A Pilot Gender Study of the Danish Parliament Corpus

Dorte Haltrup Hansen, Costanza Navarretta, Lene Offersgaard

University of Copenhagen, Centre for Language Technology

Njalsgade 136, DK-2300 Copenhagen, Denmark

dorteh@hum.ku.dk, costanza@hum.ku.dk, leneo@hum.ku.dk

Abstract

This paper describes a pilot analysis of gender differences in the revised transcripts of speeches from the sittings in the Danish Parliament in the period from 2009 to 2017. Information about the number and duration of the speeches, the gender, age, party, and role in the party was automatically extracted from the transcripts and from other data on the Danish Parliament web site. The analysis shows statistically significant differences in the number and duration of the speeches by male and female MPs, and we also found differences in speech frequency with respect to the age of the MPs. Our analysis confirms previous studies on parliamentary data in other countries showing that the role of the MPs in their party influences their participation in the debates. Furthermore, we found that female ministers were speaking more in the period with a female prime minister than they did under a male prime minister. In the future, we will determine the statistical significance of the various parameters we have analysed in this paper and automatically extract linguistic information which can further determine differences between male and female MPs of different age and from different parties.

Keywords: gender, parliament, CLARIN

1. Introduction

This paper presents a pilot analysis of gender differences in the revised transcripts of speeches from the Danish Parliament over a time span of eight years. The active participation of women in politics is historically relatively new, and women are still underrepresented in Parliament. Furthermore, researchers have found gender specific differences in e.g. the subjects addressed by female and male members of Parliament (MPs) and their speech frequency. This is also the case in Sweden (Bäck et al., 2014) which is a country culturally near to Denmark.

Our study follows this line of research in that we aim to investigate whether it is possible to determine gender differences in the Danish parliamentary speeches by analysing the speeches with respect to their duration as well as the number of the speakers, their age, party and role in the party.

The paper is organized as follows. First, we account for relevant studies that indicate that there are differences in the frequency and subject of female and male MPs. Then, we describe the Danish data and we present a preliminary analysis of the data and a discussion of some of the gender differences that can be seen in them. Finally, we conclude and shortly describe future work.

2. Gender Studies in Political Speeches

The analysis of gender differences in political speeches has involved many aspects, not only the intrinsic characteristics of the political speeches comprising lexical, morphological, syntactic, semantic, pragmatic and rhetorical aspects, but also factors including culture, society, identity definition by the politicians and the media.

Paxton et al. (2007) report different studies on the political participation of women in different countries showing that the percentage of female parliament members is generally low, but it varies by country, being highest in Scandinavia (38% in 2005) and lowest in the Middle East (8% in 2005).

Sivrić and Jurčić (2014) analyse six political speeches in which six Croatian and US male and female politicians address a large audience. They find no differences between the two gender's discourse at the syntactic and semantic level, while they notice that the three female politicians have a tendency to use more implicit meaning and more contrasting constructions than the male politicians. They conclude that some of the language characteristics traditionally attributed to male and female discourse can be found in the political speeches produced by both genders and it depends on how politicians choose to show a particular identity. For example, Hilary Clinton wants to appear strong and thus her discourse has characteristics which usually are described as masculine.

Bäck et al. (2014) investigate the role of gender in the number of speeches held by Swedish politicians in the Parliament. They find that female politicians talk less than male politicians. Furthermore, they investigate other factors than speech frequency such as the domain of expertise, the position in the party of the politicians and their personal background. The data show that women talk as much as men in cases of what they define as "softer" policy areas for which they also often are the responsible in the party. Softer areas comprise culture, education, health and social issues, while harder technical areas comprise energy, finance, macro economy, foreign affairs, national security and transport.

The statistical analysis of gender differences in various types of discourse has shown that there are small, but consistent differences in language use (Newman et al., 2008). The features were found with the text analysis program, Linguistic Inquiry and Word Count, LIWC (Pennbaker et al., 2001) which analyses texts according to 74 linguistic categories. The only consistent effects of gender through the various text types were long words, swear words, articles, pronouns, and social words. Moreover, numerous features are often used in automatic gender identification in different types of data such as parts of the general language British National Corpus (Koppel et al., 2002), scientific articles (Vogel and Jurafskyi, 2012) transcribed telephone conversations (Boulis and Ostendorf, 2005) and computer-mediated written communication, e.g. (Corney et al., 2002; Ljubešić

et al., 2017). Gender identification has been part of the tasks under the Author profiling events organized by PAN[1] where many of the efforts deal with social media data, such as blogs and tweets (Martinc et al., 2017).

The automatic identification of the gender, age (young/old) and political affiliation (left or right wing) of politicians from the edited transcriptions of speeches held by politicians in the Swedish parliament between 2003 and 2010 is addressed by Dahllöf (2012). He extracted the initial 200 words of the speeches and selected the excerpts so that for each task he had at least 77 speakers per group and each speaker had at least 21 contributions. He used features from information extraction and balanced training and test data. Words characterizing each class were used as binary features according to information gain theory. Thus no linguistic analysis of the speeches was performed. The classifier was a support vector machine and ten-fold cross-validation was the evaluation method. Accuracy for gender identification was higher for right-wing politicians than for left-wing ones and was in the range 72.8-80.1%. The fact that short excerpts were used as data restricted the number of possible features to be used for classification.

3. Description of the data

The Danish Parliament Corpus 2009 – 2017 consists of Hansards (transcripts of parliamentary speeches) from the sittings in the Chamber of the Danish Parliament, *Folketinget*. The first editions of these official reports are published immediately after the meetings as pdf-files on the website of the Parliament[2] where it is also possible to see video recordings from the meetings. The final proofread editions are published up till one year after the first ones.

The corpus consists of xml files dumped for us from the Parliament's database by the IT department at the Danish Parliament. Each file contains speeches from one parliamentary year, running from October to June. The xml files do not refer to any xml scheme.

Year	Chairman	Male	Female	*Total*
2009-2010	189,749	3,836,276	2,579,952	*6,416,228*
2010-2011	136,830	2,939,173	1,783,022	*4,722,195*
2011-2012	125,314	3,330,297	1,492,207	*4,822,504*
2012-2013	127,073	3,119,049	1,596,624	*4,715,673*
2013-2014	118,266	2,577,144	1,414,419	*3,991,563*
2014-2015	132,718	2,640,560	1,368,437	*4,008,997*
2015-2016	191,275	3,696,753	1,606,241	*5,302,994*
2016-2017	147,393	2,964,842	1,553,551	*4,518,393*
Total	1,168,618	25,104,0944	13,394,4533	*39,667,1655*

Table 1: The total number of words[3] in the corpus

The files are marked for meetings, speeches, name of speaker, party of speakers and timing of the speeches. A speech is here defined as a single intervention by a MP; it can be a question as well as a longer debate contribution. The date of birth/age and the gender of the speakers for this study have been extracted from additional sources on the web site of the Parliament[4]. The consistent mark-up makes the data a very rich source for sociological and linguistic analysis.

3.1 License and Accessibility

The Danish Parliament Corpus 2009-2017 follows the license for Open Data[5] stating:

The Danish Parliament grants a world-wide, free, non-exclusive and otherwise unrestricted right of use of the data in the Danish Parliament's open data catalogue. The data can be freely:

- copied, distributed and published,
- adapted and combined with other material,
- exploited commercially and non-commercially.

Following the copyright act the speeches can be distributed without the consent of the speaker but only in a way where the author/speaker of each text/speech is clearly stated. Furthermore, the Danish Parliament must be acknowledged as the source. To our understanding this correlates to CLARIN PUB BY or CC-BY.

It is the aim to share the corpus in the CLARIN community as soon as a new repository system is implemented in CLARIN.dk. The version of the corpus used for this study includes meetings until May 4th, 2017, and the reports for the latest parliamentary year have not been published as the final edition. The reports of all other meetings are the final editions. The first version of the Danish Parliament Corpus 2009-2017 will be shared at CLARIN.dk in the same format we received it from the Parliament. Enriched versions of the corpus will be shared later on. For reproduction of the results in this study, the current version of the data will be available upon request to the authors.

3.2 Corpus language characteristics

According to the *Office of the Folketing Hansard* the reports are verbatim (exact transcripts of the speeches), but slightly edited following the guidelines:

- The spoken language is adapted into a colloquial and syntactically coherent written language with a liberal approach to what is deemed correct language.
- The editing is done carefully to ensure that the intentions of the speaker are clear.
- Factual errors and slips of the tongue are corrected.
- The appropriate formal requirement rules are observed.

In this process punctuation marks are added, and smaller corrections are made to make the speeches compliant to Danish syntax for written language, e.g. pauses and hesitations are omitted. Therefore stylistic analysis and investigations which include factors such as "sentence" length must take into account that spoken language has been artificially converted to written language.

---

[1] http://pan.webis.de/

[2] http://www.ft.dk/

[3] The corpus is not tokenized, therefore a word is here defined as a sequence of characters delimited by white space.

[4] http://www.ft.dk/da/medlemmerhttp://www.ft.dk/da/medlemmer

[5] http://www.ft.dk/~/media/sites/ft/pdf/dokumenter/aabne-data/conditions_for_use_-of_the_danish_parliaments-_open_data.ashx?la=da

The Standing Orders of the Danish Parliament state rules for the speeches, which the Speaker (the chairman of the Parliament) enforces during the debates. According to the Standing Orders, the MPs in the Danish Parliament are not allowed to applause or express disapproval during the debates, and when debating or asking questions the MPs must be addressed "hr."(Mr.) or "fru" (Ms.) and their full names, while the ministers must be addressed using the minister title. Furthermore, the informal pronoun "du" (you) should not be used. These rules impact the language, which becomes more polite and respectful than the language used in e.g. interviews or TV debates. One could say that the language is solemn and very formal compared to spontaneous speech.

4. A pilot gender study in the Danish Parliament Corpus

The aim of the present pilot study is to provide an overview of the distributional figures in relation to male and female members of the Danish Parliament. We want to test if there are non-linguistic parameters that might be useful in determining gender differences and in automatic gender classification.

The corpus used in the study is an extract of The Danish Parliament Corpus 2009 – 2017. All comments from the Speaker (chairman or -woman) are omitted in this study, since these will affect the general gender figures.

4.1 Corpus figures in a gender perspective

The figures shown below reflect the fact that we are looking for clues to characterize the speeches of female and male MPs, focusing on how the female MPs are represented in relation to their male colleagues.

Election	Male	Female	Total	% Female
Election 2007	113	66	179	36.9
Election 2011	111	68	179	38.0
Election 2015	113	66	179	36.9
Avg.	112	67	179	37.24

Table 2: Election results from the period covered by the corpus[6]

In compliance with the observations of Paxton et al. (2007), the number of elected female members is 37-38% and quite stable over the years.

Speeches	Male	Female	Total	% Female
Speeches	119,441	62,751	182,192	34.4
Speaking time, hours	2484.59	1324.78	3809.37	34.8
Words	25,104,094	13,394,453	38,498,547	34.8

Table 3: The total number of speeches' time and words in the corpus

We also investigated whether there is a statistically significant difference in the time the female and male MPs speak in the corpus (the MPs who actually speak) and the difference is significant (two-tailed unpaired t-test, df

=1768, t = 2.4194 and p < 0.0156. Applying Welsh version of t-test, df = 1524, t = 2.486 and p = 0.13).

Table 4 compares the number of male and female MPs in the Danish Parliament in the corpus with respect to their age.

Age / MPs	Male	Female	Total	% Female
20-29	18	23	41	56.1
30-39	62	57	119	47.9
40-49	86	64	150	42.7
50-59	87	37	124	29.8
60-69	64	21	85	24.7
70-79	12	6	18	33.3
Total	329	208	537	38.7

Table 4: Members grouped by age

Table 4 shows that young women are very strongly represented in the Danish parliament in the period 2009 - 2017, and that women over 50, on the other hand, are very poorly represented.[7]

In Table 5 the number of speeches in the different age groups are given.

Age /speech	Male	Female	Total	% Female
20-29	4,839	5279	10,118	52.2
30-39	27,813	24,666	52,479	47.0
40-49	37,923	20,755	58,678	35.4
50-59	26,771	7,536	34,307	22.0
60-69	19,804	3,497	23,301	15.0
70-79	2,291	1,018	3,309	30.8
Total	119,441	62,751	182,192	34.4

Table 5: Speeches grouped by age

The table shows that women in general held fewer speeches than it would be expected from their number in the parliament. This is especially the case for female MPs in the age group 50-69.

In Table 6, the distribution of male and female MPs with respect to their party is given. Left-wing parties are marked with red, while right-wing parties are marked with blue. The central party *The Social Liberal Party* (marked with the letter RV) is in white.

MPs /party:	Male	Female	Total	% Female
EL	19	11	30	36.7
SF	21	24	45	53.3
ALT	6	4	10	40.0
S	87	50	137	36.5
Left-wing				40.1
RV	3	5	8	62.5
KF	4	2	6	33.3
V	89	38	127	29.9
DF	51	31	82	37.8
LA	19	8	27	29.6
Right-wing				32.6

Table 6: MPs in right and left-wing parties

---

[6] Extracted from https://www.dst.dk/ and http://www.ft.dk

[7] The % difference in Table 1 and 3 are caused by members stopping, taking leave of absence and new taking over. The substitutes are not necessary of the same gender.

The proportion of female MPs in the left-wing parties is higher than in the right-wing parties, and especially in *The Socialist People's Party* (SF) [8].

In Table 7, we show the number of speeches held by the MPs of the different parties.

The table indicates that the left-wing female MPs give fewer speeches than the female MPs from the right-wing-parties compared to their seats in Parliament. The women who speak less frequently compared to their seats in Parliament are those who belong to the most left-wing party, *The Red-Green Alliance* (EL) who have 36.7% seats but only give 27.7% speeches compared to their male colleagues. The female MPs of *The Social Liberal Party* (RV) give 54.5% speeches vs. 62.5% seats and the right-wing party *The Liberal Party* (V) give 27.2% vs. 33.3 % seats.

Speeches/ party:	Male	Female	*Total*	% Female
EL	15,872	5,916	*21,788*	27.2
SF	7,344	7,967	*15,311*	52.0
ALT	2,276	1,508	*3,784*	39.9
S	24,684	13,537	*38,221*	35.4
Left-wing				**36.6**
RV	5,986	7,180	*13,166*	54.5
V	26,949	11,968	*38,917*	27.2
KF	8,528	4,495	*13,023*	52.0
DF	17,839	6,504	*24,343*	39.9
LA	8,738	2,802	*11,540*	35.4
Right-wing				**33.8**

Table 7: Speeches held by MP in right and left-wing parties

The final observation leads us to investigate the relationship between the number of speeches and the role of the MPs.

In the Danish Parliament most of the bills are introduced by ministers but to some extent also by spokespersons from parties outside the government. When a bill is debated, the proposer - typically the minister - gives the first speech, followed by the spokespersons from the other parties who give their opinion or ask questions to the proposer. Therefore, the role as minister or spokesperson gives more opportunities to speak.

In Table 8 the number of male and female ministers from the various parties is shown.

MPs/ ministers:	Male	Female	*Total*	% Female
SF	4	2	*6*	33.3
S	12	7	*19*	36.8
Left-wing				**35,1**
RV	6	4	*10*	40
V	14	10	*24*	41.7
KF	5	6	*11*	54.55
LA	3	3	*6*	50
Right-wing				**48.8**

Table 8: Ministers from the different parties

[8] See all parties: http://www.thedanishparliament.dk/Members/Members_in_party_groups.aspx

Compared to the number of elected female MPs, relatively few left-wing female MPs have been ministers. The left – wing won one election while the right-wing won two elections in the period covered by the corpus (2007 and 2015).

In Table 9, we show the number of speeches held by female ministers belonging to the various parties. Although the right-wing parties have the biggest number and percentage of female ministers, these ministers do not speak as many times as their female colleagues from the left-wing parties. Especially the female MPs from *The Social Democratic Party* (S) speak much more than expected. This might be related to the period in which there was a female prime minister, which will be investigated in the following section.

Speeches/ ministers:	Male	Female	*Total*	% Female
SF	1,148	345	*1,493*	23.1
S	5,249	5,202	*10,451*	49.8
Left-wing				**36.5**
RV	1,707	1,002	*2,709*	37,0
V	8,444	5,065	*13,509*	37.5
KF	2,214	1,553	*3,767*	41.2
LA	777	488	*1,265*	38.6
Right-wing				**36.4**

Table 9: Speeches by the ministers

4.2 Prime ministers, ministers and speeches

We have further analysed two election periods: 2011-2015 with a female prime minister and 2015-2017 with a male prime minister. Although the first period is longer than the second, we can compare the percentage distribution of ministers and the duration of their speeches.

Table 10 contains the number of ministers and of MPs in the entire corpus as well as the hours they have spoken and the number of speeches they have held. The same data for the two election periods with a female and a male prime minister, respectively, are given in Tables 11 and 12.

Entire corpus (2009-2017)	Male	Female	*Total*	% Female
Ministers	44	32	*76*	42.1
MPs elected	112	67	*179*	37.4
Ministers, hours	460.05	317.58	*777.63*	40.8
Other MPs, hours	2024.54	1007.20	*3031.74*	33.2
Speeches, ministers	18,762	13,167	*31,929*	41.3
Speeches, other MPs	100,679	49,584	*150,263*	33.0

Table 10: Number of ministers, speaking time, and number of speeches in the entire corpus

Female prime minister (2011-2015)	Male	Female	Total	% Female
Ministers	23	14	*37*	37.8
MPs elected	111	68	*179*	38.0
Ministers, hours	195.76	158.47	*354.23*	44.7
Other MPs, hours	944.23	414.28	*1358.51*	30.5
Speeches, ministers	8,104	6,592	*14,696*	44.9
Speeches, other MPs	45,558	20,407	*65,965*	30.9

Table 11: Number of ministers, speaking time, and number of speeches under a female prime minister

Male prime minister (2015-2017)	Male	Female	Total	% Female
Ministers	16	10	*26*	38.5
MPs elected	113	66	*179*	36.9
Ministers, hours	128.39	55.21	*183.60*	30.1
Other MPs, hours	535.39	271.69	*807.28*	33.7
Speeches, ministers	5,942	2,658	*8,600*	30.9
Speeches, other MPs	29,138	13,802	*42,940*	32.1

Table 12: Number of ministers, speaking time, and number of speeches under a male prime minister

The comparison of Tables 11 and 12 shows a clear tendency: although the percentage of female ministers under the male prime minister was higher, the female ministers under the female prime minister spoke much more and for a longer time.

As there has only been one female prime minister in Denmark, we only have data from one election period, but the comparison of data from that election period with data from the current election period shows that the female ministers under a female prime minister speak more than the female ministers under a male prime minister.

In the next section we will consider the speaking time of all MPs, focusing on the members of the standing committees.

4.3   Committee members and subject areas

Not only the ministers play a central role in the parliament, so do the members of the standing committees who are spokespersons for the parties they represent.

In general, 62.7% of the male MPs and 67.1% of the female MPs are spokesmen in one or more committees in the current election period. Since the information about this role is not present in the corpus, it is very difficult to extract the exact distribution of speaking time for each committee policy area. It is however possible to extract the total amount of speaking time for the members of the committees.

Table 13 shows the amount of speaking time for male and female ministers, spokesmen and non-spokesmen MPs in the election period 2015-2017.

Speaking time (2015 – 2017)	Male	Female	Total	% Female
Ministers, hours	128.39	55.21	183.60	30.1
Spokesmen, hours	456.69	226.15	*682.84*	33.1
Other MPs, hours	78.70	45.54	*124.24*	36.7

Table 13: The total amount of speaking time for all MPs in the period 2015 – 2017

Table 13 shows that the majority of the debates are done by the spokespersons, followed by the ministers. The regular MPs do not have much speaking time. This follows to a great extent the Standing Orders of the Danish Parliament which states that the time allotted to speakers in general debates is a maximum of 20 minutes for ministers, a maximum of 10 minutes for spokesmen and a maximum of 5 minutes for other members.

The Danish parliament has 25 standing committees. These have been summed up to the 13 groups listed below, where e.g. *Foreign affairs* also comprises development aid, *Agriculture* comprises food, fishing and agriculture and *Social* comprises health, children, impaired and elderly people.

Since the corpus does not contain information about the spokesman role, we had to retrieve this information from the parliament website. For the same reason, we only have access to information about the present sitting and therefore only take into account the period 2015 -2017. In Table 14, the number of male and female spokesmen in the various policy areas is given. The table is sorted based on the fraction of female speakers.

Policy areas	Male	Female	Total	% Female
Foreign affairs	25	4	29	13.8
Domestic affairs	16	4	20	20.0
Labour and industry	12	3	15	20.0
Economy	13	4	17	23.5
Infrastructure	25	8	33	24.2
Defence	9	3	12	25.0
Food and agriculture	14	5	19	26.3
Church and culture	10	6	16	37.5
Environment	10	7	17	41.2
Law	7	5	12	41.7
Immigration	7	8	15	53.3
Social	23	28	51	54.9
Education	8	12	20	60.0

Table 14: The percentage of female spokespersons of the standing committees in the period 2015 - 2017

The data in the table clearly show that women are strongly represented in the "softer" subject areas, like *immigration, social* and *education,* whereas the more technical areas are dominated by men. The observations support the findings of Bäck et al. (2014).

5.   Conclusion

In this paper, we have presented a first study of the participation of female and male members of the Danish Parliament in the period 2009-2017 by looking at their number, their age, their party, and their role comparing these factors to the frequency and duration of the various politicians' speeches. More specifically, our study shows that the number of female MPs under 29 is larger than the number of male MPs from the same age group and that in general women speak less frequently and for a shorter time than male MPs in proportion to their seats in Parliament. The difference in speaking time between female and male MPs is statistically significant. The data also show that women belonging to a left-wing party speak less frequently than women from the right-left party compared to their seats in Parliament. The data also indicate that ministers and spokesmen speak more frequently than simple MPs and that female ministers

under a male prime minister give fewer speeches than female ministers under a female prime minister even though their percentages in the two periods are similar.

We also found that there were relatively more male spokesmen than female ones in the period covered by the corpus. The Danish data seem to confirm the findings of Bäck et al. (2014) in the Swedish data that female MPs often spoke about "softer" political areas for which they were responsible, while male MPs spoke about "harder" subjects.

In the future, we will combine the various features we have looked at in this study, and calculate their statistical significance. We have started to do an automatic extraction of the subject areas addressed by the MPs in order to automatically determine i.a. differences in the speeches with respect to the gender, party, role and age of the speaker. Moreover, we are looking at linguistic features that can be used to determine the gender, age and role of MPs automatically in line with the work of e.g Pennbaker et al. (2001) and Dahllöf (2012). Furthermore, since both audio- and video-recordings are available for the most recent Parliament debates, we will also address these multimodal data (speech and gestures) in the future.

6.   Acknowledgements

7.   References

Bäck, H., Debus, M. and Müller, J. (2014). Who Takes the Parliamentary Floor? The Role of Gender in Speech-making in the Swedish Riksdag. *Political Research Quarterly* 67(3), pp. 504–18.

Boulis, C. and Ostendorf, M. A. (2005). Quantitative Analysis of Lexical Differences Between Genders in Telephone Conversations. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics* , 2005 New Brunswick, NJ, Association for Computational Linguistics, pp. 435-442.

Corney, M., de Vel, O., Anderson, A. and Mohay G. (2002). Gender-preferential Text Mining of E-mail Discourse. In *Proceedings of the 18th Annual Computer Security Applications Conference*, Alamitos, IEEE Computer Society, pp. 282-289.

Dahllöf, M. (2012). Automatic prediction of gender, political affiliation, and age in Swedish politicians from the wording of their speeches - A comparative study of classifiability. *Literary and Linguistic Computing*, Volume 27, Issue 2, 1 June 2012, pp. 139-153.

Koppel, M., Argamon, S. and Shimoni, A. (2002). Automatically categorizing written texts by author gender. *Literary and Linguistic Computing* 17(4), 401–412.

Ljubešić, N., Fišer, D. and Erjavec, T. (2017). Language-independent Gender Prediction on Twitter. In *Proceedings of the Second Workshop on Natural Language Processing and Computational Social Science*, pp. 1–6.

Martinc, M., Škrjanec, I., Zupan, K. and Pollak, S. (2017). Author Profiling - Gender and Language Variety Prediction—*Notebook for PAN at CLEF 2017*. In Linda Cappellato, Nicola Ferro, Lorraine Goeuriot, and Thomas Mandl, editors, CLEF 2017 Evaluation Labs and Workshop – Working Notes Papers, pp. 11-14.

Newman, M. L., Groom, C. J., Handelman, L. D. and Pennebaker J. W. (2008). Gender Differences in Language Use: An Analysis of 14,000 Text Samples. *Discourse Processes*. 45 (3), pp. 211-236.

Paxton, P., Kunovich, S. and Hughes, M. M. (2007). Gender in Politics. In *Annual Review of Sociology*, Vol. 33, pp. 263–284.

Pennebaker, J., Francis, M. E. and Booth, R. J. (2001). Linguistic Inquiry and Word Count (LIWC): LIWC2001.

Sivrić, M. and Jurčić, D. (2014). Gender Differences in Political Discourse. In *Journal of Foreign Language Teaching and Applied Linguistics*, pp. 173-185.

Vogel, A. and Jurafsky, D. (2012). He said, she said: Gender in the ACL anthology. In *Proceedings of the ACL-2012 Special Workshop on Rediscovering 50 Years of Discoveries*, pp. 33–41.