

5-2018

Nevertheless, She Persisted: A Linguistic Analysis of the Speech of Elizabeth Warren, 2007-2017

Matthew Jennings

Follow this and additional works at: <https://dc.etsu.edu/honors>

 Part of the [Anthropological Linguistics and Sociolinguistics Commons](#), [Applied Linguistics Commons](#), and the [Computational Linguistics Commons](#)

Recommended Citation

Jennings, Matthew, "Nevertheless, She Persisted: A Linguistic Analysis of the Speech of Elizabeth Warren, 2007-2017" (2018).
Undergraduate Honors Theses. Paper 457. <https://dc.etsu.edu/honors/457>

This Honors Thesis - Withheld is brought to you for free and open access by the Student Works at Digital Commons @ East Tennessee State University. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of Digital Commons @ East Tennessee State University. For more information, please contact digilib@etsu.edu.

Nevertheless, She Persisted: A Linguistic Analysis of the Speech of Elizabeth Warren, 2007-2017

By

Matthew W. Jennings

An Undergraduate Thesis Submitted in Partial Fulfillment
of the Requirements for the
University Honors Scholars Program
Honors College
East Tennessee State University

Matthew W. Jennings Date

Dr. Theresa McGarry, Thesis Mentor Date

Dr. C. Wesley Buerkle, Reader Date

Dr. Martha Michieka, Reader Date

Acknowledgements

My thanks go to the Honors College at ETSU and all the honors faculty and staff who have shown me what scholarship truly is. Special thanks in particular are due to Dr. Karen Kornweibel, and Ms. Deanna Bryant Mroziewicz for helping me present this project at the 2018 Purdue Literature, Linguistics, and Second Language Studies Conference.

I would also like to thank the ETSU Office of Research and Sponsored Programs for funding this research and the opportunities this project has provided me, personally and professionally. Thank you to Drs. Martha Michieka and C. Wesley Buerkle as well for reading this project and offering their feedback. Drs. Edith Seier and Jean-Marie Hendrickson were instrumental in helping execute the statistical analysis portion of this project, and thanks go to Lindsey Wright for helping make the R code sing.

Finally, this project would have been impossible without the constant guidance and invaluable feedback of Dr. Theresa McGarry. I sincerely thank you for your infinite patience even when my progress – or lack thereof – was undoubtedly infinitely more frustrating. You have taught me so much this year, and your chickpea croquettes were the best dish at the linguistics minor picnic.

Abstract

A breakout star among American progressives in the recent past, Elizabeth Warren has quickly gone from a law professor to a leading figure in Democratic politics. This paper analyzes Warren's speech from before her time as a political figure to the present using the quantitative textual methodology established by Jones (2016) in order to see if Warren's speech supports Jones's assertion that masculine speech is the language of power. Ratios of feminine to masculine markers ultimately indicate that despite her increasing political sway, Warren's speech becomes increasingly feminine instead. However, despite associations of feminine speech with weakness, Warren's speech scores highly for expertise and confidence as its feminine scores increase. These findings relate to the relevant political context and have implications for presumptions of masculine speech as the standard for political power.

Keywords: LIWC, sociolinguistics, gender indexicality, Elizabeth Warren

1. Introduction

A recent star of American political life, Senator Elizabeth Warren of Massachusetts has gone from a professor of bankruptcy law at Harvard University to one of the modern progressive movement's most visible figures. Indeed, the visibility of Warren among both the American public and her Senate colleagues is exemplified by such monikers as "the Elizabeth Warren phenomenon" being used to single out the Senator's influence and opposition to certain financial reforms (Davis 2015). Despite Warren's immense public and political celebrity, however, Walker (2016) notes that "no published studies of Warren exist in any discipline beyond analysis of her published work in bankruptcy law." Given this dearth of research, this study attempts to contribute to the understanding of Warren's public political persona through an analysis of public speech events.

2. Literature Review

This work is heavily based in the social psychology research of James Pennebaker on the importance and meaning of function words and pronouns in human speech. Most importantly, when Pennebaker (2011) speaks of “masculine” and “feminine” linguistic markers, this is not a normative statement steeped in gendered language stereotypes. Rather, these statements are drawn from the work of Newman et al. (2008), who analyzed a corpus of over 14,000 texts whose author’s sex was known, using the software known as Linguistic Inquiry and Word Count, or LIWC, more commonly (Pennebaker 2015a). In rejecting ideas of essentialism, an important distinction is put forth by Cameron (1998) that what is considered “women’s language” is in fact a symbolic category, while “the language used by women” is an empirical one and thus a measurable phenomenon. The work done by Newman et al. (2008) and Pennebaker (2011) strictly limits itself to “language used by women.”

Quantifying perceived gender differences in language is hardly a novelty in linguistics. Following the assertions of Lakoff (1973) that women use allegedly tentative linguistic features such as tag questions more than men, researchers were quick to test this hypothesis. As described in Weatherall (2002:60), work by McMillan, Clifton, McGrath, and Gale (1977) and Dubois and Crouch (1974) investigated the use of tags in different contexts; McMillan et al. examined their use in a structured experimental setting, while Dubois and Crouch instead recorded spontaneous speech interactions.

Though the researchers found contradictory outcomes – more tag use by women in the structured setting and more tag use by men in the spontaneous recordings – their research spurred further investigation into the claims of the time and highlighted the importance of contextual differences in language use. Deborah Cameron highlights the difference that context

can yield, noting that language “is *radically* contextual” (Cameron 1992, cited in Holmes 2006:17). While quantitative studies can be conducive to ignoring such contextually-bound meanings, Tausczik and Pennebaker (2010) contend that despite the failure of word-counting software to detect irony and nuanced meaning, use of pronouns, for instance, can still be detected and reflect speakers’ attention and focus.

Previous research in the field of analysis of gendered language in political speech in particular has been done by Pearson and Dancey (2011), who found that the content raised by legislators in the House of Representatives differs along gendered lines, with female representatives of either party being more likely to draw attention to women’s issues in speeches. Additionally, Osborn and Mendez’s (2010) results indicated that female Senators are more likely than their male colleagues to make speeches on subjects that directly impact women’s lives, such as family issues and women’s health.

While the previous two studies analyzed content words in speech, Yu (2014) used the word-counting software LIWC to analyze the function words of speeches of legislators in the United States Congress over twenty years. Yu’s findings were largely in line with existing theories, finding that female members consistently used more emotion words of all types across the data, while males displayed a markedly higher preference for articles. Moreover, with regard to pronoun case, Yu’s findings include a higher use of possessive pronouns among women than men. This offers credence to the findings of Newman et al. (2008), which displayed a higher use of social words and references by women, as Yu cites bigrams such as “our community,” “our families,” and “our students,” all of which situate the speaker socially with reference to the group mentioned.

The foundation for the present study is Jennifer Jones's (2016) analysis, "Talk 'Like a Man': The Linguistic Styles of Hillary Clinton, 1992-2013." Following the usage-based approach of Pennebaker and Newman in which "masculine/feminine speech" is measured by the ratio of features appearing more often in the speech of men to that of features appearing more often in the speech of women, Jones took an original corpus of 567 documents and analyzed Clinton's speech with regard to gendered speech markers in order to evaluate how Clinton's gendered self-presentation varied in relation to her standing in the political world.

In order to contextualize Clinton's use of language, Jones analyzed her data in five periods related with Clinton's career and political standing: Clinton's time as First Lady, her first Senate campaign, Clinton's first term as Senator, her first presidential campaign, and her tenure as Secretary of State. Ultimately, Jones's findings were that Clinton did in fact display more masculine speech while in traditionally male settings, such as during her presidential campaign and as Secretary of State, while her language was most feminine as First Lady (Jones, 2016). More specifically, analyzing the trends among the linguistic markers analyzed, Jones concluded that Clinton's speech was not necessarily decreasingly feminine "but it is clear that her speech was increasingly masculine" (2016:633). Given the recent publication of the parent study, no major breakthroughs in the field have been identified since its publication, and thus the theoretical framework remains the same.

Eckert and McConnell-Ginet (1992:463) note that "we speak of practices... as 'gendered' where they enter in some important way into 'gendering' people and their relations," and go on to clarify that "we do not want to suggest that gendered identities and relations have any common core 'fixed' by their... link to reproductive biology." This second claim serves to support a typical Butlerian model wherein gender is not a static social category but instead a

performative construct enacted by actors and thus, flexible (Butler, 1990). Through this lens of gender as performativity, it is reasonable that a politician such as Warren or Clinton can modify her language and, thus, gendered perception in order to make necessary political and social gains.

3. Methodology

The methodology in this study is derived from that used in Jones's (2016) original study. An original corpus of twenty-six interview and debate transcripts was assembled and transcriptions were made or corrected, as needed. Of these transcripts, twenty-three of the speech events are interviews, while the remaining three are debates. As in the parent study, only spontaneous speech events were included in the data set, and thus any speeches or prepared remarks were excluded.

In the selection of Warren as a subject for study, several criteria were taken into consideration. First, the subject had to be a notable female-identifying figure in American politics. By strictly focusing the study on a speaker of American English, there are no concerns of misunderstanding gendered cultural markers in the speech of a speaker of British English, for instance. Moreover, Warren also possessed several personal and ideological characteristics akin to Clinton in the parent study. Most notably, the women are nearly equal in age – Warren is 68 and Clinton, 70 – and have educational backgrounds steeped in law. Furthermore, both women have acknowledged formerly belonging to the Republican party before eventually registering as Democrats instead. In addition to these similarities, Warren's limited political tenure – only six years in elected office – allows for research of much smaller scope than a subject with more longevity, such as Nancy Pelosi.

Much like Jones established five periods for analysis of her data, this study sorts the data into three discrete periods: Warren's life before political involvement (pre-2011), Senate campaign and early political career (2011-2015), and party leadership and progressive icon (2016-2017). These divisions serve to help contextualize Warren's use of language and to

suggest factors that could have driven her choice of linguistic style in some way. Table 1 shows the speech events analyzed in each of the three periods.

Speech Event	Date	WC
2007 Conversations with History Interview	3/8/2007	7252
2008 NPR Morning Edition Interview	12/16/2008	465
2009 Frontline PBS Interview	6/16/2009	4833
2009 Now on PBS Interview	11/13/2009	1795
2009 NPR Fresh Air Interview	5/9/2009	4564
2009 NPR Planet Money Interview	5/8/2009	8373
2010 Charlie Rose Interview	3/4/2010	2434
2010 Tavis Smiley PBS Interview	4/14/2010	1305
<i>2011 WBUR Interview</i>	9/14/2011	999
<i>2012 Senate Debate 1</i>	9/20/2012	4384
<i>2012 Senate Debate 2</i>	10/1/2012	2979
<i>2012 Senate Debate 3</i>	11/3/2012	4072
<i>2012 NPR All Things Considered Interview</i>	9/4/2012	496
<i>2013 Salon Interview</i>	8/22/2013	944
<i>2014 ABC David Muir Interview</i>	4/21/2014	3550
<i>2014 Moyers & Company Interview</i>	9/4/2014	2319
<i>2014 NPR Fresh Air Interview</i>	10/1/2014	1096
<i>2014 WBUR NPR Interview</i>	12/14/2014	611
<i>2015 NPR Politics TPP Interview</i>	5/12/2015	589
2016 Maddow Interview	6/9/2016	2360
2016 Mic Interview	5/12/2016	1641
2017 Axe Files Interview	6/12/2017	7867
2017 Charlie Rose Interview	4/19/2017	4194
2017 Maddow Interview	7/25/2017	922
2017 NPR Book Club Interview	4/18/2017	896
2017 WBUR NPR Interview	9/7/2017	1495
<i>Italics indicates second period; bold indicates third period</i>		

Table 1: List of Speech Events

Since the analysis is based on proportional word frequencies, all transcripts were processed in the previously mentioned software Linguistic Inquiry and Word Count (LIWC) to calculate percentages of total word count in the data. As discussed in the introduction, the categories of “masculine” and “feminine” speech as measured in LIWC in this study are operationally aggregates of words and word categories that appear more often in the speech of

females and males rather than features essentially linked to sex or gender. Accordingly, the two categories are independent, and it is entirely possible for a person's speech to score high or low in both or neither. Thus, "feminine" speech in this study should be taken to mean a high ratio of feminine to masculine indicators, and "masculine speech" is shorthand for a low ratio.

In contrast to Jones's work, the 2015 version of the software was used instead of the version used in the parent study, 2007. Per notes following the software's release, the "cognitive mechanisms" category of the 2007 version was replaced with "cognitive processes," a conceptually similar marker that limits its elements to "true markers of cognitive activity" (Pennebaker et al. 2015b). Beyond this, the software added four new summary variables based on research published by Pennebaker Labs – analytical thinking, clout, authenticity, and emotional tone – which were not available to Jones at the time of her research. These four variables do not count percentages of text as do the other variables, and are instead scored from 0 to 100; the mechanics behind this scoring represent "the only non-transparent dimensions in the... output" (Pennebaker et al. 2015b).

In one departure from Jones's (2016) methodology, instead of merely the aggregate of all masculine and feminine markers being taken, some categories were calculated in order to minimize overlap. For instance, "pronoun" necessarily includes all first-person plural pronouns – so-called "we" words. Thus, so as to not count these pronouns in the wrong category, the percentage of text composed of this marker was subtracted from the broader "pronoun" category. This methodology was also applied in the categories of "negative emotion" and "anger," as all words marked for "anger" shared the "negative emotion" tag, so again, all "anger" words were subtracted from the broader "negative emotion" category in order to more accurately reflect the true values of word usage in Warren's speech.

Replicating Jones's methods, the sums of both the masculine and feminine variables were calculated and the feminine total was divided by the masculine total in order to calculate the feminine/masculine ratio in Warren's speech. These calculated ratios were then plotted over time to observe any general trends in Warren's gendered self-presentation. Debates were not analyzed separately from interview transcripts as the only debate data available was a series of three debates during Warren's campaign for Massachusetts Senator, and there were no differences in content among the three debates so substantial as to warrant further investigation.

As in the parent study, R was used to perform the statistical analysis of the data. Jones's original code was used with minor alterations, namely the deletion of extraneous code relating to Clinton's presidential campaign in 2007-2008 and the replacement of a now-deprecated function.

4. Results and Discussion

In contrast to the data yielded by Jones's work which showed that Clinton developed markedly more masculine tendencies in her speech over time, the findings presented here show that Warren's speech displays a higher feminine/masculine ratio in the first period of analysis, becomes more masculine in the second period, and actually displays a higher feminine/masculine ratio on average in the third period than in both others, despite her having the greatest amount of political involvement at this point. That said, the ratio model results in Table 3 display a general decrease in the feminine/masculine ratio, but without statistical significance. While the generalized linear model measures change on a per-year basis across all years of analysis, Table 2 instead shows Warren's use of each linguistic token delineated by period with the data in a given year weighted by word count.

Table 2: Weighted Average for all Linguistic Markers (percent of word count)

	<i>Examples</i>	<i>Pre-2011</i>	<i>2011-2015</i>	<i>2016-2017</i>
Feminine style				
Pronouns	I, you, they, it	14.30	14.44	16.46
First-person singular	I, me, my	2.52	3.76	3.18
Verbs	Is, do, make	20.13	20.56	20.39
Auxiliary verbs	Have, will, is	11.13	10.95	10.36
Social references	Help, family, we	9.60	9.70	12.23
Positive emotion	Fair, good, love	2.32	2.78	2.77
Negative emotion	Shaken, wrong, worry	1.01	0.70	0.59
Tentative words	May, probably, possible	3.02	2.18	2.03
Cognitive processes	Because, think, bet	11.78	12.1	10.36
Masculine style				
Words > 6 letters		15.25	16.53	13.36
First-person plural	We, our, let's	1.92	2.13	2.39
Articles	A, an, the	7.48	6.88	6.99
Prepositions	In, over, of, by	13.26	13.58	13.84
Anger words	Hate, argue, kill	0.09	0.24	0.70
Swear words	Heck, damn, shit	0.00	0.01	0.03
Feminine/Masculine ratio		2.00	1.96	2.10

Word count	31021	22039	19375
No. documents	8	11	7
Total word count	72435		
Total Documents	26		

Table 3 Generalized linear model results

	Full model	Ratio Model
Intercept	2205.00**** (32.99)	2021.799**** (14.49)
Pronouns	2.40**** (0.31)	
First-person singular	-24.04**** (2.85)	
Verbs	3.48**** (0.33)	
Auxiliary Verbs	10.09**** (1.36)	
Social references	-4.31**** (0.51)	
Positive emotion	-15.38**** (1.81)	
Negative emotion	-24.31**** (2.39)	
Cognitive processes	-1.12** (0.49)	
Tentative words	-6.88**** (0.69)	
<i>Words > 6 letters</i>	5.66**** (0.81)	
<i>First-person plural</i>	-9.06**** (1.38)	
<i>Articles</i>	-42.64**** (5.70)	
<i>Prepositions</i>	2.55**** (0.46)	
<i>Anger words</i>	-17.89**** (2.79)	
<i>Swear words</i>	320.29**** (46.94)	
Feminine/Masculine ratio		-5.21 (7.31)
N	43	10
Log Likelihood	-26.06	-25.43
AIC	84.12	54.87
****p<.0001; ***p<.001; **p<.01; *p<.05; italics indicate masculine variable		

Standard errors are in parentheses. Both models are based on time series data. The full model is a quarterly time series; the ratio model is a yearly time series.

The generalized linear model results as seen in Table 3 show mixed results for Warren's use of masculine and feminine markers across all years of analysis. While her use of only three masculine markers – prepositions, swear words, and words of more than six letters – increases a statistically significant amount, her use of feminine speech markers almost uniformly drops at the same time, with the exception of pronouns, verbs, and auxiliary verbs. All of these results display immense statistical significance, holding true at $p < 0.0001$, while the decrease in cognitive processes is significant at $p < 0.01$. Given the general negative trend in feminine variables and slight positive trend in masculine variables, it appears that Warren's speech was overall decreasingly feminine and slightly more masculine across the entire analysis. While all results are statistically significant, some variables display more pronounced change than others in the model results. Swear words, for instance, did indeed see a significant increase in use, but when consulting the data it is evident that these words were only a negligible fraction of the words Warren used overall.

In the first period, Warren held no elected political office, with her only public office being her tenure as the chair of the Congressional Oversight Panel (COP), established in 2008 to oversee the allocation of funds per the Treasury Department's Trouble Asset Relief Program (TARP). Besides one interview conducted in 2007 about her life and the state of the American middle class, all speech events analyzed here were interviews conducted with Warren giving opinions from the perspective of a financial expert or the chairwoman of the COP. One element of Warren's use of language that stands out and contributes to the comparatively higher feminine/masculine ratio in this period is her use of tentative words, a typically feminine marker. Words labeled as tentative do not necessarily indicate verbal hesitation on the speaker's part (e.g.

uh) but rather express possibility or uncertainty in quantity or actor (i.e. “*Most* of the [money] has already been committed, but Treasury *may* want the second \$350 billion.”)

In fact, the highest value for tentative words in the unweighted data set – 5.38 percent of words in the text – comes in an interview given in 2008. This high use of tentative language during the period is likely indicative of the uncertainty and shaken confidence of the American public in the midst of the Great Recession, a phenomenon explained by Owens and Cook (2013). Despite speaking from a place of power, Warren voices a question on the efficacy of the bailout and its nature, noting that “one of the real *questions* we're asking here is whether *or* not *any* of that money is in *any* way helping end the mortgage crisis,” highlighting the lack of transparency from financial institutions and the Treasury department (emphasis added). In an exchange with Tavis Smiley, for instance, Warren can only speculate on the possibility of a Consumer Financial Protection Agency, saying “I *hope* so, but Tavis, I don't know. Right now this consumer agency is our one *hope to try* to straighten out a consumer credit market that's broken... Whether *or* not it'll go through *or* not... that's up in the air” (emphasis added). While some of this tentative language could be attributed to feminine speech alone, it seems more likely that some such language is the result of the relevant economic and historical factors.

In the second period of analysis, Warren's speech initially displays a fairly similar level of feminine to masculine markers as at the end of first period. Her speech, however, quickly drops in terms of relative femininity, reaching the lowest point in the data set during her third debate with then-Senator Scott Brown during the Massachusetts Senate race, as seen in the dramatic drop in 2012 in the graph below.

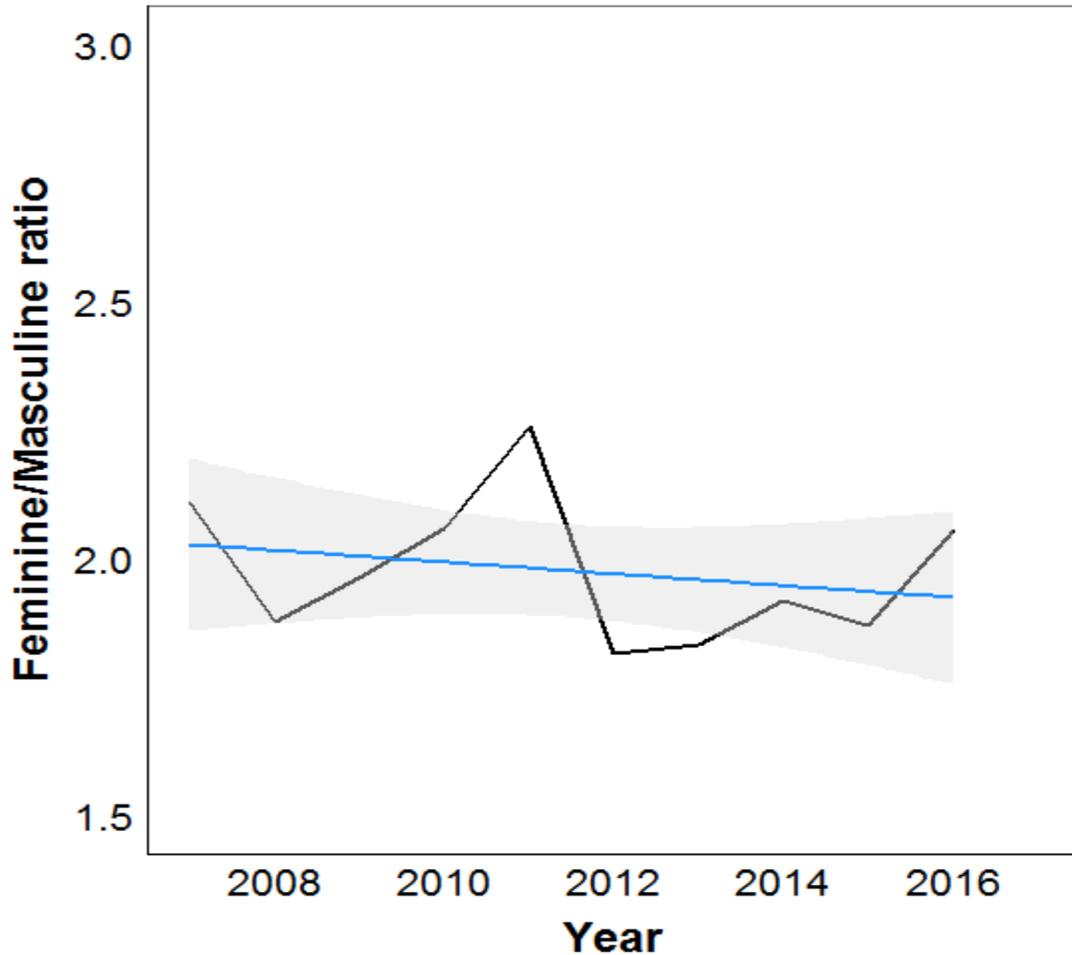


Figure 1: Ratio of Masculine to Feminine Styles over Time

This lowest value does seem in stark contrast to the incredibly high feminine/masculine ratio found in an interview following Warren’s campaign announcement in 2011, but the general trend towards masculine speech at the time is not. Rather, it is consistent with Jones’s own findings in her analyses of Clinton’s Senate races, which argue that female candidates’ self-presentation skews more masculine in order to appear “tough enough” for the position in question. Notably, Warren’s use of big words rises at this time, again peaking in the third debate, where these words represent approximately 21 percent of all words used in the text.

Additionally, following her very positive campaign announcement, Warren begins to display an increase in anger words over the preceding period that only continues to rise going

into the third period. This rise can be attributed to Warren's impassioned style of speaking and preferred way of pitching her economic beliefs to her audience, namely her vivid descriptions of the middle class, a group whom she describes as "hammered" and "cheated" by a wealthy elite (Madigan, 2012). Walker (2016:9) encapsulates the spirit of Warren's arguments, noting that her narrative offers "an idealized image of citizenry made even clearer by the ongoing presence of the villain... [a] powerful, wealthy and corrupt financial sector... referred to with the short hand [sic] of 'Wall Street.'"

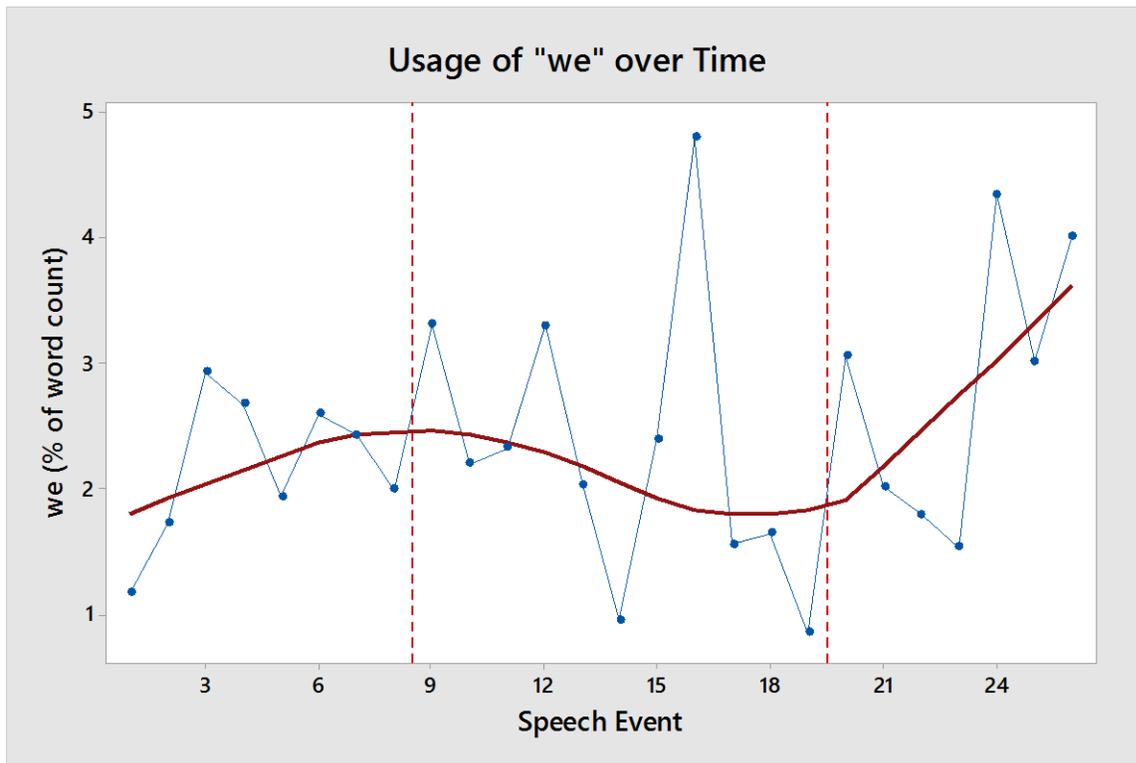
Perhaps the most interesting results can be seen with the data from the third period. Warren, now a Senator in her fourth and fifth years in office, has established herself as an outspoken critic of Donald Trump and the Republican majority in Congress, and for most of 2016 her endorsement for president was highly sought by both Democratic contenders, Bernie Sanders and Hillary Clinton. Presumably, then, Warren's speech should decrease in relative femininity per Jones's findings. Instead, Warren's speech actually displays a higher feminine/masculine ratio on average than in either of the preceding periods. Moreover, this increase occurs despite the fact that Warren's use of "we" words, a masculine speech marker, also sees a sizeable increase in this period, as seen in Figure 2.

This heightened use of "we" can be easily attributed to a change in Warren's messaging strategy, which strongly emphasizes Democratic unity at the time and tends to define her party in contrast to the opposition. In all previous periods, Warren's use of "we" had a much less consistent referent, ranging from speaking on behalf of the members of the COP to aligning herself with Massachusetts voters in order to win over their support. In contrast, Warren's use of "we" forms in the third period is almost uniquely partisan. In one exchange, for instance, Warren says to the interviewer "As a Democrat, one of the things that frustrates me the most is there are

a lot of times *we* [emphasis added] just don't get in the fight. *We* ask pretty please if *we* can have things or *we* make the argument for why it is the best thing to do..." (Wolff & Gnazzo, 2016).

Warren's marked frequent identification with her party at this point in her career is an important implicit marker of her status of authority in the party, as Reicher and Hopkins (2001:386) note that people "will agree with and follow a would-be leader to the extent that the individual is seen as prototypical of the in-group and acts in terms of in-group norms." That said, Warren's shift towards a more collective style of speech could also be in line with the theory put forth by Volden, Wiseman, and Wittmer (2013) that the general tendency of women in minority parties to emphasize cooperation is a useful asset in forming coalitions with members of the majority party. Nonetheless, it is telling of Warren's status as a leader and figurehead of the party that her use of tentative words is lowest in the third period, suggesting certainty in her speech and status.

Simultaneously, despite Warren's increasingly feminine language, the clout dimension, which measures a speaker's perceived confidence and expertise, is consistently at its highest values in this period. Although the exact means by which this variable are calculated are unclear, as the equations are proprietary in LIWC, clout is associated with work done by Kacewicz et al. (2014) on standing in social hierarchies as measured by pronoun use (Pennebaker et al., 2015b).

Figure 2: Usage of "we" over Time (unweighted data)

The interesting question raised therein, then, is why Warren's speech reflects higher confidence and expertise when adhering to party messaging than at previous points, such as her tenure as the chair of the COP. Per the findings of Kacewicz et al. (2014), speakers with higher status more frequently use first-person plural and second-person singular pronouns than those with lower status, who more frequently use first-person singular pronouns. In those transcripts where Warren's clout scores are highest, her speech exhibits two or more of these characteristics, most frequently high "we" and low "I" use. A standard linear correlation calculation between the use of "we" and Warren's clout scores, seen below, yields $r = 0.76$, a result which confirms the strong relationship between the two.

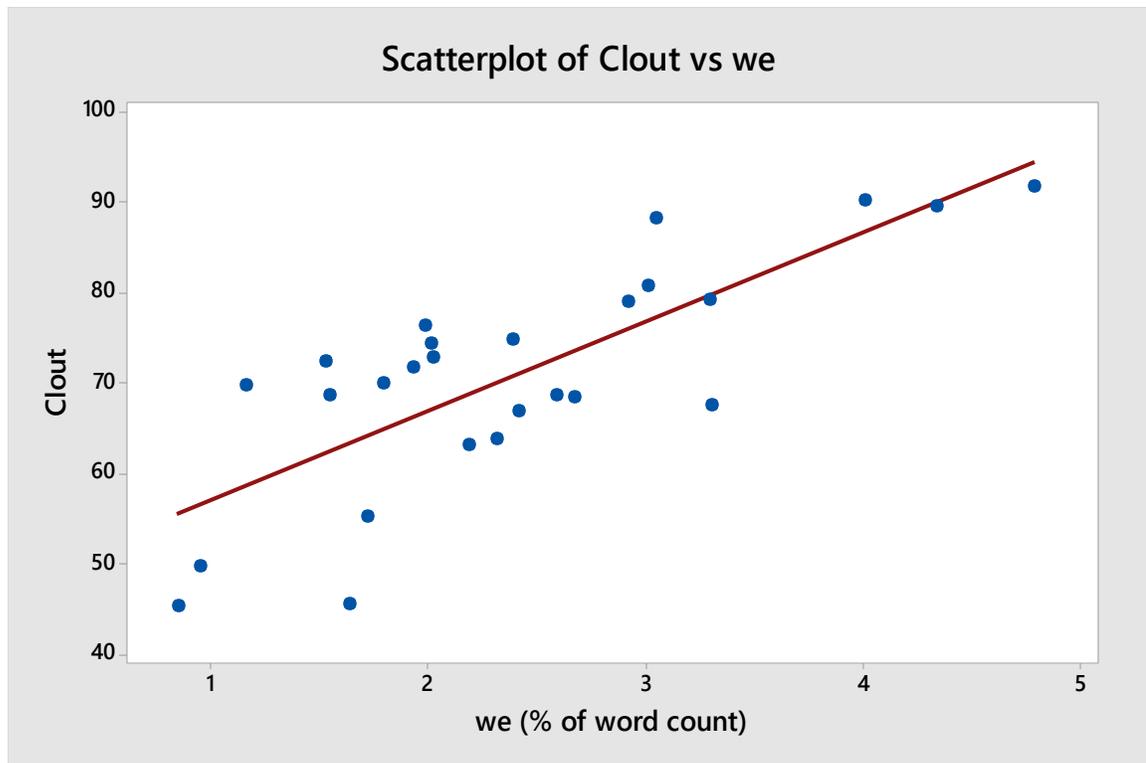


Figure 3: Correlation of We Usage and Clout over Time

One notable exception to the clustering of the peak values for clout in the third period actually comes in the midst of the second period in a 2014 interview. In this interview, Warren’s speech exhibits the same characteristics of her later speech, situating herself clearly as a member of the Democratic party and highlighting the actions taken by her in-group members as beneficial to the American people (i.e. “I think *we* got a lot of Democrats lined up to do this. You know, you ask *us* to get organized, I do want to be clear. *We* got this bill out there. *We* made them vote on it.” [emphasis added]) (Warren, 2014). In many ways, the interview serves as a prototype for her future speech, as this same tendency towards strong in-group identification and the creation of a clear linguistic “other” with repeated references to “the Republicans” and “them” is later reprised in her speech in 2016 and 2017.

Although a rise in “we” use is seen, there is a marked drop of over 3 percent in Warren’s use of words greater than six letters between the second and third periods in particular, another

element contributing to the increase in relatively feminine speech. In contrast from previous years where Warren needed to prove herself against a male competitor and establish herself as credible, Warren speaks with credibility here, speaking on behalf of herself and other Democratic senators. Moreover, instead of attempting to navigate the intricacies of the financial system and crisis, Warren here acts as sort of a party “pitch woman,” attempting to cast Democratic actions and values in the most positive light possible through rigid contrast with “the other guys.” Although the full model results showed a statistically significant overall decrease in “we” use, no analysis was conducted on the change on a per-period basis and this phenomenon bears further scrutiny.

5. Conclusions and Further Research

Just as gender is enacted in all other social contexts, so too is it enacted in the political sphere by those seeking to appeal to certain qualities invoked by a certain linguistic style. This study found that throughout her relatively limited political career, Elizabeth Warren has displayed a noticeable but not statistically significant trend towards masculine speech. Despite the generally masculine trend overall, her speech beginning in 2016-2017 appears to become increasingly feminine despite having the highest public profile and most influence in her career at that point. Despite social and linguistic expectations that men hold power and that their language is the language of authority, Warren's speech shows a high "Clout" score in LIWC, which suggests a high degree of confidence and expertise in her speech, an effect mirrored by her low score for tentative language in this period.

Given Warren's heightened use of "we" variants despite their status as a masculine speech marker, it is worth investigating further if other notable female politicians exercise authority with this same collective strategy that Warren does or if it may merely be a linguistic trait of the language used by those in a minority party. In particular, Eckert and McConnell-Ginet (1992:464) argue for researchers to adopt a community of practice model, which is based in shared practices because it differs from traditional community models "primarily because it is defined simultaneously by its membership and by the practices in which that membership engages." Additionally, Mendelberg and Karpowitz (2016:492) note that "in a world where parties are especially polarized, partisan differences... may swamp gender." Recent polling results from the Pew Research Center (2017:65) confirm a record-breaking partisan divide nationally between Democrats and Republicans, both with respect to political values and views of the opposite party. Thus, given that the United States Congress is a highly ritualized body with

very strong partisan ties that govern interaction, it may bear further scrutiny to see if this shift towards collective language in the form of “we” variants is unique to Warren or if other Democratic Senators – male or female – shift their language accordingly during this same period.

The prevailing limitation of this study is the limited size of the corpus, which contains only twenty-six speech events, a dramatically smaller quantity than Jones’s original corpus of 567. Consequently, some years and periods are more heavily represented than others in the data. While there is at least one speech event represented per year, almost half of the years under analysis – 2007, 2008, 2011, 2013, and 2015 – have only one speech event in total, compared to more robust years, such as 2014 and 2017. Given that this study, following the parent study’s methodology, uses a per-year basis for analysis, these isolated speech events could incorrectly be assumed to be indicative of the whole year. Expanding the corpus to include such events as Senate floor debates or questions in confirmation hearings would help to alleviate such limited representation in the data, in addition to providing other contexts in which one could analyze spontaneous speech.

A limitation discovered in performing the statistical analysis in R was that the data displayed autocorrelation under the Durbin-Watson test at lags 1 and 2. Effectively this means that a different model for analysis may be called for in order to account for this discovery. Moreover, it indicates that the results from the analysis could be misleading, as some factors for which statistical significance was determined may in fact be incorrectly labeled as such.

It is perhaps equally important to note that though the principal object for comparison in this study is Hillary Clinton, especially with the limited corpus size of this study, extrapolating findings too broadly to apply to all female politicians would be a mistake. As Jones (2016:631) herself notes of Clinton, she “has experienced a unique trajectory into politics and, arguably, her

career is not a 'typical' case." That said, by expanding the size of the corpus as discussed above, one could more reasonably make claims about Warren's gendered self-presentation over time than with the current data. Ultimately, the results of the study have interesting implications for understanding female leadership styles and how they exist in relation to conventional understandings of women's speech and leadership as cooperative.

Works Cited

- Butler, J. (1990). *Gender Trouble: Feminism and the Subversion of Identity*. New York: Routledge.
- Cameron, D. (1998). Gender, Language, and Discourse: A Review Essay. *Signs*, 23(4), 945-973.
- Davis, P. (2015, April 22). 'Warren Phenomenon' Causing Reg Relief Delay: House Republican. *American Banker*, 1(61). Retrieved from http://link.galegroup.com.iris.etsu.edu:2048/apps/doc/A410578640/AONE?u=tel_a_etsul&sid=AONE&xid=66088fa7
- Dubois, B. L., & Crouch, I. (1974). The question of tag questions in women's speech: they don't really use more of them, do they? *Language in Society*, 4, 289-294.
- Eckert, P., & McConnell-Ginet, S. "Think Practically and Look Locally: Language and Gender as Community- Based Practice." *Annual Review of Anthropology*, 21, 1992, pp. 461-490.
- Fox, R., & Lawless, J. (2011). Gendered Perceptions and Political Candidacies: A Central Barrier to Women's Equality in Electoral Politics. *American Journal of Political Science*, 55(1), 59-73. Retrieved from <http://www.jstor.org/stable/25766254>
- Holmes, J. (2006). *Gendered Talk at Work*. Cornwall: Blackwell Publishing.
- Jones, J. (2016). Talk "like a man": The linguistic styles of Hillary Clinton, 1992-2013. *Perspectives on Politics*, 14(3), 625-642.
- 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.
- Lakoff, R. (1973). Language and woman's place. *Language in Society*, 2, 45-80.

- Madigan, J. (Director). (10 October 2012). *Massachusetts Senate Debate* [Television broadcast]. Springfield, MA. Cable-Satellite Public Affairs Network.
- McMillan, J.R., Clifton, A.K., McGrath, D., & Gale, L. (1977). Women's language: uncertainty or interpersonal sensitivity and emotionality. *Sex Roles*, 3(6), 545-559.
- Mendelberg, & Karpowitz. (2016). *Women's authority in political decision-making groups*. *The Leadership Quarterly*, 27(3), 487-503.
- Newman, M.L., Groom, C.J., Handelman, L.D., & Pennebaker, J.W. (2008). Gender Differences in Language Use: An Analysis of 14,000 Text Samples, *Discourse Processes*, 45:3, 211-236, DOI: [10.1080/01638530802073712](https://doi.org/10.1080/01638530802073712)
- Osborn, T., & Mendez, J.M. (2010). Speaking as women: Women and Floor Speeches in the Senate. *Journal of Women, Politics, & Policy*, 31(1), 1-21.
- Owens, L. A, & Cook, K. S. (2013). The Effects of Local Economic Conditions on Confidence in Key Institutions and Interpersonal Trust after the Great Recession. *The ANNALS of the American Academy of Political and Social Science*, 650(1), 274-298.
- Pearson, K. & Dancey, L. (2011). Speaking for the Underrepresented in the House of Representatives: Voicing Women's Interests in a Partisan Era. *Politics & Gender*, 7(04), 493-519.
- Pennebaker, J.W. (2011) *The Secret Life of Pronouns: What Our Words Say About Us*. New York, Bloomsbury Press.
- Pennebaker, J.W., Booth, R.J., Boyd, R.L., & Francis, M.E. (2015a). Linguistic Inquiry and Word Count: LIWC2015. Austin, TX: Pennebaker Conglomerates (www.LIWC.net).
- Pennebaker, J.W., Boyd, R.L., Jordan, K., & Blackburn, K. (2015b). The development and psychometric properties of LIWC2015. Austin, TX: University of Texas at Austin.

Pew Research Center. (2017). "The Partisan Divide on Political Values Grows Even Wider."

Retrieved from [http://assets.pewresearch.org/wp-](http://assets.pewresearch.org/wp-content/uploads/sites/5/2017/10/05162647/10-05-2017-Political-landscape-release.pdf)

[content/uploads/sites/5/2017/10/05162647/10-05-2017-Political-landscape-release.pdf](http://assets.pewresearch.org/wp-content/uploads/sites/5/2017/10/05162647/10-05-2017-Political-landscape-release.pdf)

Reicher, S., & Hopkins, N (2001). Psychology and the End of History: A Critique and a Proposal for the Psychology of Social Categorization. *Political Psychology*, 22(2), 383-407.

Tausczik, Y.R., & Pennebaker, J.W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*, 29(1), 24-54. DOI: 10.1177/0261927X09351676

Volden, C. Wiseman, A.E., & Wittmer, D.E. (2013). When are women more effective lawmakers than men? *American Journal of Political Science*, 57(2), 326-341.

Walker, Z.S. (2016). *Populism and Myth in the Rhetoric of Elizabeth Warren*, ProQuest Dissertations and Theses.

Weatherall, A. (2002). *Gender, Language, and Discourse*. New York: Routledge.

Wolff, B., & Gnazzo, C. (Executive Producers). (2016, June 9). *The Rachel Maddow Show* [Television broadcast]. New York, NY. Microsoft News Broadcasting Company.

Yu, B. (2014). Language and gender in congressional speech. *Literary and Linguistic Computing*, 29(1), 118-132. [fqs073]. DOI: 10.1093/lc/fqs073