asking questions and tentativeness and supported the previous approach by claiming that when asking questions, women encouraged others to speak thus aiming for rapport-building. To compare, Newman and colleagues (2008) did not find any gender differences in asking questions.

In our research, we used the Mann-Whitney test to examine the gender differences in asking questions. The results ($U = 12314$, $Z = -2.707$, $p = .007$) pointed to a statistically significant gender difference. However, contrary to all the previous research, the mean ranks showed that the men ($M = 206.97$) asked more questions than the women ($M = 172.23$). Also, the post hoc Kruskal-Wallis test showed that the men Representatives ($M = 209.60$) asked questions significantly more than the women Representatives ($M = 168.90$, $p = .030$). We did not want to make any conclusions without studying the types of raised questions. Firstly, in spite of not studying question tags, we were curious and briefly scanned the corpus to find

them. Interestingly, in neither the men's nor women's speeches question tags were found at high rates. However, since we did not study them in detail, we will not make conclusions about their usage. Secondly, our contextual analysis revealed that both the men and the women asked rhetorical questions with two main purposes – to ask for future actions or criticize for past actions. Given the nature of the speeches, which did not require an immediate direct answer, rhetorical questions were an understandable choice. There were no gender differences in the purposes choice. To put it differently, the only gender difference was that the male politicians asked more questions thus criticizing and urging their colleagues to actively plan future actions.

In addition, we conducted several Spearman correlation two-tailed tests, which further supports our findings of the purposes of questions. Firstly, when raising questions, both the male [r_s (292) = .415, p = .000] and the female politicians [r_s (100) = .209, p = .035] were tentative, yet, according to the p values, the men were more tentative than the women. Tentativeness in those cases may have been connected with the second correlation examples we found. Precisely, either when they were concerned about future or when they criticized previous actions, the men expressed negative emotions [r_s (292) = .173, p = .003], anger [r_s (292) = .155, p = .008] and swore [r_s (292) = .171, p = .003]. Once again, the female politicians showed their emotional neutrality by expressing neither negative emotions [r_s (100) = .045, p = .656] nor anger [r_s (100) = .117, p = .242]. However, they swore [r_s (100) = .234, p = .018], which was possibly their way of expressing emotions. Further, criticisms and questions asking for future actions positively correlated with the issue of money in both the men's [r_s (292) = .235, p = .000] and the women's speeches [r_s (100) = .205, p = .039]. Lastly, the women's questions also correlated with the topic of death [r_s (100) = .210, p = .034], whereas the men's did not [r_s (292) = .037, p = .529], which indicated that the women were concerned about lethal consequences and wanted to proactively work on the problem.

4.5.2 Exclamation mark

The usage of exclamation marks in language has been described as an indicator of emotive force (Quirk et al., 1985) or strong assertion (McArthur, 1992). Hence, using these markers of excitability implies emotional instability which is usually associated with women.

Numerous researchers studied the usage of exclamation marks in different discourses. One of the first was Hiatt's (1977) who found that women used more exclamations in prose because they were more emotional than men, which was confirmed by Scates (1981) in her doctoral dissertation, and also by Winn and Rubin (2001), in whose study women used exclamation marks three times more than men. Based on the studies conducted by Rafaeli and Sudweeks (1993), Savicki et al. (1996), Colley and Todd (2002) found that female college students used excitability exclamation marks, especially multiple ones, far more than male students. The same results were found by Rubin and Green (1992), who offered an alternative interpretation that exclamation marks were used as intensifiers (*"I really mean this!"*) thus pointing to assertiveness of a speaker/writer. Waseleski (2006) criticized all the previous studies because they were based on a mere calculation, i.e. contexts were completely ignored. Therefore, she conducted a content analysis of 200 exclamation marks and found that the most exclamations were statements of fact followed by expressing thanks and greetings and finally friendliness, cordiality and helpfulness messages. Nevertheless, women dominated in all the contexts. Regarding excitability, she found only 19 examples 10 of which were used by women and 9 by men, which led to a conclusion that exclamation marks were not necessarily indicators of emotionality. More recent research on internet linguistics (Crystal, 2011; Webb & Lee, 2012) continued reporting women as more emotional and expressive not examining the context as Waseleski (2006) suggested.

In a corpus of official transcripts, transcription of exclamation marks was based on a transcriber's stylistic choice so the data might have been skewed. Nevertheless, if skewed, it worked for both the men and women equally. To examine the gender differences in exclamations, we conducted the Mann-Whitney test ($U = 13906.5$, $Z = -2.346$), $p = .019$) which clearly demonstrated the existence of the gender difference. Yet, contrary to all the previous research, the male politicians ($M = 201.54$) used significantly more exclamations than the female (187.84) politicians. More specifically, the men Senators ($M = 240.10$) used exclamations significantly more than the men Representatives ($M = 8.78$, $p = .000$), the women Senators ($M = 194.05$, $p = .002$) and the women Representatives ($M = 86.32$, $p = .000$) provided by the post hoc Kruskal-Wallis test.

Inspired by Waseleski's (2006) study, we did not want to make any conclusions based solely on quantitative data, so we closely examined the contexts. There were only 5

occurrences in the women's and 64 in the men's speeches so the corpus might have been too limited to make gender differences conclusions. Nevertheless, exclamation marks were used as parts of personal names (Purple Up!), in quotations and when expressing the feelings of anger and dissatisfaction with past and current actions by both the men and the women. Additionally, the men used exclamations when they were cheering up (usually sports teams) and in the formulaic expressions such as *God bless America*. Hence, we could not support previous claims that women were more emotional and expressive than men since it was proved otherwise in our study. Furthermore, our contextual analysis was supported by the Spearman correlation tests. Namely, exclamation marks from the men's speeches positively correlated with the topics of work [r_s (292) = .139, p = .017], religion [r_s (292) = .130, p = .026] and leisure [r_s (292) = .122, p = .037] compared to the women's speeches [r_s (100) = .052, p = .601; r_s (100) = .085, p = .395; r_s (100) = .118, p = .238] where no such correlations, probably due to a very limited number of occurrences, were not found.

4.5.3 Quotation marks

During the preparatory and data cleansing process, we noticed a tendency of politicians quoting different sources so we decided to examine possible gender differences in it. Since previous sociolinguistic research did not examine the gender distribution in the use of quotation marks, our findings may be valuable to future researchers should they find it inspirational and thought-provoking. Even though the mean ranks were pointing that the male politicians (M = 201.22) quoted sources more than their female counterparts (M = 188.76), the Mann-Whitney test results (U = 14000.5, Z = -1.08, p = .285, two-tailed) showed that the difference was not statistically significant. However, the post hoc Kruskal-Wallis test showed one intragroup difference. Namely, the women Republicans (M = 244.14) quoted sources significantly more than the women Democrats (M = 173.53, p = .023). Therefore, we may conclude that both the male and the female politicians supported their claims equally from a quantitative perspective.

In addition to a statistical analysis, we examined the contexts of 856 occurrences in the men's and 138 in the women's speeches aiming to detect possible types of quotations. We were able to select five different types of quotes used by both groups of the politicians. The

most frequently quoted people were other politicians, both domestic and foreign, even though domestic were quoted more frequently. Usually, those were the politicians who performed important state roles or fellow politicians who served in Congress. Occasionally, the quoted people were late famous American politicians. The second most used type of quotes were letters, e-mails, and telephone conversations of people from states a certain politician represented. The texts of those quotations ranged from supports to personal problems. The third and fourth types were citations from the respected American newspapers and research journals. Here, we noticed a gender difference with the men using more of these two types. Finally, the fifth type, which was more used by the women, were religious quotes from different (mainly Biblical) religious sources or religious representatives' sermons, blogs and letters. These findings were supported by the Spearman correlations tests which showed that quotation marks from the men's speeches positively correlated with money-related words [r_s (292) = .115, p = .049] which was not the case in the women's speeches [r_s (100) = .050, p = .617]. To put it differently, the usage of quotation marks from the women's speeches correlated with the topic of religion [r_s (100) = .240, p = .015] unlike in the cases of the men's speeches [r_s (292) = .099, p = .089].

5. Conclusion

The conclusion section is divided into four subsections. Firstly, in Section 5.1, a list of conclusions is given. The list is followed by a review of the research questions and the key objectives of the study in Section 5.2. This subsection summarizes the main empirical findings from our study with respect to the individual research questions. Section 5.3 contains self-evaluation of our research by listing the strengths and limitations of the study. Finally, some recommendations for future research in the field are suggested in Section 5.4.

5.1 Conclusions

The study was set to explore the gender differences on the corpus of congressional speeches from the 113th United States Congress. The reason for choosing the 113th Congress

was the fact that it was composed of the record number of female political representatives. Precisely, it was composed of 450 male and 103 female participants. Another reason for conducting our study on the 113th Congress corpus was in its contemporariness, i.e. it is the most recent completed American Congress. Furthermore, the corpus was compiled with all uninterrupted speeches which contained at least 100 words. More specifically, it contained 2,983 speeches by the male and 672 by the female politicians. The official transcripts were analyzed with the text analysis software *Linguistic Inquiry and Word Count*, whose applicability to the study of the gender differences in congressional speeches was tested. Finally, the computational analysis results were processed with the software for statistical analysis *Statistical Package for the Social Sciences*, which was used to perform the independent sample t-test, Mann-Whitney U test, Kruskal-Wallis, one-way ANOVA and Spearman correlation test.

Upon conducting the pilot and the main research, we came to the following conclusions:

- The simple random sampling method is not appropriate for this type of research. Even though the method is well recognized and used in various types of research, we believe its disadvantages outweigh its advantages. Simple random sampling means that every member of the population has equal chances to be selected in the sample. However, this method cannot be applied if the members of the population are too heterogeneous in their nature. To paraphrase, if you want to get an unbiased sample, the list of members from the population should not be widely dispersed. When observing the population, we did not only take gender of the participants into account; rather, we took other sociodemographic factors (such as race, age, educational level and religion), party affiliation, chamber, topics, etc. into consideration. After having done that, we realized that our population was too heterogeneous to apply the simple random sampling method. Hence, in order to avoid skewness of the results, we decided to include all the speeches which had met the technical prerequisite of 100 words in our sample.

- The text analysis software *Linguistic Inquiry and Word Count* is a useful tool for the analysis of a large corpus. The software categorizes words and calculates the frequency of

their usage. The key advantage of LIWC is that it is easy to use and the pre-established and independently rated categories disable a researcher from being biased during the categorization process. Also, LIWC offers 80 categories ranging from word count and grammatical categories to different topics, spoken and punctuation categories, which gives a researcher a wide choice while doing research. However, LIWC cannot recognize irony, sarcasm, idioms and context. Hence, LIWC is an excellent tool for the fast and accurate computational categorization of words on a large corpus.

- LIWC output results in the form of raw numbers are useless if they are not analyzed by statistical methods. We decided to use SPSS, which has proven to be an excellent choice because it is compatible with LIWC. LIWC results are easily uploaded and do not require coding. The pull-down menu, effective data management and a wide choice of (non)parametric tests make SPSS an excellent supplemental tool.

- The combination of computational LIWC and statistical SPSS analysis is sufficient if the research aim is to examine which group of participants uses the selected LIWC categories more than other group(s) and if that difference is statistically significant. When interpreting quantitative analysis results, one might want to examine the contexts LIWC categories or specific words are used in. Since LIWC cannot recognize contexts, a researcher has to closely examine the corpus. The application of these two tools provides the information about the difference in the usage frequency and not the contextual usage itself.

- In addition to the gender differences, we recorded numerous gender similarities as far as the frequency of the LIWC categories is concerned, i.e. the male and the female politicians from our study used numerous LIWC categories at the same rate. However, a more detailed contextual analysis revealed the gender differences in the ways the categories were used. The gender differences will be elaborated on in the following subsection.

5.2. Review of the objectives and research questions

The main objective of this thesis was to examine if the linguistic practices of the male and female politicians who served in the 113th United States Congress differed. In order to pursue our objective, we used the text analysis software LIWC and the tool for the statistical analysis SPSS. We found that the tools, in addition to more critical analysis, could be successfully applied to the research on the gender differences in linguistic expressions. The specific research objectives have also been successfully met:

- We have found the differences in the linguistic practices of the male and the female politicians from the 113th United States Congress;
- The gender differences in the usage frequency were found for 26 variables which have proven to be statistically significant;
- We have determined which of the two groups of participants used the 26 variables significantly more;
- We have found the similarities in the linguistic practices of the male and the female politicians from the 113th United States Congress;
- The differences in the usage frequency in 44 variables were not statistically significant;
- By applying the Critical Discourse Analysis approach, we have found the gender differences in the ways some word categories were used even in those variables that did not record statistically significant difference in the usage frequency;
- We have managed, to a considerable extent, to provide underlying reasons for the differences in the variable usage.

The initial research questions have been answered as follows:

1) Is there a statistically significant difference in the usage of 70 linguistic categories tracked by LIWC between the male and female speakers on the corpus of the political speeches made in the 113th American Congress?

A statistically significant difference between the male and female speakers was recorded for 26 variables. The variables are as follows: *word count, six-letter words, function words, pronouns, pronoun you, impersonal pronouns, verbs, auxiliary verbs, past tense, adverbs, conjunctions, family, affective processes, negative emotions, sadness, causation, tentativeness, exclusive words, achievement, home, religion, death, nonfluencies, fillers, question mark* and *exclamation mark.*

To put it differently, a statistically significant difference was not recorded in the following variables: *words per sentence, personal pronouns, pronoun I, pronoun we, pronoun he/she, pronoun they, articles, present tense, future tense, prepositions, quantifiers, numbers, swear words, social processes, friends, humans, positive emotions, anxiety, anger, cognitive processes, insight, discrepancy, certainty, inhibition, inclusive words, perceptual processes, visually-related words, auditory-related words, tactile-related words, biological processes, body parts, health, sexual, ingestion, relativity, motion, space, time, work, leisure, money, assents* and *quotation mark.*

2) Which linguistic categories are predominately used by the male politicians?

Out of 26 variables where we recorded a statistically significant difference, 17 were predominately used by the male politicians. Those are: *word count, function words, pronouns, pronoun you, impersonal pronouns, verbs, auxiliary verbs, past tense, adverbs, conjunctions, tentativeness, exclusive words, religion, death, nonfluencies, question mark* and *exclamation mark.*

3) Which linguistic categories are predominantly used by the female politicians?

9 out of 26 statistically significant variables were predominately used by the female politicians. The variables are as follows: *six-letter words, family, affective processes, negative emotions, sadness, causation, achievement, home* and *fillers.*

4) How can the determined differences be interpreted?

The computational and more detailed corpus analysis of the LIWC variables have led to the following conclusions:

- The female politicians were more formal in their addressing Congress. They more frequently used the formulaic expressions for addressing Congress and were more elaborative clearly stating their motivation for the speech, while some male politicians started their speeches *in medias res*.

- The male politicians occupied the floor significantly more than the female politicians. Precisely, the total number of word count in the male speeches corpus was 2,198,364 words compared to 405,533 words in the female speeches corpus. The statistical evidence showed that the male politicians felt more comfortable speaking in the public setting. To put it in Deborah Tannen's terminology, a rapport, mostly used by women, is used for negotiating relationships and establishing connections, whereas, the purpose of a report, used by men, is to maintain independence and negotiate status. Since the purpose of the public setting of congressional speeches is more related to a report, which men are naturally better at, the fact that the male politicians used longer utterances could be interpreted as their attempt to establish themselves in a hierarchical order.

- The female politicians used six-letter words significantly more than the male politicians, which indicated that the women's vocabulary choice and a style were more informational and uninvolved. The reason why the female politicians made such scholarly choices of vocabulary might be to establish themselves as valuable participants in the political society.

- Words can be categorized in the category of content or function words. The latter category, which consists of pronouns, prepositions, articles, auxiliary verbs, conjunctions, negations and quantifiers, is used to organize and connect content words. The analysis of the category of function words and pronouns showed that the male politicians used them significantly more than the female politicians, which might point to the male politicians being

more involved and trying to build a relationship between themselves and the audience in comparison to more formal and less socially oriented female politicians.

- Even though there was no statistically significant difference, a contextual analysis of the pronoun *I* showed that the female politicians used the pronoun to separate themselves from the group/party they belonged to and establish themselves as independent politicians. Also, the male politicians shared private, while the female politicians focused on the public experience, which created the impression that the female politicians were more formal. Similarly, the analysis of the pronoun *we* indicated that the male politicians emphasized their affiliation more than the female politicians. While doing that, the male politicians reported their group's achievements, while the female politicians expressed the need for actions to be done by someone thus confirming their tendency of separating themselves from a group.

- In addition to the pronoun *you*, which was significantly more used by the male politicians, the gender differences were found in the types of the pronoun usage. Namely, the male politicians used the "intimate you", while the female politicians used the "critical you". These results were confirmed with the results for the pronoun *you + modal verb*, i.e. the female politicians expressed their stronger opinion than their male counterparts, thus asserting their authority.

- The male politicians used verbs and auxiliary verbs significantly more than the female politicians. The higher usage of verbs and auxiliary verbs, accompanied by pronouns, show that speakers pay attention to other people. These results suggested that the male politicians were socially and other-oriented, while the female politicians were task-focused. The gender differences were found in the usage of past tenses, which were used significantly more by the male politicians. That illustrated that they acknowledged past actions. It is in line with the results for the pronoun *we* (the male politicians recognizing their party's achievements).

- Even though there were no gender differences in mentioning their own or other people's family members, both the male and the female politicians referred to male family members more than to female, which indicated that men receive more public attention than

women. Furthermore, when they were speaking about family members, the politicians expressed their positive emotions and their speech became more tentative. It is possible that when they spoke about family members (a more personal topic), the politicians became more emotional and did not read prepared speeches, which resulted in a more tentative speech.

- When expressing positive feelings, the male politicians focused on the categories of family members, friends and their accomplishments, while the female politicians expressed their positive feelings in more categories. The female politicians expressed the majority of their positive feelings when they talked about their accomplishments, which pointed to higher formality and job orientation. In comparison, the female politicians expressed more negative feelings (sadness, anger and anxiety) than men. However, while doing that, the women focused on giving support to family members who experienced traumatic events, whereas the men focused on the consequences of tragic events, which is in accordance with Tannen's *advice versus understanding* dichotomy.

- The male politicians were more tentative and more emotional, which is reflected in fewer number references and more vulgar language. To compare, the female politicians were extremely assertive when they spoke about work and achievement topics. Additionally, the women who have PhDs were the least tentative, while the women with secondly education the most, which was not the case with the male politicians. The tentativeness results, accompanied by the nonfluency results, showed that the female politicians were more work-oriented and better prepared for speeches, which resulted in a higher degree of assertiveness.

- When speaking about health-related issues, the female politicians spoke about the issues such as eating disorder (anorexia and bulimia) and sexual offence (incest, rape and abortion), while the male politicians concentrated on the rehabilitation of war veterans issue, i.e. each gender dealt with the problems related to them. Tannen's *advice versus understanding* dichotomy was supported because the male politicians dealt with the consequences and possible solutions to the problems, whereas the female politicians focused on recognizing problems.

- The male politicians were more precise using more statistical data and measuring units in order to back up their speeches.

- The female politicians talked about theirs and other people's business achievements significantly more than the male politicians. When they talked about achievements, the male politicians used the verbs *beat, complete, control, win* and *lose*, while the female politicians used the verbs *achieve, succeed* and *improve*. The choice of verbs illustrated that the male politicians were more competitive and perceived achievement in the matter of defeating their opponent, while the female politicians focused on the task completion which did not imply someone's defeat. Similarly, when discussing money-related issues, the male politicians used the words *tax, bargain* and *bank*, while the female politicians used *debt* and *insurance*. This word choice pointed to the female politicians' desire to insure financial stability, whereas the male politicians focused on negotiation and competition.

- The male politicians used more references to ball sports and playing video games when they recognized someone's success. While doing that, the men expressed their positive feelings, as opposed to the women who kept the higher level of formality even when speaking about less serious topics.

- The male politicians used more religious references, especially the ones related to Islam and ISIL. They also spoke about different death consequences when they spoke about the number of victims. While doing that, the male politicians expressed their emotions in comparison to the female politicians who stayed unemotional and formal.

- The male politicians asked significantly more questions than their female colleagues. Both genders mainly asked rhetorical questions in order to criticize past actions and ask for future ones. Furthermore, the female politicians used significantly less exclamations marks, which means that they signaled fewer emotions.

The objectives of the research have been fulfilled and the research questions answered. Hence, the research has been successfully conducted. However, the research has a number of strengths and weaknesses, which will be elaborated on in the following subsection.

5.3. Strengths and limitations of the research

It is very difficult to be objective while evaluating one's own work; however, we will try to self-evaluate the research by listing its strong and weak points.

In our opinion, the strongest point of this research is the combination of the quantitative and the qualitative approach. The quantitative part of our research consisted of two subparts. Firstly, the corpus was studied by using the computational technique – the text analysis software *Linguistic Inquiry and Word Count* which categorized the words automatically. It safeguarded against bias in the categorization process. Secondly, the output of the computational analysis was examined with the statistical tool *Statistical Package for the Social Sciences*, which provided us with the useful tests for statistical significance of gender differences and correlations of different variables. Another strong point is the qualitative approach, i.e. in addition to detecting the gender differences, we provided the underlying reasons for their occurrence. Finally, by compiling a corpus using all the speeches which had met the technical prerequisite, we created a representative sample which allowed us to draw conclusions for the 113th United States Congress.

There are several weak points of our research. The text analysis software LIWC fails to recognize sarcasm, irony or contexts words are used in. Consequently, words can be miscategorized by the software. We tried to overcome this problem by more close examination of specific words and contexts they were used in. Secondly, the software can only recognize and categorize words which match the words in its internal dictionary. Even though the internal LIWC dictionary consists of 4,500 words and word stems, there is always a possibility that a word is not listed, which might skew the results. Our research results are limited in a way that they do not allow us to make generalizations about gender differences in language or congressional speeches since the research was done on the corpus of a two-year congressional speeches in the United States Congress. However, we got useful insights and a possible direction for future research.

5.4. Recommendations for future research

The present study suggests some new directions of research on gender differences in language. One way to proceed is to expand the corpus by adding more speeches from the previous, ongoing and future meetings of the legislative branch on the United States Congress. A more immediate way to proceed is to compile and analyze a corpus of speeches from the current 114th legislative branch meeting and compare research findings with ours in order to test our hypotheses of the current tendencies regarding the gender differences in language. Another way is to study a number of congressional meetings individually and compare them. Given the size of a potential corpus, this could be turned into a large-scale research project, which could benefit from contributions of more researchers.

During the analysis, several new research questions and hypotheses emerged:

1) In our corpus, the politicians who served in the House of Representatives (both the male and the female politicians) spoke less than the politicians who served in the Senate. They also used more function words, pronouns, verbs and auxiliary verbs whose usage suggested less formality. Are Representatives generally less formal than Senators? Hypothesis: Yes, Representatives are less formal and more other-oriented than Senators.

2) In different contexts (tributes, recognitions, politicians' personal examples, electorate's personal examples), men receive more public attention. Why does that happen? Why do female politicians recognize men more than women? Is there a difference in giving men more public attention by Representatives and Senators?

3) By the different usage of the personal pronouns, the female politicians tried to establish themselves as independent politicians, i.e. they avoided affiliating themselves to a group. Is the collective identity more avoided by the female politicians in the House of Representatives or the Senate? Hypothesis: the female politicians are trying to create an image and gain their independence more in the Senate than in the House of Representatives. The Senate represents a broader constituency than the House of Representatives, has more decision-making power and fewer members (two

from each state). Hence, the female politicians can easier develop a collective identity in the House of Representatives (even unintentionally) when jointly working with other female (or male) politicians.

4) The male politicians in our study were more precise. They used the definite article *the*, more statistics and measuring units. Is it universal? Is there a difference regarding the party affiliation or chamber?

5) In our analysis, the male politicians were more past-oriented and acknowledged their party's or other people's actions, while the female politicians called for future actions. Are the female politicians more reluctant to recognize someone's actions? If yes, why?

6) Hypothesis: when expressing emotions, politicians (regardless of their gender) become more tentative; they use shorter words and sentences and more vulgar language. It is possible to assume that when speaking about emotional topics, politicians do not read prepared speeches, which results in tentativeness and less formal speech.

7) Hypothesis: dealing with topics is gender related. Women talk about women's issues (eating disorders, incest, rape, abortion) and men address issues men are usually more concerned with, such as the treatment of war veterans.

8) In our study, the male politicians talked about different illness and war consequences (treatment costs and fatal injuries), while the female politicians tried to raise awareness and give their support for patients/victims and their families. Is it universal that women show compassion, while men try to solve problems? If yes, why?

9) Both genders presented their or someone else's achievements. Hypothesis: men are more competitive and perceive achievement in terms of beating an opponent, while women focus on the result which does not imply someone's defeat.

10) In our analysis, the male politicians talked about violent deaths (massacre, genocide), while the female politicians used general death references. Why does that happen?

The present research can be broadened by examining other categorical variables (chamber, party affiliation, education level, race, religion, etc.). The methodology can be applied to any national congressional meeting. It would be interesting to examine if our findings are culture related (by examining some other national congressional meetings) or they are universal. In addition to broadening the scope, the research topic can be made more specific – it can focus on several variables (such as pronouns, current topics, etc.) and study more correlations looking for their motivations and implications.

References

Abbar, S., Mejova, Y., Weber, I. (2015). You Tweet What You Eat: Studying Food Consumption Through Twitter. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3197-3206

Abbas, L. (2014). *Gender Differences in Expressing Wishes in Quotations from Dutch and English Sports News Articles: A Corpus-Based Study* (Master thesis). Universiteit Gent

Adams, K. L. (1992). Accruing power on debate floors. In Hall, K., Bucholtz, M., Moonwomon, B. (eds.) *Locating Power: Proceedings of the Second Berkeley Women and Language Conference.* Berkeley, CA: Berkeley Women and Language Group, 1-10

Adler, E. (2013). *Sex Differences in Positive Emotions within Appraisal Theory* (doctoral thesis). Vanderbilt University

Ahad, N.A., Yin, T.S., Othman, A.R., Yaacob, C.R. (2011). Sensitivity of Normality Tests to Non-normal Data. *Sains Malaysiana*, 40 (6), 637–641

Ahmad, I., Mehmood, A. (2015). Gender Differences in the Use of Parts of Speech in Pakistani English Newspaper Columns: A Critical Discourse Analysis. *New Media and Mass Communication*. 36, 10-15

Ahrens, K. (ed.) (2009). *Politics, Gender and Conceptual Metaphors*. New York: Palgrave Macmillian

Ainsworth-Vaughn, N. (1992). Topic transitions in physician-patient interviews: power, gender and discourse change. *Language and Society*, 21, 409-26

Alam, F., Riccardi, G. (2013). Comparative Study of Speaker Personality Traits Recognition in Conversational and Broadcast News Speech. *Proceedings of 14th Annual Conference of the International Speech Communication Association Interspeech*, 2851-2855

Alexander, D., Andersen, K. (1993). Gender as a Factor in the Attribution of Leadership Traits. *Political Research Quarterly*, 46, 527–45

Al-Faki, I.M. (2014). Political Speeches of Some African Leaders from Linguistic Perspective (1981-2013). *International Journal of Humanities and Social Science*, 4 (3), 180-198

Andersen, E. (1984). The acquisition of sociolinguistic knowledge: some evidence from children's verbal role-play. *Western Journal of Speech Communication*, 48, 125-144

Anderson, T.W., Darling, D.A. (1954). A Test of Goodness of Fit. *Journal of the American Statistical Association*, 49 (268), 765-769

Andersson, L. (2012). Personal Pronouns in Editor's Letters – A gender-based study. *Linnaeus University*, School of Language and Literature

Anshen, F. (1973). *Sex and obscenity at Stony Brook*. (Unpublished manuscript). State University of New York at Stony Brook

Ardila, A., Rosselli, M., Matute, E., Inozemtseva, O. (2011). Gender differences in cognitive development. *Developmental Psychology*, 47 (4), 984-90

Aries, E. (1976). Interaction patterns and themes of male, female, and mixed groups. *Small Group Behavior*, 7, 7-18

Aries, E. (1996). *Men and Women in Interaction: Reconsidering the Differences*. New York: Oxford University Press

Aries, J., Johnson, F. L. (1983). Close friendship in adulthood: Conversational content between same-sex friends. *Sex Roles*, 9, 1183-1196

Argamon, S., Koppel, M., Fine, J., Shimoni, A. (2003). Gender, genre, and writing style in formal written texts

Argamon, S., Koppel, M., Pennebaker, J.W., Schler, J. (2007). Mining the blogosphere: Age, gender and the varieties of self-expression. *First Monday*, 12 (9)

Argamon, S., Koppel, M., Pennebaker, J.W., Schler, J. (2009). Automatically profiling the author of an anonymous text. *Communications of the ACM*, 52 (2), 119–123

Arustamyan, R. (2014). The Manifestation of Gender Peculiarities in Political Discourse. *Armenian Folia Anglistika*, 1 (12), 88-95

Ashby Plant, E., Hyde, J.S., Keltner, D., Devine, P.G. (2000). The gender stereotyping of emotions. *Psychology of Women Quarterly*, 24 (1), 81-92

Austin, J.L. (1962). *How to do things with words*. London: Oxford University Press

Bailey, L.A., Timm, L.A. (1976). More on women's and men's expletives. *Anthropological Linguistics*, 18, 438-449

Baker, P. (2014). *Using corpora to analyse gender*. London: Bloomsbury

Bamman, D., Eisenstein, J., Schnoebelen, T. (2014). Gender in Twitter: Styles, Stances, and social networks. *Journal of Sociolinguistics*, 18, 135-160

Basow, S. A., Rubenfeld, K. (2003). "Troubles talk": Effects of gender and gender typing. *Sex Roles: A Journal of Research*, 48, 183-187

Baumann, A., Krasnova, H., Veltri, N., Yusni, Y. (2015). Men, Women, Microblogging: Where Do We Stand? *Proceedings of 12th Internationale Tagung Wirtschaftsinformatik*, 857-871

Baxter, J. (1999a). Teaching girls to speak out: the female voice in public contexts. *Language and Education*, 13 (2), 81-98

Baxter, J. (1999b). *Teaching Girls to Speak Out: An investigation of the extent to which gender is a pertinent discourse for describing and assessing girls' and boys' speech in public contexts.* Unpublished doctoral thesis. University of Reading

Baxter, J. (2003). *Positioning Gender in Discourse: A Feminist Methodology*. New York: Palgrave Macmillan

Baxter, J. (ed.) (2006). *Speaking Out: The Female Voice in Public Contexts*. New York: Palgrave Macmillan

Baxter, J. (2010). *The Language of Female Leadership*. New York: Palgrave Macmillan

Beattie, G. (1981). Interruption in Conversational Interaction and its Relation to the Sex and Status of the Interactants. *Linguistics*, 19, 15-35

Bell, C., McCarthy, P., McNamara, D. (2006). Variations in Language Use across Gender: Biological versus Sociological Theories. In Sun, R., Miyake, N. (eds.). *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, Mahwah New Jersey: Erlbaum, 1009-1013

Bell, C., McCarthy, P., McNamara, D. (2012). Using LIWC and Coh-Metrix to Investigate Gender Differences in Linguistic Styles. In McCarthy, P., Boonthum-Denecke, C. (eds.) *Applied Natural Language Processing: Identification, Investigation, and Resolution*. Hershey, PA: IGI Global, 545-556

Bengoechea, M. (2011). Media Interpretation of a Leading Woman Politician's Performances and Dress Code Challenges. *Journal of Interdisciplinary Feminist Thought*, 5 (1), Article 5

Berger, C. (2003). *The myth of gender-specific swearing*. Berlin: Verlag für Wissenschaft und Forschung

Bergvall, V. (1996). The Question of Questions: Beyond Binary Thinking. In Bergvall, V., Bing, J., Freed, A. (eds.) *Rethinking Language and* Gender *Research: Theory and Practice*. New York: Longman, 1-30

Berry, D.S., Hiller, W.S., Mueller, J.S., Pennebaker, J.W. (1997). Linguistic bases of social perception. *Personality and Social Psychology Bulletin*, 23 (5), 526-537

Best, D.L., Williams, J.E. (1982). *Measuring Sex Stereotypes: A Thirty- Nation Study*. Beverly Hills, CA: Sage

Biber, D., Conrad, S., Reppen, R. (1998). *Corpus Linguistics. Investigating language structure and use*. Cambridge: Cambridge University Press

Biesen, J.N., Schooler, D.E., Smith, D.A. (2015). What a Difference a Pronoun Makes I/We Versus You/Me and Worried Couples' Perceptions of Their Interaction Quality. *Journal of Language and Social Psychology*

Billig, M. (2003). Critical Discourse Analysis and the rhetoric of critique. In Weiss, G., Wodak, R. (eds.) *Critical Discourse Analysis: theory and interdisciplinarity*. Palgrave Macmillan, 35-46

Bischoping, K. (1993). Gender Differences in Conversation Topics, 1922-1990. *Sex Roles*, 28 (1/2), 1-18

Black, M., Coward, R. (1998). Linguistic, social and sexual relations: a review of Dale Spender's Man Made Language. In Cameron, D. (ed.) *The Feminist Critique of Language: a Reader*. Routledge London: Routledge, 100-118

Blaisdell, B. (ed.) (2011). *Great Speeches of the 20th Century*. Washington: Dover Thrift Editions

Bodine, A. (1975). Sex differentiation on language. In Thorne, B., and Henley, N. (eds.) *Language and sex: Difference and dominance*. Rowley MA: Newbury House Publishers, 130-151

Bohan, J. (1993). Regarding gender: essentialism, constructionism and feminist psychology. *Psychology of Women Quarterly*, 17, 5-22

Bradley, P.H. (1981). The folk-linguistics of women's speech: An empirical examination. *Communication Monographs*, 48, 73-90

Bramley, N.R. (2001). *Pronouns of politics: the use of pronouns in the construction of 'self' and 'other' in political interviews* (doctoral thesis). Australian National University

Brewer, M. B., Gardner, W. (1996). Who is this "we"? Levels of collective identity and self-representations. *Journal of Personality and Social Psychology*, 71, 83-93

Brónnimann, R., Herlihy, J., Müller, J., Ehlert, U. (2013). Do testimonies of traumatic events differ depending on the interviewer? *The European Journal of Psychology Applied to Legal Context*, 5 (1), 97-121

Brouwer, D., Gerritsen, M., De Haan, D. (1979). Speech Differences between Women and Men: On the Wrong Track? *Language in Society*, 8 (1), 33-50

Brown, W. (1988). *Manhood and Politics: A Feminist Reading in Political Theory.* Totowa, NJ: Rowman and Littlefield

Brown R., Gilman, A. (1960). The Pronouns of Power and Solidarity. In Sebeok, T. (ed.) *Style in Language*. Cambridge, Massachusetts: MIT Press, 253-276

Brownlow, S., Rosamond, J.A., Parker, J.A. (2003). Gender-Linked Linguistic Behavior in Television Interviews. *Sex Roles*, 49 (3), 121-132

Broughton, S. and Palmieri, S. (1999). Gendered contributions to parliamentary debates: The case of euthanasia. *Australian Political Studies* Association, 34 (1), 29-45

Bublitz, W. (1988). *Supportive fellow-speakers and cooperative conversations. Discourse topics and topical actions, participant roles and "recipient action" in a particular type of everyday conversation*. Amsterdam: John Benjamins

Burns, N., Lehman Schlozman, K., Verba, S. (eds.) (2001). The Private Roots of Public Action: Gender, Equality, and Political Participation. London: Harvard University Press

Butler, J. (1990). *Gender Trouble.* London: Routledge

Butler, J. (1993). *Bodies that Matter: On the Discursive Limits of "Sex"*. New York: Routledge

Bystrom, D.G., DeRosa, K.L. (1999). The Voice of and for Women in the 1996 Presidential Campaign: Style and Substance of Convention Speeches. In Bystrom, D.G. and Lee Kaid, L. (eds.) *The Electronic Election*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers, 97-111

Caldas-Coulthard, C.R. (1996). *Texts and practices: Readings in critical discourse analysis*. Psychology Press

Caldwell, M. A., Peplau, L. A. (1982). Sex differences in same-sex friendship. *Sex Roles*, 8, 721-732

Cambridge dictionary available at http://dictionary.cambridge.org

Cameron, D. (1995a). *Verbal Hygiene*. London: Routledge

Cameron, D. (1995b). Rethinking Language and Gender Studies: Feminism into the 1990s. In Mills, S. (ed.) *Language and Gender: Interdisciplinary Perspectives*. London: Longman, 31-44

Cameron, D. (1996). The Language-gender Interface: Resisting Co-optation. In Bergvall, V., Bing, J., Freed, A. (eds.) *Rethinking Language and Gender Research: Theory and Practice*. New York: Longman, 31-53

Cameron, D. (1997). Theoretical debates in feminist linguistics: Questions of sex and gender. In Wodak, R. (ed.) *Gender and Discourse*. London: Sage, 21-36

Cameron, D. (ed.) (2002). *The Feminist Critique of Language: A Reader.* London: Routledge Taylor and Francis Group

Cameron, D. (2005). Language, Gender and Sexuality: Current Issues and New directions. *Applied linguistics*, 26 (4), 682-502

Cameron, D. (2007). *The Myth of Mars and Venus*. London: Oxford University Press

Cameron, D., McAlindon, F., O'Leary, K. (1989). Lakoff in context: the social and linguistic functions of tag questions. In Coates, J. and Cameron, D. (eds.) *Women in their Speech Communities.* London: Longman, 74-93

Caplan, P., Crawford, M., Shibley Hyde, J., Richardson, J. (1997). *Gender Differences in Human Cognition*. Oxford: Oxford University Press

Carli, L. (1990). Gender, language, and influence. *Journal of Personality and Social Psychology*, 59, 941-951

Carrol, S.J., Dodson, D.L., Mandel, R.B. (1991). *The Impact of Women in Public Office: An Overview*. New Brunswick, NJ: Center for American Women in Politics

Carroll, D.W. (2007). Patterns of student writing in a critical thinking course: A quantitative analysis. *Assessing writing*, 12, 213–227

Carvalho, A.M., Child. M. (2011). Subject pronoun expression in a variety of Spanish in contact with Portuguese. In Michnowicz, J. and Dodsworth, R. (eds.). *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics. Somerville,* MA: Cascadilla Press, 14-25

Case, S.S. (1988). Cultural differences, not deficiencies: an analysis of managerial women's language. In Rose, S., Larwood, L. (ed.) *Women IS Careers: Pathways and Pitfalls.* New York: Praeger, 41-63

Cassell, J., Huffaker, D., Tversky, D., Ferriman, K. (2006). The language of online leadership: Gender and youth engagement on the Internet. *Developmental Psychology*, 42 (3), 436-449

Cavazza, N., Guidetti, M. (2014). Swearing in Political Discourse: Why Vulgarity Works. *Journal of Language and Social Psychology*, 33 (5), 537-547

Chafe, W. (1982). Integration and Involvement in Speaking, Writing and Oral Literature. In Tannen, D. (ed.) *Spoken and Written Language*. Norwood, NJ: Ablex, 35-53

Chafe, W. (1995). Polyphonic topic development. In Givón, T. (ed.) *Conversation: Cognitive, Communicative and Social Perspectives*. Typological Studies in Language 34. Amsterdam: John Benjamins, 41-55

Charteris-Black, J. (2005). *Politicians and Rhetoric: The Persuasive Power of Metaphor*. New York: Palgrave Macmillan

Chentsova-Dutton, Y., Tsai, J. (2007). Gender differences in emotional response among European Americans and Hmong Americans. *Cognition and Emotion*, 21 (1), 162-181

Childs, S. (2000). The new Labour women MPs in the 1997 British Parliament. *Women's History Review*, 9 (1), 55-73

Chilton, P. (2004). *Analysing Political Discourse: Theory and Practice*. London and New York: Routledge Taylor and Francis Group

Chraif, M., Anitei, M. (2013). Gender differences in measuring positive and negative emotions self-perception among Romanian high school students–A pilot study. *Procedia - Social and Behavioral Sciences*, 76 (15), 181–185

Christie, C. (2000). *Gender and Language: Towards a Feminist Pragmatics*. Edinburgh: Edinburgh University Press

Christenfeld, N. (1995). Does it hurt to say um? *Journal of Nonverbal Behavior*, 19, 171-186

Chung, C., Pennebaker, J. (2007). *The Psychological Functions of Function Words*. In Fiedler, K. (ed.) (2007). *Social Communication*. New York: Psychology Press, 343-359

Cicourel, A. (1974). *Cognitive Sociology: Language and Meaning in Social Interaction*. London: Penguin

Coates, J. (1989). Gossip revisited: language in all-female groups. In Coates, J., Cameron, D. (eds.) *Women in their Speech Communities.* London and New York: Longman, 75-94

Coates, J. (1993). *Women, men, and language*. London: Longman

Coates, J. (1994). The language of the professions: discourse and career. In Evetts, J. (ed) *Women and Career: Themes and Issues in Advanced Industrial Societies.* London and New York: Longman, 72-86

Coates, J. (1996). *Women Talk: Conversation Between Women Friends*. Oxford: Blackwell

Coates, J. (1997). The construction of a collaborative floor in women's friendly talk. In T. Givón, T. (ed.) *Conversation: Cognitive, Communicative and Social Perspectives*. Philadelphia: John Benjamins, 55-89

Coates, J., Jordan, M.E. (1997). Que(e)rying Friendship: Discourses of Resistance. In Livia, A. and Hall, K. (eds.) *Queerly Phrased: Language, Gender and Sexuality*. Oxford and New York: Oxford University Press, 214-232

Coates, J. (2003). *Women, Men and Language: A Sociolinguistic Account of Gender Differences in Language*. London: Longman

Cohen, A., Alpert, M., Nienow, T., Dinzeo, T., Docherty, N. (2008). Computerized measurement of negative symptoms in schizophrenia. *Journal of psychiatric research*, 42 (10), 827–836

Cohen, A.S., St-Hilaire, A., Aakre, J.M., Docherty, N.M. (2009). Understanding anhedonia in schizophrenia through lexical analysis of natural speech. *Cognition and emotion,* 23, 569–586

Colley, A., Todd, Z. (2002). Gender-linked differences in the style and content of e-mails to friends. *Journal of Language and Social Psychology*, 21 (4), 380–392

Colley, A., Todd, Z., Bland, M., Holmes, M., Khanom, N., Pike, H. (2004). Style and content in e-mails and letters to male and female friends. *Journal of Language and Social Psychology*, 23, 369–378

Costa, P. Jr., Terracciano, A., McCrae, R. (2001). Gender Differences in Personality Traits across Cultures: Robust and Surprising Findings. *Journal of Personality and Social Psychology*, 81, 322–31

Craig, A. (2008). Gender Stereotypes about Emotion Portrayed in Children's Picture Books. (Master thesis). *NC State University*

Crawford, M. (1995). *Talking Difference: On Gender and Language*. London: Sage

Crosby, F., Nyquist, L. (1977). The female register: an empirical study of Lakoff's hypotheses. *Language in Society*, 6 (3), 313-322

Cruse, A. (2006). *A Glossary of Semantics and Pragmatics*. Edinburgh: Edinburgh University Press

Crystal, D. (2011). *Internet linguistics – A Student Guide*. London: Routledge

Cuming, J. (2013). Gender differences in the recall of neutral, positive and negative emotive words. *Institutional Repository and Scholarly Communication*: University of Johannesburg

Cunha, E., Magno, G., Marcos André Gonçalves, M.A., Cambraia, C., Almeida, V. (2013). A Linguistic Characterization of Google+ Posts across Different Social Groups. *5th Workshop on Information in Networks*

Dabs, J. M., Chang, E. L., Strong, R. A., Milun, R. (1998). Spatial ability, navigation strategy, and geographic knowledge among men and women. *Evolution and Human Behavior*, 19, 89–98

D'Agostino, R.B., Stephens, M.A. (1986). *Goodness-of-Fit Techniques*. New York: Marcel Dekker, Inc.

Danescu-Niculescu-Mizil, C., Gamon, M., Dumais S. (2011). Mark my words!: linguistic style accommodation in social media. *Proceedings of the 20th international conference on World Wide Web*. ACM: New York, 745-754

Davies, B., Harre, R. (1990). Positioning: The Discursive Production of Selves. *Journal for the Theory of Social Behaviour*, 20 (1), 43-63

Davis, D., Brock, T.C. (1975). Use of First Person Pronouns as a Function of Increased Objective Self-Awareness and Performance Feedback. *Journal of Experimental Social Psychology*, 11, 381-388

DeFrancisco, V.L. (1989). *Marital communication: A feminist qualitative analysis* (doctoral dissertation). Champaign/Urbana: University of Illinois

DeFrancisco, V.L. (1991). The Sounds of Silence: How Men Silence Women in Marital Relations. *Discourse and Society*, 2 (4), 413-423

De Klerk, V. (1991). Expletives: Men only? *Communication Monographs*, 58, 156–169

Delgado, A. R., Prieto, G. (1996). Sex differences in visuospatial ability. Do performance factors play such an important role? *Memory and Cognition*, 24, 504–510

Delucchi, K.L., Bostrom, A. (2004). Methods for analysis of skewed data distributions in psychiatric clinical studies: working with many zero values. *American Journal of Psychiatry*, 161 (7), 1159-68

DeWall, N., Buffardi, L.E., Bonser, I., Campbell, W.K. (2011). Narcissism and implicit attention seeking: Evidence from linguistic analyses of social networking and online presentation. *Personality and Individual Differences*, 51, 57-62

Dittmar, N. (1976). *Sociolinguistics: A critical survey of theory and application*. London: Arnold

Dodson, D.L. (2006). *The Impact of Women in Congress*. Oxford University Press

Dolan, K. (2008). Women as Candidates in American Politics - The Continuing Impact of Sex and Gender. In Wolbrecht, C., Beckwith, K., Baldez, L. (eds.) *Political Women and American Democracy*. Cambridge: Cambridge University Press, 110-127

Dolgin, K. G., Minowa, N. (1997). Gender differences in self-presentation: A comparison of the roles of flatteringness and intimacy in self-disclosure to friends. *Sex Roles*, 36 (5/6), 371-380

Dovi, S. (2008). Theorizing Women's Representation in the United States. In Wolbrecht, C., Beckwith, K., Baldez, L. (eds.) *Political Women and American Democracy*. Cambridge: Cambridge University Press, 148-166

Driscoll, I., Hamilton, D.A., Yeo, R.A., Brooks, W.M., Sutherland, R.J. (2005). Virtual navigation in humans: the impact of age, sex, and hormones on place learning. *Hormones and Behavior*, 47, 326-35

Duan, A., Dowell, N.M., Graesser, A.C., Li, H. (2014). Linguistic Style and Social Historical Context: An Automated Linguistic Analysis of Mao Zedong's Speeches. *Proceedings of the Twenty-Seventh International Florida Artificial Intelligence Research Society Conference*. AAAI Publications, 43-46

Dubois, B.L., Crouch, I. (1975). The question of question tags in women's speech: They don't really make more use of them, do they?, *Language and Society,* 4, 289-94

Durant, A. (1999). Clarifying notions of 'meaning' at stake in linguistic evidence as to meaning: libel as illustration. *International Association of Forensic Linguists 4th Biennial Conference*, Birmingham, UK

Eagly, A. H., Johannesen Schmidt, M. C., van Engen, M. (2003). Transformational, Transactional, and Laissez-Faire Leadership Styles: A Meta-Analysis Comparing Women and Men. *Psychological Bulletin*, 95, 569-591

Eagly, A., Diekman, A., Johannesen-Schmidt, M., Koenig, A. (2004). Gender Gaps in Sociopolitical Attitudes: A Social Psychological Analysis. *Journal of Personality and Social Psychology*, 87, 796–816

Edelsky, C., Adams, K.L. (1990). Creating equality: Breaking the rules in debates. *Journal of Language and Social Psychology*, 9 (3), 171-90

Edwards, J. (ed.) (2009). *Gender and Political Communication in America: Rhetoric, Representation and Display*. Maryland: Lexington Books

Eckert, P. (1989). The whole woman: sex and gender differences in variation. *Language Variation and Change*, 3, 245-267

Eckert, P. (1998). Gender and sociolinguistic variation. In Coates, J. (ed.) *Language and Gender: A Reader*. Oxford: Blackwell, 64-75

Eckert, P., McConnell-Ginet, S. (1992). Communities of Practice: Where language, gender and power all live. In Hall, K., Bucholtz, M., Moonwomon, B. (eds.) *Locating Power*. Berkley: Berkley Women and Language Conference, 89-99

Eckert, P., McConnell-Ginet, S. (1995). Constructing Meaning, Constructing Selves: Snapshots of Language, Gender and Class from Belten High. In Bucholtz, M., Hall, K. (eds.) *Gender Articulated: Language and the Culturally Constructed Self*. New York: Routledge

Eckert, P., McConnel-Ginet, S. (2003). *Language and Gender*. Cambridge: Cambridge University Press

Edelsky, C. (1981). Who's Got the Floor?. *Language in Society*, 10 (3), 383-421

Edwards, D., Potter, J. (1992). *Discursive psychology*. London: Sage

Eggins, S., Slade, D. (1997). *Analysing Casual Conversation*. London: Cassell

Eisenman, R. (1997). *Men, women and gender differences: the attitudes of three feminists – Gloria Steinem, Gloria Allred and Bella Azbug*

Eliason, C. (2007). Language and gender as reflected in the advertisements of wedding magazines (thesis). *Kalmar University*, 41-60

Epstein, E.M., Sloan, D.M., Marx, B.P. (2005). Getting to the Heart of the Matter: Written Disclosure, Gender, and Heart Rate. *Psychosomatic Medicine*, 67, 413–419

Erickson, B., Lind, E.A., Johnson, B.C., O'Barr, W.M. (1978). Speech style and impression formation in a court setting: The effects of "powerful" and "powerless" speech. *Journal of Experimental Social Psychology*, 14, 266-279

Fairclough, N. (1989). *Language and Power*. London: Longman

Fairclough, N. L., Wodak, R. (1997). Critical discourse analysis. In van Dijk, T.A. (ed.) *Discourse Studies. A Multidisciplinary Introduction, Vol 2. Discourse as Social Interaction*. London: Sage, 258-284

Falk, J. (1980). The conversational duet. *Proceedings of the 6th Annual Meeting of the Berkeley Linguistics Society*, 6, 507–514

Farrel, P.J., Stewart, K.R. (2006). Comprehensive Study Of Tests For Normality And Symmetry: Extending The Spiegelhalter Test. *Journal of Statistical Computation and Simulation*, 76 (9), 803–816

Fast, L.A., Funder, D.C. (2008). Personality as Manifest in Word Use: Correlations With Self-Report, Acquaintance Report, and Behavior. *Journal of Personality and Social Psychology*, 94 (2), 334-346

Fehlman, B. (2015). Gender and Dialogue Differences in American Novels. *English Class Publications*, Paper 6

Fernández-Cabana, M., Jiménez-Félizb, J., Alves-Pérezc, M.T., Mateosa, R., Gómez-Reino Rodríguezd, I., García-Caballeroa, A. (2015). *The European Journal of Psychiatry*, 29 (2), 145-155

Fisher, H. (1995). The nature and evolution of romantic love. In Jankowiak, W. (ed.) *Romantic passion: A universal experience?* New York: Columbia University Press, 23-42

Fishman, P. (1977). Interactional shiftwork. *Heresies,* 2, 99–101

Fishman, P. (1978). Interaction: The work women do. *Social Problems,* 25, 397–406

Fishman, P. (1980). Conversational insecurity. In Howard, G., Robinson, R.W., Smith, P. (eds.) *Language: social psychological perspectives*. New York: Pergamon Press, 127-132

Fivush, R., Buckner, J. P. (2000). Gender, sadness, and depression: The development of emotional focus through gendered discourse. In Fischer, A.H. (Ed.) *Gender and emotion: Social psychological perspectives*. Cambridge, England: Cambridge University Press, 232-253

Flannery, R. (1946). Men's and women's speech in Gros Ventre. *International Journal of American Linguistics*, 12, 133-135

Flekova, L., Gurevych, I. (2013). Can We Hide in the Web? Large Scale Simultaneous Age and Gender Author Profiling in Social Media. *Notebook for PAN at CLEF*

Flexner, S.B., Wentworth. H. (1975). *Dictionary of American Slang*. New York: Crowell

Foote, R., Woodward, J. (1973). A Preliminary Investigation of Obscene Language. *The Journal of Psychology Interdisciplinary and Applied*, 83 (2), 263-275

Foucault, M. (1975). *Discipline and Punish: The Birth of the Prison*. Harmondsworth: Penguin Books

Fowler, R., Hodge, B., Kress, G., Trew, T. (1979). *Language and Control*. London: Routledge and Kegan Paul

Francis, W.N., Kučera, H. (1982). *Frequency analyses of English usage: Lexicon and grammar*. Boston: Houghton Mifflin

Frattaroli, J. (2006). Experimental disclosure and its moderators: A meta-analysis. *Psychological Bulletin*, 132, 823-865

Freed, A. (1992). We understand perfectly: A critique of Tannen's view of Cross-sex Communication. In Hall, K., Bucholtz, M., Moonwomon, B. (eds.) *Locating Power: Proceedings of the Second Berkeley Women and Language Conference.* Berkeley: Berkeley Women and Language Group, 144-152

Freed, A. (1996). Language and gender in an experimental setting. In Bergvall, V., Bing, J., Freed, A. (eds.) *Re-thinking Gender and Language Research: Theory and Practice.* London and New York: Longman, 54-76

Freed, A.F., Greenwood, A. (1996). Women, men, and type of talk: What makes the difference? *Language in Society*, 25 (1), 1-26

Freeman, R.D., McElhinny, B. (1996). Language and Gender. In McKay, S., Hornberger, N. (eds.) *Sociolinguistics and Language Teaching*. Cambridge: Cambridge University Press, 218-280

Friginal, E. (2009). The language of Outsourced Call Centers. Amsterdam/Philadelphia: John Benjamin's Publishing

Gal, S. (1991). Between speech and silence: the problematics of research on language and gender. In di Leonardo, M. (ed.) *Gender at the Crossroads of Knowledge: Feminist Anthropology in the Postmodern Era*. Berkeley, CA: University of California Press, 175-203

Gauthier, M. (2012). *Profanity and Gender: A Diachronic Analysis of Men's and Women's Use and Perception of Swear Words* (MA thesis). Universite Lumiere Lyon 2

Giddens, A. (1984). *The Constitution of Society*. Cambridge: Polity Press

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press

Gilligan, C. (1987). Moral orientation and moral development. In Feder Kittay, E., Meyers, D.T. (eds.) *Women and moral theory*. Totowa, NJ: Rowman & Littlefield

Gleason, Berko, J. (1975). Fathers and Other Strangers: Men's speech to Young Children. *26th Annual Roundtable*. Georgetown University Press, 289-297

Gleser, G.C., Gottschalk, L.A., Watkins, J. (1959). The relationship of sex and intelligence to choice of words: A normative study of verbal behavior. *Journal of Clinical Psychology*, 15, 183–191

Glovacki-Bernardi, Z. (2008). *Kad student zatrudni*. Zagreb: Alfa

Goddard A., Patterson, L.M. (2000). *Language and Gender*. The Eihosha LTD

Goedert, N., Karpowitz, C.F., Mendelberg, T. (2013). Does Descriptive Representation Facilitate Women's Distinctive Voice? How Group Gender Composition and Decision Rules Affect the Content of Deliberation. *American Journal of Political Science*, 58 (2), 291-306

Goffman, E. (1974). *Frame Analysis: An essay on the organisation of experience*. New York: Free Press

Goffman, E. (1981). *Forms of Talk*. Philadelphia: University of Pennsylvania Press

Goodwin, M.H. (1980). Directive-Response Speech Sequences in Girls' and Boys' Task Activities. In Borker, R., Furman, N. and McConnell-Ginet, S. (eds.) *Women and Language in Literature and Society*. New York: Praeger Publishers, 157-173

Gomm, I. (1981). *A study of the inferior image of the female use of the English language as compared to that of the male*. (BA dissertation). Edge Hill College, Ormskirk

Gorbatai, A., Nelson, L. (2015). *The Narrative Advantage: Gender and the Language of Crowdfunding*

Gottschalk, L.A., Defrancisco, D., Bechtel, R.J. (2002). Computerized Content Analysis of Some Adolescent Writings of Napoleon Bonaparte: A Test of the Validity of the Method. *The Journal of Nervous and Mental Disease*, 190, 542-548

Gottschalk, L.A., Gleser, G.C. (1969). The measurement of psychological states through the content analysis of verbal behavior. Berkeley: University of California Press

Graells-Garrido, E., Lalmas, M., Menczer, F. (2015). First Women, Second Sex: Gender Bias in Wikipedia. *Proceedings of 26th ACM Conference on Hypertext and Social Media*, 165-174

Griswold, D. (2007). *The politics of speech: engendering the public sphere*. Carnegie Mellon University

Graham, J., Haidt, J., Nosek, B.A. (2009). Liberals and Conservatives Rely on Different Sets of Moral Foundations. *Journal of personality and social psychology*, 96, 1029–1046

Greenwald, A. G., Breckler, S.J. (1985). To whom is the self presented? In Schlenker, B. (ed.) *The self and social life*. New York: McGraw-Hill, 126-145

Greif, E.B. (1980). Sex differences in parent—child conversations. *Women's Studies International Quarterly*, 3 (2-3), 253-258

Gross, J. J., John, O. P. (1998). Mapping the domain of expressivity: Multimethod evidence for a hierarchical model. *Journal of Personality and Social Psychology*, 74, 170-191

Guiller, J., Durndell, A. (2006). "I totally agree with you": Gender interactions in educational online discussion groups. *Journal of Computer Assisted Learning*, 22, 368–381

Gumperz, J. (1981). Conversational inference and classroom learning. In Green, J. and Wallat, C. (eds.) *Ethnography and Language in Educational Settings*. Norwood, NJ: Ablex, 3-23

Gumperz, J. (1982a). *Discourse Strategies.* Cambridge: Cambridge University Press

Gumperz, J. (ed.) (1982b). *Language and Social Identity*. Cambridge: Cambridge University Press

Gunsch, M.A., Brownlow, S., Haynes, S.E., Mabe, Z. (2000). Differential linguistic content of various forms of political advertising. *Journal of Broadcasting and Electronic Media,* 44, 27-42

Gunta R., Karapetjana, I. (2009). *The Use of Language in Political Rhetoric: Linguistic Manipulation*

Haas, A., Sherman, M. A. (1982). Conversational topic as a function of role and gender. *Psychological Reports*, 51, 453-454

Hakansson, J. (2012). The Use of Personal Pronouns in Political Speeches: A comparative study of the pronominal choices of two American presidents (Bachelor thesis). *Linnaeus University, School of Language and Literature*

Hall, C. (1985). Private persons versus public someones: class, gender and politics in England. 1780-1850. In Steedman, C., Urwin, C., Walkerdine, V. (eds.) *Language, Gender and Childhood.* London: Routledge and Keegan Paul, 10-33

Hall, K. (1995). Lip service on the fantasy lines. In Hall, K. and Bucholtz, M. (eds.) *Gender articulated: language and the socially constructed* self. London and New York: Routledge, 183-216

Hall, K. (2003). Exceptional Speakers: Contested and Problematized Gender Identities. In Holmes, J. and Meyerhoff, M. (eds.) *The Handbook of Language and Gender*. Malden, MA: Blackwell Publishing, 353-380

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

Hamdi, S., Dabaghi, A. (2012). Gender Differences in Iranian EFL Students' Letter Writing. *International Journal of Applied Linguistics and English Literature*, 1 (7), 155-169

Hanafiyeh, M., Afghari, A. (2014). Gender differences in the use of hedges, tag questions, intensifiers, empty adjectives, and adverbs: A comparative study in the speech of men and women. *Indian Journal of Fundamental and Applied Life Sciences*, 4 (4), 1168-1177

Hancock, A. B., Rubin, B. A. (2014). Influence of Communication Partner's Gender on Language. *Journal of Language and Social Psychology*, 34 (1), 46-64

Hansen, S. B. (1997). Talking about politics: gender and contextual effects on political proselytizing. *The Journal of Politics*, 51 (1), 73-103

Harding, S. (1975). Women and words in a Spanish village. In Reiter, R.R. (ed.) *Toward an anthropology of women*. New York: Monthly Review Press, 283-308

Harris, S. (1984). Questions as a Mode of Control in Magistrates Courts. *International Journal of the Sociology of Language*, 49, 5-28

Hartman, M. (1976). A descriptive study of the language of men and women born in Maine around 1900 as it reflects the Lakoff hypotheses in language and woman's place. In Dubois, B.L., Crouch, I. (eds.) *The sociology of the languages of American women*. San Antonio, TX: Trinity University Press, 81-90

Hearn, J. (1992). *Men in the Public Eye.* London: Routledge

Heath, R.A. (2006). *The Praeger Handbook of Transsexuality: Changing Gender to Match Mindset.* London: Praeger Publishers

Heilman, M. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues*, 57, 657–674

Hellinger, M., Bußmann, H. (eds.) (1984). *Gender Across Language: The Linguistic Representation of Women and Men*. Amsterdam/Philadelphia: John Benjamins Publishing Company

Henley, N. M., Kramarae, C (1991). Gender, power, and miscommunication. In Coupland, N., Giles, H., Wieman, J.M. (eds.) Miscommunication and problematic talk. Newbury Park, CA: Sage, 18-43

Herring, S.C., Paolillo, J.C. (2006). Gender and genre variation in weblogs. *Journal of Sociolinguistics*, 10 (4), 439-459

Hiatt, M. (1977). *The Way Women Write*. New York: Teachers College Press

Hirsh, J.B., Peterson, J.B. (2009). Personality and language use in self-narratives. *Journal of Research in Personality*, 43, 524-527

Hirschman, L. (1994). Female-Male Differences in Conversational Interaction. *Language in Society*, 23 (3), 427-442

Holmes, J. (1992). Women's talk in public contexts. *Discourse and Society*, 3 (2), 121-50

Holmes, J. (1995). *Women, men and Politeness.* London and New York: Longman

Holmes, J. (2001). *An Introduction to Sociolinguistics* (2nd edn.). London: Longman

Holmes, J., Meyerhoff, M. (eds.) (2003). *The Handbook of Language and Gender*. Blackwell Publishing

Holmes, J. (2006). *Gendered Talk at Work: Constructing Social Identity Through Workplace Interaction*. Blackwell Publishing

Holmes, J. (1993). Women's talk: The question of sociolinguistic universals. *Australian Journal of Communications*, 20, 125-149

Howe, C. (1997). *Gender and Classroom Interaction. A Research Review*. Edinburgh: The Scottish Council for Research in Education

Huddy, L., Terkildsen, N. (1993). Gender Stereotypes and the Perception of Male and Female Candidates. *American Journal of Political Science*, 37, 119-147

Houvardas, J., and Stamatatos, E. (2006). N-gram feature selection for authorship identification. *Artificial Intelligence: Methodology, Systems, and Applications*. Springer Berlin Heidelberg, 77-86

Howell, D.C. (2013). *Statistical Methods for Psychology*. Wadsworth

Huber-Carol, C., Balakrishnan, N., Nikulin, M.S., Mesbah, M. (2002). *Goodness-of-Fit Tests and Model Validity*. Boston: Birkhäuser

Huddy, L., Terkildsen, N. (1993). Gender Stereotypes and the Perception of Male and Female Candidates. *American Journal of Political Science*, 37, 119–147

Hughes, S.E. (1992). Expletives of lower working-class women. *Language in Society*, 21, 291-303

Iachini, T., Sergi, I., Ruggiero, G., Gnisci, A. (2005). Gender differences in object location memory in a real three-dimensional environment. *Brain and Cognition*, 59, 52-59

Ide, S. (1991). How and why do women speak more politely in Japanese? In Ide, S. and McGloin, N.H. (eds.) *Aspects of Japanese women's language*. Tokyo: Kurosio Publishers, 63-79

Iosub, D., Laniado, D., Castillo, C., Fuster Morell, M., Kaltenbrunner, A. (2014). Emotions under Discussion: Gender, Status and Communication in Online Collaboration. *PLos One*, 9 (8)

Ireland, M. (2008). *Candidates' Language in Speeches and Interviews: Summary Comparisons*

Ireland, M.E., Slatcher, R.B., Eastwick, P.W., Scissors, L.E., Finkel, E.J., Pennebaker, J.W. (2011). Language style matching predicts relationship initiation and stability. *Psychological science*, 22, 39-44

Ishikawa, Y. (2015). Gender Differences in Vocabulary Use in Essay Writing by University Students. *The Proceedings of 2nd Global Conference on Conference on Linguistics and Foreign Language Teaching*, 192, 593-600

Jaffe, M., Lee, Y., Huang, L., Oshagan, H. (1995). Gender, pseudonyms, and cmc: masking identities and baring souls. *Gender, Language and CMC*

James, D., Clarke, S. (1993). Women, men, and interruptions: A critical review. In Tannen, D. (ed.) *Gender and conversational interaction*. Oxford: Oxford University Press, 231-280

James, D., Drakich, J. (1993). Understanding gender differences in amount of talk: a critical review of research. In Tannen, D. (ed.) *Gender and Conversational Interaction.* New York and Oxford: Oxford University Press, 281-312

Jäger, S., Maier, F. (2009). Theoretical and Methodological Aspects of Foucaldian Critical Discourse Analysis and Dispositive Analysis. In Wodak, R., Meyer, M. (eds.) *Methods of Critical Discourse Analysis*. London: Sage, 34-61

Januschek, F. (1992). Rechts-populismus und NS-Anspielungen am Beispiel des österreichischen Politikers Jörg Haider. *DISS-Texte*, 15

Jay, T.B. (1992). *Cursing in America*. Philadelphia: John Benjamins Publishing Company

Jay, T.B. (1986). *2,170 Obscene words. Who, what, where, when*. Presented at Eastern Psychological Association, New York

Jay, T.B., Janschewitz, K. (2007). The pragmatics of swearing. *Journal of Politeness Research*, 4, 267–288

Jespersen, O. (1922). *Language; Its Nature, Development and Origin*. London: Unwin Brothers Ltd.

Jieun, J., Jae-Woong, C. (2009). A Key Word Analysis of English Intensifying Adverbs in Male and Female Speech in ICE-GB. *Proceedings of the 23rd Pacific Asia Conference on Language, Information and Computation*, 210-219

Jones, J.J. "Talk like a man" The linguistic appeal of Hillary Rodham Clinton. Revise and resubmit requested at Perspectives on Politics on 8/18/15

Jones, J., Stilwell Peccei, J. (2004). Language and Politics. In Thomas, L., Wareing, S., Singh, I., Stilwell Peccei, J., Thornborrow, J. and Jones, J. (eds.) *Language, Power and Solidarity*. London: Routledge, 35-54

Johnson, F. L., Aries, E. J. (1983). Conversational patterns among same-sex pairs of late-adolescent close friends. *The Journal of Genetic Psychology*, 142, 225-238

Jordan, K.N., Pennebaker. J.W. (2015). *Past, Present, and Future: Time Orientation in Republican Frontrunners*

Jordan, J.V., Surrey, J.L., Kaplan, A.G. (1991). Women and empathy: Implications for psychological development and psychotherapy. In Jordan, J.V., Kaplan, A.G., Miller, J.B., Stiver, I.P., Surrey, J.L. (eds.) *Women's growth in connection: Writings from the Stone Center*. New York: Guilford Press, 27-50

Jordan Brooks, D. (2013). *He Runs, She Runs: Why Gender Stereotypes Do Not Harm Women Candidates*. Princeton University Press

Johnston, A., Barton White, A. (1994). Communication Styles and Female Candidates: A Study of the Political Advertising During the 1986 Senate Elections. *Journalism Quarterly*, 71, 321-329

Kacewicz, E., Pennebaker, J., Davis, M., Jeon, M., Graesser, A. (2014). Pronoun Use Reflects Standings in Social Hierarchies. *Journal of Language and Social Psychology*, 33 (2), 125-143

Kahn, K. (1996). *The Political Consequences of Being a Woman: How Stereotypes Influence the Conduct and Consequences of Political Campaigns*. New York: Columbia University Press.

Kandiah, T. (1991). Extenuatory Sociolinguistics: Diverting Attention from Issues to Symptoms in Cross-Cultural Communication Studies. *Multilingua*, 10 (4), 345–380

Kangas, S.E.N. (2014). What can software tell us about political candidates? A critical analysis of a computerized method for political discourse. *Journal of Language and Politics*, 77-97

Kapidžić, S., Herring, S.C. (2011). Gender, Communication, and Self-Presentation in Teen Chatrooms Revisited: Have Patterns Changed? *Journal of Computer-mediated Communication*, 17 (1), 39-59

Karapetjana, I. (2011). Pronominal Choice in Political Interviews. *Baltic Journal of English Language, Literature and Culture*, 1, 36–45

Karpowitz, C.K., Mendelberg, T. (2014). *The Silent Sex: Gender, Deliberation, and Institutions*. London: Princeton University Press

Kathlene, L. (1994). Power and influence in state legislative policy making: the interaction of gender and position in committee hearing debates. *American Political Science Review*, 88 (3), 560-576

Kathlene, L. (1995). Position power versus gender power: Who holds the floor? In Duerst-Lahti, G., Kelly, R (eds.) *Gender, Power, Leadership and Governance.* Michigan: University of Michigan Press, 167-193

Keenan, E.O. (1974). Norm-makers, norm-breakers: Uses of speech by men and women in Malagasy community. In Bauman, R, Sherzer, J. (eds.) *Explorations in the Ethnography of Speaking*. Cambridge: Cambridge University Press, 125-143

Kendall, S., Tannen, D. (1997). Gender and language in the workplace. In Wodak, R. (ed.) *Gender and discourse*. London: SAGE Publications, 81-106

Kerswill, P. (1994). *Dialects converging: Rural speech in urban Norway*. Oxford: Oxford University Press

Keskin, S. (2006). Comparison of Several Univariate Normality Tests Regarding Type I Error Rate and Power of the Test in Simulation Based Small Samples. *Journal of Applied Science Research*, 2 (5), 296-300

Kirk, U. (1992). Evidence for early acquisition of visual organization ability: A developmental study. *The Clinical Neuropsychologist*, 6, 171-177

Kivran-Swaine, F., Brody, S., Diakopoulos, N., Naaman, M. (2012). Of joy and gender: emotional expression in online social networks. *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work Companion*, 139–142

Klein, J. (1965). *Samples from English cultures*. New York: Routledge & Kegan Paul

Kloch, Z. (2000). Language and Gender: Social and Psychological Determinants in Communication. *Psychology of Language and Communication*, 4 (2), 45-58

Koller, V. (2005). Critical discourse analysis and social cognition: evidence from business media discourse. *Discourse and Society*, 16 (2), 199-224

Komarovsky, M. (1962). *Blue collar marriage*. New York: Random House

Koppel, M., Argamon, S., Shimoni, A. (2002). Automatically categorizing written texts by author gender. *Literary and Linguistic Computing*, 17 (4), 401–412

Kotthoff, H., Wodak, R. (eds.) (1997). *Communicating Gender in Context*. Amsterdam/Philadelphia: John Benjamins Publishing

Kramer, A.D.I., Guillory, J.E., Hancock, J.T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111, 8788-8790

Kramer, C. (1974). Stereotypes of Women's Speech: The Word from Cartoons. *The Journal of Popular Culture*, 8 (3), 624-630

Kress, G. (1993). Against arbitrariness. *Discourse and Society*, 4 (2), 169–191

Kryeziu, L. (2015). Language and Gender Difference in Discourse. *Proceedings of the 5th International Conference on Foreign Language Teaching and Applied Linguistics and International Forum on Slavic Studies*. International Burch University

Kučera, J., Francis, W.N. (1967). *Computational Analysis of Present-Day American English.* Providence: Brown University Press

Kuo, S. (2003). Involvement vs detachment: gender differences in the use of personal pronouns in televised sports in Taiwan. *Discourse Studies*, 5 (4), 479-494

Krenn, B., Schreitter, S. Emo Dialogue (2015). Differences in male and female ways of communicating with affective autonomous conversational systems (project). *Austrian Research Institute for Artificial Intelligence*

Kress, G., Hodge, R. (1979). *Language as ideology*. London: Routledge and Kegan Paul

Kucuktunc, O., Cambazoglu, B.B., Weber, I., Ferhatosmanoglu, H. (2012). A large-scale sentiment analysis for Yahoo! answers. *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining.* WSDM '12, 633–642

Kupier, K. (1991). Sporting Formulae in New Zealand English: Two Models of Male Solidarity. In Coates, J. (ed.) (1998). *Language and Gender: A Reader*. Oxford: Blackwell, 285-293

Laberge, S., Sankoff, G. (1980). Anything you can do. In Sankoff, G. (ed.) *The Social Life of Language*. New York: Academic Press, 271-293

Labov, W. (1966). *The Social Stratification of (r) in New York City Department Stores.* Washington, D.C.: Center for Applied Linguistics

Lakoff, R. (1975). *Language and Woman's Place: Text and Commentaries*. New York: Harper and Row Publishers

Lakoff, R. (1990). *Talking Power: The Politics of Language.* San Francisco: Basic Books

Lakoff, R. (2003). Language, Gender, and Politics: Putting "Women" and "Power" in the Same Sentence. In Holmes, J., Meyerhoff, M. (eds.) *The Handbook of Language and Gender*. Blackwell Publishing, 161-178

Laserna. C., Seih. Y., Pennebaker, J. (2014). Um…Who Like Says You Know: Filler Word Use as a Function of Age, Gender, and Personality. *Journal of Language and Social Psychology*, 33 (3), 328-338

Larner, L.R. (2009). The Role of Feminine Rhetoric in Male Presidential Discourse: Achieving Speech Purpose. *CUREJ: College Undergraduate Research Electronic Journal*, University of Pennsylvania

Lave, J., Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press

Leaper, C., Ayres, M.A. (2007). A Meta-Analytic Review of Gender Variations in Adults' Language Use: Talkativeness, Affiliative Speech, and Assertive Speech. *Personality and Social Psychology Review*, 11, 328-363

Leaper, C., Robnett, R.D. (2011). Women Are More Likely Than Men to Use Tentative Language, Aren't They? A Meta-Analysis Testing for Gender Differences and Moderators. *Psychology of Women Quarterly*, 35 (1), 129-142

Lee, Y. (2009). Measures of student attitude on aging. *Educational Gerontology*, 35, 121-134

Lee Kaid, L (ed.) (2004). Handbook of Political Communication. New Jersey: Lawrence Erlbaum Associates, Inc.

Leeper, M. (1991). The Impact of Prejudice on Female Candidates: An Experimental Look at Voter Inference. *American Politics Quarterly*, 19, 248–61

Lemke, J. (2003). Texts and Discourses in the Technologies of Social Organization. In Weiss, G., Wodak, R. (eds.) *Critical Discourse Analysis: theory and interdisciplinarity*. Palgrave Macmillan, 130-149

Lester, D. (2004). *Katie's diary: Unlocking the Mystery of a Suicide*. London: Routledge

Lester, D., Haines, J., Williams, C. L. (2010). Content differences in suicide notes by sex, age, and method: A study of Australian suicide notes. *Psychological Reports*, 106, 475–476

Li, H., Cai, Z., Graesser, A.C., Duan, Y. (2014). A Comparative Study on English and Chinese Word Uses with LIWC. *Proceedings of the Twenty-Fifth International Florida Artificial Intelligence Research Society Conference.* The AAAI Press: Palo Alto, California, 190-195

Li, G., Papadopoulos, A. (2002). A note on goodness of fit test using moments. *Statistica*, 62 (1), 72–86

Limbrick, P. (1991). A study of male and female expletive use in single and mixed-sex situations. *Te Reo*, 34, 71-89

Lin, H., Qiu, L. (2013). Two Sites, Two Voices: Linguistic Differences between Facebook Status Updates and Tweets. *Lecture Notes in Computer Science*, 8024, 432-440

Ling, R., Baron, Naomi S. (2007). Text Messaging and IM Linguistic Comparison of American College Data. *Journal of Language and Social Psychology*, 28 (3), 291-298

Link, J. (1983). Was ist un was bringt Diskurstaktik. *kulturRRevolution*, 1, 6-21

Lipman-Blumen, J. (1984). *Gender Roles and Power*. Englewood Cliffs, NJ: Prentice-Hall

Litosseliti, L. (2006). *Gender and Language: Theory and Practice*. Oxford: Oxford University Press Inc.

Lepore, S.J., Smyth, J.M. (2002). *The Writing Cure: How Expressive Writing Promotes Health and Emotional Well-Being*. Washington, DC: American Psychological Association

Lockhart, M., Mollick, K. (eds.) (2013). *Political Women: Language and Leadership*. Lanham: Lexington

Lovenduski, J. (1996). Sex, gender and British politics. In Lovenduski, J., Norris, P. (eds.) *Women in Politics.* Oxford: Oxford University Press, 3-18

Lovenduski, J. (2005). *Feminizing Politics*. Cambridge: Polity

Ludu, P.S. (2014). *Inferring gender of a Twitter user using celebrities it follows*

Lyons, J. (1977). *Semantics*. Cambridge: Cambridge University Press

Maas, U. (1984). Als der Geist der Gemeinschaft eine Sprache fand. *Sprache im Nationalsozialismus. Westdeutscher Verlag*

Macaulay, R.K.S. (1997). *Standards and variation in urban speech: Examples from Lowland Scots*. Amsterdam: Benjamin

Mairesse, F., Walker, M.A., Mehl, M.R., Moore, R.K. (2007). Using linguistic cues for the automatic recognition of personality in conversation and text. *Journal of artificial intelligence research*, 30 (1), 457–500

Malone, M.J. (1997). *Worlds of Talk: The presentation of self in everyday conversation*. Oxford: Polity Press

Maltz, D., Borker, R. (1982). A cultural approach to male-female miscommunication. In Gumberz, J. (ed.) *Language and Social Identity*. Cambridge: Cambridge University Press, 196-216

Manjavacas, E. (2015). *Statistical description of gender-related differences in language use in a Dutch blog corpus*

Manning, J.J. (2014). Membership of the 113th Congress: A Profile. *Congressional Research Service*. Retrieved from https://www.senate.gov/CRSReports/crs-publish.cfm?pid=%260BL%2BR%5CC%3F%0A

Manson, M. (2014). *A brief history of male/female relations.* Retrieved from http://markmanson.net/male-female-relations

Marmolejo-Ramos, F., González-Burgos, J. (2012). A Power Comparison of Various Tests of Univariate Normality on Ex-Gaussian Distributions. In Hox, J. and Balluerka, N. (eds.) *Methodology. European Journal of Research Methods for the Behavioral and Social Sciences*. 1-15

Martin, E. (1991). The egg and the sperm: how science has constructed a romance based on stereotypical male-female roles. *Signs: Journal of Women in Culture and Society*, 16, 485-501

Martin Rojo, L., Gomez Esteban, C. (2005). The Gender of Power: The Female Style in Labour Organizations. In Lazar, M.M. (Ed.) *Feminist Critical Discourse Analysis*. Palgrave Macmillan, 61-89

Mautner, G. (2005). Time to get wired: Web as corpus in CDA. *Discourse and Society*, 16 (6), 809-829

McArthur, T. (ed.). (1992). *The Oxford Companion to the English Language*. Oxford: Oxford University Press

McClelland, D. C. (1979). Inhibited power motivation and high blood pressure in men. *Journal of Abnormal Behavior*, 88, 182-109

McElhinny, B.S. (1998). I Don't Smile Much Anymore: Affect, Gender and the Discourse of Pittsburgh Police Officers'. In Coates, J. (ed.) *Language and Gender: A Reader*. Oxford: Blackwell, 309-327

McFadyen, R. G. (1996). Gender, status and 'powerless' speech: Interactions of students and lecturers. *British Journal of Social Psychology*, 35, 353-367

McKinlay, A., Potter, J. (1987). Social representations: A conceptual critique. *Journal for the Theory of Social Behaviour*, 17, 471-487

McMillan, J.R., Clifton, A.K., McGrath, D., Gale, W.S. (1977). Women's language: Uncertainty or interpersonal sensitivity and emotionality. *Sex Roles*, 3 (6), 545-559

Medilexicon dictionary available at http://www.medilexicon.com/

Mehl, M.R., Gosling, S.D., Pennebaker, J.W. (2006). Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology*, 90 (5), 862-877

Mehl, M.R., Pennebaker, J.W. (2003). The sounds of social life: A psychometric analysis of students' daily social environments and natural conversations. *Journal of Personality and Social Psychology*, 84, 857–870

Mendes, M., Pala, A. (2003). Type I Error Rate and Power of Three Normality Tests. *Pakistan Journal of Information and Technology*, 2 (2), 135-139

Merchant, K. (2012). *How Men And Women Differ: Gender Differences in Communication Styles, Influence Tactics, and Leadership Styles*. CMC Senior Theses.

Merten, C.H., Beal, C. (1999). An analysis of the Hooper Visual Organization test with neurological patients. *The Clinical Neuropsychologist*, 13, 521-529

Meyner, S. (2013). *Do the Speeches of Japanese the female politicians Feature Japanese Women's Language?* (thesis) Goteborgs Univerzitet

Miculicich Werlen, L. (2015). Statistical Learning Methods for Profiling Analysis (poster). *Conference handbook of Empirical Methods in Natural Language Processing*. Lisbon: Association for Computational Linguistic

Mills, S. (1995). *Feminist Stylistics*. London: Routledge

Moir, A., Jessel, D. (1992). *Brain sex: The real difference between men and women*. New York: Dell Publishing

Mondorf, B. (2002). Gender Differences in English Syntax. *Journal of English Linguistics*, 30 (2), 158-180

Moore, H. T. (1922). Further data concerning sex differences. *Journal of Abnormal Psychology*, 17, 210-214

Mougeon, R., Beniak, E. (1996). Social class and language variation in bilingual speech communities. In Guy, G.R., Feagin, C., Schiffrin, D., Baugh, J. (eds.) *Towards a Social Science of Language: Papers in honor of William Labov. Variation and Change in Language and Society*. Amsterdam: Benjamins, 69-99

Mulac, A. (1989). Men's and Women's Talk in Same-Gender and Mixed-Gender Dyads: Power or Polemic? *Journal of Language and Social Psychology*, 8 (3-4), 249-270

Mulac, A. (1998). The gender-linked language effect: Do language differences really make a difference?. In Canary, D.J., Dindia, K. (eds.) *Sex differences and similarities in communication*. Mahwah, NJ: Erlbaum, 127-153

Mulac, A. (2006). The gender-linked language effect: Do language differences really make a difference? In Dindia, K., Canary, D.J. (eds.) *Sex differences and similarities in communication*. Mahwah, NJ: Lawrence Erlbaum Associates, 195-215

Mulac, A., Giles, H., Bradac, J., Palomares, N. (2013). The gender-linked language effect: an empirical test of a general process model. *Language Sciences*, 38, 22–31

Mulac, A., Lundell, T.L., Bradac, J.J. (1986). Male/female language differences and attributional consequences in a public speaking situation: Toward an explanation of the gender-linked language effect. *Communication Monographs*, 53, 115-129

Mulac, A., Lundell, T.L. (1986). Linguistic contributors to the gender-linked language effect. *Journal of Language and Social Psychology*, 5, 81-101

Mulac, A., Lundell, T.L. (1994). Effects of gender-linked language differences in adults' written discourse: Multivariate tests of language effects. *Language and Communication*, 14, 299-309

Mulac, A., Seibold, D.R., Farris, J.L. (2000). Female and male managers' and professionals' criticism giving: Differences in language use and effects. *Journal of Language and Social Psychology*, 19 (4), 389–415

Mulac, A., Studley, L.B., Blau, S. (1990). The gender-linked language effect in primary and secondary students' impromptu essays. *Sex Roles*, 23, 439-469

Mulac, A., Wiemann, J.M., Widenmann, S.J., Gibson, T.W. (1988). Male/female language differences and effects in same-sex and mixed-sex dyads: The gender-linked language effect. *Communication Monographs*, 55 (4), 315-335

Nagarajan, M., Hearst, M.A. (2009). An Examination of Language Use in Online Dating Profiles. *Proceedings of the Third International ICWSM Conference*. Association for the Advancement of Artificial Intelligence, 266-269

Nardi, B. A., Whittaker, S., Bradner, E. (2000). Interaction and outeraction: Instant messaging in action. *Paper presented at the Conference on Computer Supported Cooperative Work*, Philadelphia, PA

Ndambuki, J., Janks, H. (2010). Political Discourses, Women's Voices: Mismatches in Representation. *Critical Approaches to Discourse Analysis across Disciplines*, 4 (1), 73-92

Newman, M., Groom, C., Handelman, L., Pennebaker, J. (2008). Gender Differences in Language Use: An Analysis of 14 000 Text Samples. *Discourse Processes*, 45, 211–2362-1990

Newman, M.L., Pennebaker, J.W., Berry, D.S., Richards, J.M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and social psychology bulletin*, 29, 665-675

Ng, W., Leung, M. D. (2015). For love or money? Gender differences in how one approaches getting a job. *Institute for Research on Labor and Employment*, UC Berkeley, 1-47

Nicholson, L. (1994). Interpreting gender. *Signs*, 20 (1), 79-105

Nicolau, M. (2013). Gender Differences towards the Use of Conjunctions (doctoral dissertation). *Atma Jaya Catholic University*, Jakarta, Indonesia

Norris, P. (1996). Women Politicians: Transforming Westminster?, In Lovenduski, J., Norris, P. (eds.) *Women in Politics.* Oxford: Oxford University Press, 91-104

Noughabi, H.A., Arghami, N.R. (2011). Monte Carlo comparison of seven normality tests. *Journal of Statistical Computation and Simulation*, 81 (8), 965–972

Nowson, S. (2006). *The Language of Weblogs: A study of genre and individual differences* (doctoral dissertation). Institute for Communicating and Collaborative Systems: School of Informatics. University of Edinburgh

O'Barr, W. M. (1982). *Linguistic evidence: Language, power, and strategy in the courtroom*. New York: Academic

O' Barr, W., Atkins, B. (1980). Women's language" or "powerless language? In Coates, J. (ed.) (1998) Language and gender: A reader. Oxford: Blackwell Publishing

O' Barr, W. M., Atkins, B. (1998). "Women's language" or "powerless language"? In Coates, J. (ed.) *Language and gender: A reader*. Oxford: Blackwell Publishing

Oberlander, J., Gill, A. (2004). Individual Differences and Implicit Language: Personality, Parts-of-Speech and Pervasiveness. In *Proceedings of the 26th Annual Conference of the Cognitive Science Society*, Chicago, IL, USA, 1035-1040

O'Brien, R., Hunt, K., Hart, G. (2005). 'It's caveman stuff, but that is to a certain extent how guys still operate': Men's accounts of masculinity and help seeking. *Social Science and Medicine*, 61, 503–516

Ochs, E. (1993). Constructing Social Identity: A Language Socialization Perspective. *Research on Language and Social Interaction*, 26 (3), 287-306

O'Kearney, R., Dadds, M. (2004). Developmental and gender differences in the language for emotions across the adolescence years. *Cognition and emotion*, 18 (7), 913-938

Oliver, M.M., Rubin, J. (1975). The use of expletives by some American women. *Anthropological Linguistics*, 17, 191-197

Otheguy, R., Zentella, A.C. (2012). *Spanish in New York: Language contact, dialectal leveling, and structural continuity*. Oxford: Oxford University Press

Ottoni, R., Pesce, J.P., Las Casas, D., Franciscani Jr., G., Meira Jr., W., Kumaraguru, P., Almeida, V. (2013). Ladies First: Analyzing Gender Roles and Behaviors in Pinterest. *Association for the Advancement of Artificial Intelligence*, 1-9

Palmer, F.R. (1986). *Mood and Modality*. Cambridge: Cambridge University Press

Palomares, N. A. (2004). Gender schematicity, gender identity salience, and gender-linked language use. *Human Communication Research*, 30, 556-588

Palomares, N. A. (2008). Explaining gender-based language use: Effects of gender identity salience on references to emotion and tentative language in intra- and intergroup contexts. *Human Communication Research*, 34, 263-286

Palomares, N. A. (2009). Women Are Sort of More Tentative Than Men, Aren't They?: How Men and Women Use Tentative Language Differently, Similarly, and Counterstereotypically as a Function of Gender Salience. *Communication Research*, 36 (4), 538-560

Palomares, N. A., Reid, S. A., Bradac, J. J. (2004). A self-categorization perspective on gender and communication: Reconciling the gender-as-culture and dominance explanations. In Ng, S.H., Candlin, C.N., Chiu, C.Y. (eds.) *Language matters: Communication, identity, and culture*. Hong Kong: City University of Hong Kong Press, 85-109

Parsons, T.D., Larson, P., Kratz, K., Thiebaux, M., Bluestein, B., Buckwalter, J.G., Rizzo, A.A. (2004). Sex differences in mental rotation and spatial rotation in a virtual environment. *Neuropsychologia*, 42, 555-62

Pasupathi, M. (2007). Telling and the remembered self: Linguistic differences in memories for previously disclosed and previously undisclosed events. *Memory*, 15, 258-270

Pateman, C. (1998). *Ženski nered: demokracija, feminizam i politička teorija*. Standford: Standford University Press

Pennebaker, J.W. (1997). Writing about emotional experiences as a therapeutic process. *Psychological Science*, 8, 162-166

Pennebaker, J.W. (2013). *The Secret Life of Pronouns: What Our Words Say About Us*. New York: Bloomsbury Press

Pennebaker, J.W., Chung, C.K., Ireland, M., Gonzales, A., Booth, R.J. (2007). The Development and Psychometric Properties of LIWC2007. Austin, Texas: LIWC.net

Pennebaker, J.W., Chung, C.K. (2011). Expressive writing and its links to mental and physical health. In Friedman, H.S. (ed.), *Oxford handbook of health psychology*. New York, NY: Oxford University Press

Pennebaker, J.W., Francis, M.E., Booth, R.J. (2001). *Linguistic Inquiry and Word Count (LIWC): LIWC2001*. Mahwah: Lawrence Erlbaum Associates

Pennebaker, J.W., Groom, C.J., Loew, D., Dabbs, J.M. (2004). Testosterone as a Social Inhibitor: Two Case Studies of the Effect of Testosterone Treatment on Language. *Journal of Abnormal Psychology*, 113 (1), 172–175

Pennebaker, J.W., Ireland, M.E. (2011). Using literature to understand authors: The case for computerized text analysis. *Scientific Study of Literature*, 1, 34-48

Pennebaker, J.W., Jordan, K.N. (2015). *Through Their Words: Hillary, Bernie, and their Republican Friends - Text analysis of the first Democratic Debate*

Pennebaker. J.W., King, L. (1999). Linguistic Styles: Language Use as an Individual Difference. *Journal of Personality and Social Psychology*, 77 (6), 1296-1312

Pennebaker, J.W., Lay, T. C. (2002). Language use and personality during crises: Analyses of Mayor Rudolph Giuliani's press conferences. *Journal of Research in Personality*, 36, 271–282

Pennebaker, J.W., Mehl, M.R., Niederhoffer, K.G. (2003). Psychological Aspects of Natural Language Use: Our words, Our Selves. *Annual Review of Psychology*, 54, 547-577

Pennebaker J.W., Stone L. (2003) Words of wisdom: Language use over the life span. *Journal of personality and social psychology*, 85 (2), 291-301

Pennycook, A. (1993). The politics of pronouns. Oxford University Press: *ELT Journal*, 48 (2), 173-178

Pilkington, J. (1992). 'Don't try and make out that I'm nice!' The Different Strategies that Women and Men Use While Gossiping. In Coates, J. (ed.) (1998). *Language and Gender: A Reader*. Oxford: Blackwell, 254-269

Ponocny, I. (2001). Nonparametric goodness-of-fit tests for Rasch models. *Psychometrika*, 66 (3), 437-460

Poole, M. (1979). Social class, sex, and linguistic coding. *Language and Speech*, 22, 49-67

Popova, E. (2015). Linguistic Differences between Genders: The use of Past-tense verbs in Online reviews (thesis). *Tilburg University*.

Potter, J., Billig, M. (1992). Re-representing representations. Ongoing *Production on Social Representations*, 1, 15-20

Potter, J., Edwards, D. (2001). Sociolinguistics, cognitivism and discursive psychology. In Coupland, N., Sarangi, S., Candlin, C.N. (eds.) *Sociolinguistics and Social Theory*. London: Longman, 88-104

Potter, J., Wetherell, M. (1987). *Discourse and social psychology: Beyond attitudes and behaviour*. London: Sage

Potter, J., Wetherell, M. (1998). Social representations, discourse analysis and racism. In Flick, U. (ed.) *The psychology of the social*. Cambridge: Cambridge University Press, 138-155

Power, M., Berardone, M. (1998). Speaking in Parliament: first speeches of men and women. *Journal of Applied Social Behaviour*, 4 (2), 42-55

Quirk, R., Greenbaum, S., Leech, G., Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman

Rafaeli, S., Sudweeks, F., Konstan, J., Mabry, E. (1993). *Project overview: A collaborative quantitative study of computer-mediated communication*

Razali, N.M., Wah, Y.B. (2010). Power comparisons of some selected normality tests. *Proceedings of the Regional Conference on Statistical Sciences*, 126-138

Raumolin-Brunberg, H. (1998). Social factors and pronominal change in the seventeenth century: The Civil-War effect? In Fisiak, J., Krygier, M. (eds.) *Advances in English Historical Linguistics*. Mouton de Gruyter, 361-388

Reid, S. A., Keerie, N., Palomares, N. A. (2003). Language, gender salience, and social influence. *Journal of Language and Social Psychology*, 22, 210-233

Reiman, T. *Gender differences in Communication. Body Language University*

Reiter, R. R. (1975). Men and women in the south of France: Public and private domains. In Reiter, R.R. (ed.) *Toward an Anthropology of Women*. New York: Monthly Review Press, 252-282

Rilea, S.L., Roskos-Ewoldsen, B., Boles, D. (2004). Sex differences in spatial ability: a lateralization of function approach. *Brain and Cognition*, 56, 332-43

Risch, B. (1987). Women's derogatory terms for the men: That's right, dirty words. *Language in Society*, 16, 353-358

Robinson, R., Navea, R., Ickes, W. (2013). Predicting Final Course Performance From Students' Written Self-Introductions: A LIWC Analysis. *Journal of Language and Social Psychology*, 20 (2), 1-12

Romaine, S. (1994). *Language in society: An introduction to sociolinguistics*. Oxford: Oxford University Press

Romaine, S. (1999). *Communicating gender*. London: Oxford University

Rosenberg, S.D., Tucker, G.J. (1978). Verbal behavior and schizophrenia: The semantic dimension. *Archives of General Psychiatry*, 36, 1331-1337

Rosenwasser, S., Dean, N. (1989). Gender Role and Political Office: Effects of Perceived Masculinity/Femininity of Candidate and Political Office. *Psychology of Women Quarterly*, 13, 77-85

Rubin, D., Greene, K. (1992). Gender-typical style in written language. *Research in the Teaching of English*, 26 (1), 7–40

Rudman, L. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. *Journal of Personality and Social Psychology*, 74, 629–645

Ruijuan, Y. (2010). The Interpersonal Metafunction Analysis of Barack Obama's Victory Speech. *English Language Teaching*, 3 (2), 146-151

Sachs, J., Goldman, J., Chaille, C. (1984). Planning in pretend play: Using language to coordinate narrative development. In Pellegrini, A.D., Yawkey, T.D. (eds.) *The development of oral and written language in social contexts*. Norwood, NJ: Ablex, 119-128

Sacks, H. (1992). *Lectures on Conversations*. Oxford: Blackwell

Sadker, M., Sadker, D. (1985). Sexism in the Schoolroom of the 80's. *Psychology Today*, 54-57

Saeed, A., Hussain, S., Surhan, F. (2015). Prepositional Errors in the Writings of Pakistani ESL Students. *International Journal of English Linguistics*, 5 (3), 120-128

Sapir, E. (2013). *Jezik: uvod u istraživanje govora*. Zagreb: Institut za hrvatski jezik i jezikoslovlje

Savicki, V., Lingenfelter, D., Kelley, M. (1996). Gender language style in group composition in Internet discussion groups. *Journal of Computer-Mediated Communication*, 2 (3)

Scates, C. (1981). *A Sociolinguistic Study of Male/Female Language in Freshman Composition* (doctoral dissertation). University of Southern Mississippi, Hattiesburg, Mississippi

Schler, J., Koppel, M., Argamon, S., and Pennebaker, J.W. (2006). Effects of age and gender on blogging. *Proceedings of 2006 AAAI Spring Symposium on Computational Approaches for Analyzing Weblogs*, 199-205

Schmader, T., Whitehead, J., Wysocki, V.H. (2007). A Linguistic Comparison of Letters of Recommendation for Male and Female Chemistry and Biochemistry Job Applicants. *Sex Roles*, 57 (7-8), 509-514

Schultz, P. (2013). Gender variation in writing: Analyzing online dating ads. *Coyote Papers*, 21

Schwartz, H.A., Eichstaedt, J.C., Kern, M.L., Dziurzynski, L., Ramones, S.M., Agrawal, M., Shah, A., Kosinski, M., Stillwell, D., Seligman M.E.P., Ungar, L.H. (2013). Personality, Gender, and Age in the Language of Social Media: *The Open-Vocabulary Approach*. PLoS ONE 8 (9)

Schwartz, S., Rubel, T. (2005). Sex Differences in Value Priorities: Cross-Cultural and Multimethod Studies. *Journal of Personality and Social Psychology*, 89, 1010–1028

Scollon, R., Scollon, S. (2002). *Nexus analysis: Discourse and the Emerging Internet*. Routledge

Seale, C. (2006). Gender accommodation in online cancer support groups. *Health*, 10 (3), 345-360

Searle, J.R. (1969). *Speech Acts*. Cambridge University Press

Seidlitz, L., Diener, E. (1998). Sex differences in the recall of affective experiences. *Journal of Personality and Social Psychology*, 74, 262-271

Selnow, G. W. (1985). Sex differences in uses and perceptions of profanity. *Sex Roles*, 12 (3/4), 303–312

Severance, L.E. (2012). *The Gender Dynamics of Dissent in Organizations* (doctoral dissertation). University of Maryland

Shapiro, J. (1981). Anthropology and the study of gender. *Soundings: An Interdisciplinary Journal*, 64, 446-65

Shapiro, S.S., Francia, R.S. (1972). An Approximate Analysis of Variance Test for Normality. *Journal of the American Statistical Association*, 67 (337), 215-216

Shapiro, S.S., Wilk, M.B. (1965). An analysis of variance test for normality. *Biometrika*, 52 (3/4), 591–611

Shapiro, S.S., Wilk, M.B., Chen, H.J. (1968). A comparative study of various tests for normality. *Journal of the American Statistical Association*, 63 (324), 1343–1372

Sherman, J. (1978). *Sex-related Cognitive Differences: An Essay on Theory and Evidence.* Springfield, IL

Shima Ahmadi-Azad, S., Azad, I. (2015). Gender Differences in Emotional Content of EFLWritten Narratives. *Journal of Language Teaching and Research*, 6 (3), 619-626

Shofwan, H., Musli, A., Sugeng, A. (2013). An analysis on the gender-based difference of the eight Year students' recount paragraph writing based on the grammatical errors at smpn 3 Jnggawah Jember. *Jurnal Pendidikan dan Pembelajaran*, 2 (2), 75-84

Shotter, J., Gergen, K. (1994). Inquiries in social constructionism. In Sarbin, T.R., Kitsuse, J.I. (eds.) *Constructing the Social.* London: Sage, 1-15

Singh Ludu, P. (2014). *Inferring gender of a Twitter user using celebrities it follows*

Skarphol Kaml, S. (2000). The Fusion of Populist and Feminine Styles in the Rhetoric of Ann Richards. In DeVore, B. and Mayhead, M.A. (eds.) *Navigating Boundaries: The Rhetoric of Women Governors*. Westport, CT: Praeger Publishers, 59-82

Slatcher, R.B., Chung, C.K., Pennebaker, J.W., Lori D., Stone, L.D. (2007). Winning words: Individual differences in linguistic style among U.S. presidential and vice presidential candidates. *Journal of Research in Personality*, 41, 63–75

Smith, F.L. (1993). The pulpit and woman's place: Gender and the framing of the "exegetical" self in sermon performance. In Tannen, D. (ed.) *Framing in discourse*. New York: Oxford University Press, 147-175

Soler, J. (2013). *Author Profiling: Gender Identification* (Master thesis). Universitat Pompeu Fabra

Solheim, B.O. (2000). On *Top of the World: Women's Political Leadership in Scandinavia.* Wesport CT: Greenwood Press

Speer, S.A., Stokoe, E. (eds.) (2011). *Conversation and Gender*. Cambridge: Cambridge University Press

Spender, D. (1980). *Man Made Language*. London: Routledge

Steele, M., Chaseling, J. (2006). Powers of discrete goodness-of-fit test statistics for a uniform null against a selection of alternative distributions. *Communication in Statistics – Simulation and Computation*, 35, 1067–1075

Staley, C. M. (1978). Male-female use of expletives: A heck of a difference in expectations. *Anthropological Linguistics*, 20, 367-380

Stepney, C. (2014). *What's Important to Me: Identifying At-Risk and Resilient Students Through Narrative Writing about Personal Values* (MA thesis). The State University of New Jersey

Steinberg, P. (2012). Can Social Media Help Organize Public Opinion? A Case Study of Iranian Public Opinion After the 2009 Election. In Elson, S.B., Yeung, D., Roshan, P.,

Bohandy, S.R., Nader, A. (eds.) *Using Social Media to Gauge Iranian Public Opinion and Mood After the 2009 Election.* RAND Corporation: Santa Monica

Stiles, W.B. (1992). *Describing talk: A taxonomy of verbal response modes*. Newbury Park, CA: Sage

Stipek, D.J., Galinsky, J.H. (1991). Gender differences in children's achievement-related beliefs and emotional responses to success and failure in mathematics. *Journal of Educational Psychology*, 83 (3), 361-371

Suborn, F. (2013). Gender Differences in the Use of Linguistic Forms in the Speech of Men and Women in the Malaysian Context. *Journal Of Humanities And Social Science*, 13 (3), 67-79

Sunderland, J. (2004). *Gendered Discourse*. New York: Palgrave Macmillan

Swacker, M. (1976). Women's Verbal Behavior at Learned and Professional Conferences. In Dubois, B.L., Crouch, I. (eds.) *The Sociology of the Languages of American Women*. San Antonio: Trinity University, 155-160

Talbot, M. (1998). *Language and gender: an introduction*. Cambridge: Polity Press

Tamale, S. (2000). "Point of Order Mr. Speaker": African women claiming their space in parliament. *Gender and Development*, 8 (3), 8-25

Tanaka, L. (2004). *Gender, Language and Culture – A study of Japanese television interview discourse*. Amsterdam/Philadelphia: John Benjamins Publishing Company

Tannen, D. (1983). "I Take Out the Rock – Dok!": How Greek Women Tell about Being Molested (and Create Involvement). *Anthropological Linguistics*, 25 (3), 359–374

Tannen, D. (1986). *That's not what I meant: how conversational style makes or breaks your relationship with others*. London: Virago Press

Tannen, D. (1989). *Talking Voices*. Cambridge: Cambridge University Press

Tannen, D. (1990). *You Just Don't Understand: Women and Men in Conversation*. New York: Ballantine Books

Tannen, D. (1993). The Relativity of Linguistic Strategies: Rethinking power and solidarity in gender and dominance. In Tannen, D. (ed.) *Gender in conversational language*. New York: Oxford University Press, 165-188

Tannen, D. (1994). *Gender and Discourse*. Oxford: Oxford University Press

Tannen, D. (1997). The Power of Talk: Who Gets Heard and Why. In Oaks, D. (ed.) *Linguistics at Work: A Reader of Applications*. Fort Worth: Harcourt, 242-259

Tausczik, Y., 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

Thelwall, M. (2008). Fk yea I swear: Cursing and gender in a corpus of MySpace pages. *Corpora*, 3 (1), 83-107

Thode, H. Jr. (2002). *Testing for Normality*. New York: Marcel Dekker

Thomas, S. (1994). *How Women Legislate.* New York: Oxford University Press

Thomson, R., Murachver, T. (2001). Predicting gender from electronic discourse. *British Journal of Social Psychology*, 40, 193–208

Thorne, B., Henley, N. (1975). Difference and dominance: An overview of language, gender, and society. In Thorne, B., Henley, N. (eds.) *Language and sex: Difference and dominance*. Rowley, MA: Newbury House Publisher, 5-42

Trew, T. (1979). Theory and ideology at work. In Fowler, R., Hodge, B., Kress, G., Trew, T. (eds.) *Language and Control*. Boston: Routledge and Kegan Paul, 94-116

Triandis, H.C. (1989). The self and social behavior in differing cultural contexts. *Psychological Review*, 96, 506-520

Troemel-Ploetz, S. (1991). Review Essay: Selling the apolitical. *Discourse and Society*, 2 (4), 489-502

Trudgill, P. (1974). Linguistic change and diffusion: Description and explanation in sociolinguistic dialect geography. *Language in Society*, 3, 215-246

Trudgill, P. (2000) *Sociolinguistics: An Introduction to Language and Society* (4th edn.). London: Penguin

Trudgill, P. (2003). *A Glossary of Sociolinguistics*. Oxford: Oxford University Press

Turner, L., Dindia, K., Pearson, J. (1995). An Investigation of Female/Male Verbal Behaviors in Same-Sex and Mixed-Sex Conversations. *Communication Reports*, 8 (2), 86-96

Uchida, A. (1992). When 'difference' is 'dominance': A critique of the 'antipower-based' cultural approach to sex differences. *Language in Society*, 21, 547-568

Umberson, D. (1992). Gender, marital status and the social control of health behavior. *Social Science and Medicine*, 34 (8), 907–917

van Dijk, Teun A. (1983). *Minorities in the media*. Amsterdam: Socialistische Uitgeverij Amsterdam

van Dijk, T.A. (1988). *News Analysis: Case Studies of International and National News in the Press*. Hillsdale, N.J.: Lawrence Erlbaum Associates

van Dijk, T.A. (1993). *Elite Discourse and Racism.* London: Sage Publications

van Dijk, T.A. (1995). Discourse Analysis as Ideology Analysis. In Schaffner, C., Wenden, A.L. (eds.) *Language and Peace*. Dartmouth: Aldershot, 17-33

van Dijk, T.A. (1998). Opinions and Ideologies in the Press. In Bell, A., Garrett, P. (eds.) *Approaches to Media Discourse*. Oxford: Blackwell, 21-63

van Dijk, T.A. (2006). Discourse and Manipulation. *Discourse and Society*, 17 (3), 359–83

Vali, S., Kianiparsa, P. (2010). Gender differences in using cohesive devices while writing (research paper). *Weekly eNewsletter for the Teachers of English*

Vuorinen, A. (2002). The Gender Role of Queen Elizabeth I as Reflected by her Language. *Helsinki English Studies,* 2, 1-14

Walsh, C. (2001). *Gender and Discourse: Language and Power in Politics, the Church and Organisations.* London: Longman

Wang, M. R., Hsieh, S. C. Y. (2007). Gender Differences in the Language for Emotions. *Asian Journal of Management and Humanity Sciences*, 2, (1-4), 89-97

Waseleski, C. (2006). Gender and the Use of Exclamation Points in Computer-Mediated Communication: An Analysis of Exclamations Posted to Two Electronic Discussion Lists. *Journal of Computer-Mediated Communication*, 11, 1012-1024

Watson, D., Clark, L.A., Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063-1070

Watson, D., Pennebaker, J.W. (1989). Health complaints, stress, and distress: exploring the central role of negative affectivity. *Psychological review*, 96, 234–254

Weatherall, A. (2002). *Language, Gender and Discourse*. New York: Routledge

Webb, L., Lee, B. (2012). Blogs and blogging. In Kosut, M. (ed.) *Encyclopedia of gender in media*. Thousand Oaks, CA: SAGE Publications, 25-27

Wehrwein, E.A., Lujan, H.L., DiCarlo, S.E. (2007). Gender differences in learning style preferences among undergraduate physiology students. *Advances in Physiology Education*, 31 (2), 153-157

Wells, J.W. (1989). Sexual language usage in different interpersonal contexts: A comparison of gender and sexual orientation. *Archives of Sexual Behavior*, 18, 127-143

Wenger, E. (2000). *Communities of practice*. New York: Cambridge University Press

West, C., Zimmerman, Don. H. (1987). Doing Gender. *Gender and Society*, 1 (2), 125-151

West, C. (1990). Not just 'doctor's orders: Directive-response sequences in patients' visits to women and men physicians. *Discourse and* Society, 1 (1), 85-112

Westbrook Eakins, B., Eakins, R.G. (1976). Verbal turn-taking exchanges in faculty dialogue. In Dubois, B.L., Crouch, I. (eds.) *The Sociology of the Languages of American Women.* Papers in Southwest English IV. San Antonio, TX: Trinity University, 53-62

Westbrook Eakins, B., Eakins, R.G. (1978). *Sex Differences in Human Communication*. Boston: Houghton Mifflin

Wierzbicka, A. (1991). *Cross-Cultural Pragmatics: The Semantics of Human Interaction*. Berlin: Mouton de Gruyter

Wiley, M.G., Eskilson, A. (1985). Speech style, gender stereotypes, and corporate success: What if women talk more like men? *Sex Roles*, 12, 993-100

Wilson, J. (1990). *Politically Speaking: The pragmatic analysis of political language*. Oxford: Blackwell

Winn, L., Rubin, D. (2001). Enacting gender identity in written discourse: Responding to gender role bidding in personal ads. *Journal of Language and Social Psychology*, 20 (4), 393–418

Wodak, R. (ed.) (1989). *Language, Power and Ideology: Studies in Political Discourse*. Amsterdam/Philadelphia: John Benjamins Publishing Company

Wodak, R. (1991). Turning the Tables: Anti-Semitic Discourse in Post-War Austria. *Discourse and Society*, 2 (1), 65-83

Wodak, R. (ed.) (1997). *Gender and Discourse*. London: SAGE Publications

Wodak, R., Meyer, M. (eds.) (2001). *Methods of Critical Discourse Analysis*. London: Sage Publications

Wodak, R., Koller, V. (eds.) (2008). *Handbook of Communication in the Public Sphere*. Berlin: Walter de Gruyter GmbH & Co.

Wodak, R., Ludwig, C. (ed.) (1999). *Challenges in a changing world: Issues in Critical Discourse Analysis*. Vienna: Passagenverlag

Wodak, R. (2008). Multiple Identities: The Roles of Female Parliamentarians in the EU Parliament. In Holmes, J., Meyerhoff, M. (eds.) *The Handbook of Language and Gender*. Blackwell Publishing, 671-699

Wolbrecht, C., Beckwith, K., Baldez, L. (eds.) (2008). *Political Women and American Democracy*. Cambridge: Cambridge University Press

Wolf, M., Horn, A., Mehl, M., Haug, S., Pennebaker, J.W., Kordy, H. (2008). Computergestützte quantitative Textanalyse: Äquivalenz und Robustheit der deutschen Version des Linguistic Inquiry and Word Count Diagnostica, 285–98

Wolfram, W., Schilling-Estes, N. (1998). *American English: Dialects and variation.* Oxford: Blackwell

Wolfson, I. (2005). Adjective participles as emotion words (MA thesis). *Hamline University Saint Paul, Minnesota*

Xiufang, X. (2013). Gender Differences in Using Language. *Theory and Practice in Language Studies*, 3 (8), 1485-1489

Xu, W.W., Li, L., Stefanone, M.A., Fu, Y. (2014). Does social media users' commenting behavior differ by their local community tie? A computer-assisted linguistic analysis approach. *First Monday*, 19 (1)

Yale, N.R. (2007). *Instant Messaging Communication: A Quantitative Linguistic Analysis*. VDM Verlag Dr. Mueller

Yanagisako, S., Collier, J.F. (1990). The mode of reproduction in anthropology. In Rhode, D. (ed.) (1990). *Theoretical Perspectives on Sexual Difference*. New Haven, CT: Yale University Press, 131-44

Yap, B.W., Sim, C.H. (2011). Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation*, 81 (12), 2141-2155

Yarkoni, T. (2010). Personality in 100,000 Words: A large-scale analysis of personality and word use among bloggers. *Journal of research in personality*, 44, 363-373

Yu, B. (2013). Language and Gender in Congressional speech. *Literary and Linguistic Computing*, 29 (1), 118-132

Yuki, M. (2001). Does 'Women's Language' Really Exist?: A Critical Assessment of Sex Difference Research in Sociolinguistics. *Multicultural Studies*, 1, 141-159

Zahn-Waxler, C., Shirtcliff, E. A., Marceau, K. (2008). Disorders of childhood and adolescence: Gender and psychopathology. *Annual Re-view of Clinical Psychology*, 4, 275-303

Zaini, A. Hazirah, A., Saadiyah, D., Kemboja, I. (2012). Gender Differences In The Language Use Of Malaysian Teen Bloggers. *Journal of Language Studies*, 12 (1), 105-124

Zimmerman, Don H., West, C. (1975). Sex Roles, Interruptions and Silences in Conversation. In Thorne, B., Henly, N. (eds.) *Language and Sex: Difference and Dominance*. Newbury: Rowley Mass, 105-129

Curriculum Vitae

Dragana Božić Lenard was born in Zagreb on 3rd July 1986. In 2008, she received a Bachelor's degree in English language and literature and Croatian language and literature from the Faculty of Philosophy in Osijek. Two years later, she received a Master's degree in English language and literature and Croatian language and literature from the aforementioned faculty (today known as the Faculty of Humanities and Social Sciences) and in December the same year, she enrolled into the doctoral study program in *Linguistics* at the same institution. From September 2010 to September 2012 she worked as a teacher of English and Croatian language and literature at A. G. Matoš High School in Đakovo. Furthermore, since March 2012, she has been working as a teaching assistant at the Faculty of Electrical Engineering, Computer Science and Information Technology in Osijek where she teaches English for Specific Purposes Courses.

Since 2012, she has participated in seven international scientific conferences on translation, phraseology, cognitive linguistics, sociolinguistics, rhetorics and applied linguistics. She presented her doctoral work at the PhD training program at Middlesex University London. Additionally, she attended six international and three domestic conferences. Her research interests include sociolinguistics and rhetorics. From 28th July 2014 to 31st July 2014 she attended Sociolinguistics Summer School at the University College Dublin. From 4th August 2014 to 10th August 2014 she participated in Second Networds Summer School on *Words: structure, meaning, acquisition and processing* at the Norwegian University of Science and Technology in Trondheim, Norway for which she won a scholarship. She was awarded with a research scholarship by the British Scholarship Trust and spent two months (October 2015 – December 2015) conducting her doctoral research at Middlesex University in London.

Dragana Božić Lenard published four scientific papers in international publications. She proofread the university textbook entitled *Simulacija komunikacijskih sustava – Priručnik za laboratorijske vježbe* written by Marijan Herceg, PhD and Tomislav Matić, PhD. In 2016, she participated in three international projects entitled *Senior Academy Project*, *Medico-Tech research network cooperation for public health advancements related to cardio-vascular issues* and *Smart sustainable public buildings with high share of renewable energy sources in Pannonian region* which are currently under review. She also participated in three domestic

and two international workshops. She mentored electrical and computer engineering students who won 2nd and 3rd place in ESP at the international competition of electrical engineering students *Elektrijada* in 2014 and 2015.

Conferences:

- May 22-24, 2015 Faculty of Electronic Engineering, Niš, Serbia
  Second International Conference Synergies of Language Learning, Teaching English for Specific Purposes and New Language Learning Technologies

- July 28-31, 2014 University College Dublin, Ireland
  Sociolinguistics Summer School 5

- May 9-10, 2014 International Burch University, Sarajevo, Bosnia and Herzegovina
  Fourth International Conference on Foreign Language Teaching and Applied Linguistics

- April 23-26, 2014 Brač, Postire
  Second International Conference Days of Ivo Škarić

- March 13, 2013 Faculty of Philosophy, Novi Sad, Serbia
  Second Conference on English Language and Anglophone Literatures Today

- December 1, 2012 Faculty of Philosophy, Novi Sad, Serbia
  First International Interdisciplinary Conference for Young Scholars in Social Sciences and Humanities Contexts

- October, 15-16, 2012 Faculty of Philosophy, Osijek, Croatia
  Current Approaches to English Studies International Conference to Mark the 35th Anniversary of English Studies

Workshops:

- September 22, 2014 Faculty of Humanities and Social Sciences, Osijek
  Types of evaluating students' work in higher education

- March, 28, 2014 Faculty of Teacher Education, Osijek
  Alternative Approaches to Education

- March 17, 2014 Embassy of the United States of America, Osijek
  Academic Writing for English Language Teachers

- April 5, 2013 Faculty of Teacher Education, Osijek
  Languages in Urban Communities - Diversity and Integration for Europe

- October 24, 2012 Faculty of Humanities and Social Sciences, Osijek
  Constructing Exams

PUBLICATIONS

Božić Lenard, Dragana, Liermann-Zeljak, Yvonne, Ferčec, Ivanka (2015). A Corpus-based Linguistic Analysis of Errors in Final Paper Abstracts. In Stojković, Nadežda, Nejković, Valentina (eds.). *Synergies of Language Learning*. Niš: Faculty of Electronic Engineering, 10-19

Liermann-Zeljak, Yvonne, Ferčec, Ivanka, Božić Lenard, Dragana (2015). Employing Experiments and Interactive Materials in Teaching Technical English Vocabulary. In Stojković, Nadežda, Nejković, Valentina (eds.). *Synergies of Language Learning*. Niš: Faculty of Electronic Engineering, 355-363

Božić Lenard, Dragana (2014). A Contrastive Study of Some Lakoff and Johnson's Metaphorical Expressions from Love is a Journey Metaphor and Their Croatian Equivalents. In Akbarov, Azamat (ed.). *Linguistics, Culture and Identity in Foreign Language Education*. Sarajevo: International Burch University, 519-528

Božić Lenard, Dragana (2014). Frazeologizmi u nekim djelima Mire Gavrana. In Živančević Sekeruš, Ivana et al. (eds.). *Prvi međunarodni interdisciplinarni skup mladih znanstvenika društvenih i humanističkih znanosti Konteksti*. Novi Sad: Impressum, 439-470

Božić, Dragana (2010). Proces zaključivanja: kako slušatelji uspijevaju razabrati pravo tumačenje. *Hrvatistika*. 4, 99-109