Detecting Gender Bias in the Coverage of Politicians in Irish Newspapers Using Automated Text Classification

Thesis submitted for the degree of Doctor of Philosophy

By

Susan Leavy

School of Computer Science and Statistics

Trinity College Dublin

April 2014

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work. I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

Susan Leavy

Abstract

This thesis presents an investigation of whether there is evidence of gender bias in the coverage of politicians in Irish newspapers. Text-analysis techniques including natural language processing and machine learning are used to analyse newspaper content and highlighting differences in the coverage of male and female politicians. These differences are then analysed to identify any evidence of gender bias.

The corpus analysed in this research contained Irish newspaper articles featuring male and female politicians over a 15 year period between 1997 and 2011 and coverage of candidates in the 2011 Irish Presidential Election. Machine learning algorithms were used to identify differences in the coverage of male and female politicians. These differences were analysed for evidence of gender bias and where appropriate, the context in which these differences occurred in the corpus were investigated. A broad range of text classification experiments were explored and the best approaches to using text classification to identify gender bias in text were identified.

This research found that the best approaches to text classification involved using a support vector machine learning algorithm along with a binary representation of the features of the articles. The features that were found to be most useful in identifying gender bias were single words, adjectives, verbs and domain specific

lexicons.

Evidence of gender bias was identified in the newspaper coverage of politicians in Ireland. Some of these findings align with findings of previous research on the portrayal of female politicians in the media. Other findings are new and highlight some concerns regarding how female politicians are evaluated by the media and stereotypical portrayals regarding personal and policy issues.

This research presents a corpus-driven approach to analysing gender bias in the coverage of politicians in the media. They show how machine learning algorithms can be used to highlight differences in the coverage of male and female politicians and how these patterns can then be analysed to identify gender bias. This approach facilitates large scale analysis of texts thus addressing some critiques regarding the generalisability of some studies in gender and language that focus on smaller samples of text (Baker, 2014; Neuendorf, 2011). These findings contribute to the current body of literature on the representation of female politicians in the media and present new topics for the analysis of gender bias in media content.

This research presents a broad exploration of patterns of difference in how male and female politicians are represented in the media. These findings could form the basis for further in-depth studies of gender bias. In evaluating a range of approaches to using text classification for identifying gender bias, this research presents a new methodology for researchers interested in examining the representation of female politicians in the media.

Acknowledgements

The first person I would like to thank for helping me through the PhD process is my husband, Stephen. I would never have done it without his unfailing belief in me.

I would also like to thank my sons, Félim and Cathal, who were my motivation to complete the PhD. My parents Deirdre and Jimmy Leavy have been constant in their support and encouragement not only for the past four years, but throughout my professional and academic career. I would also like thank Margaret Donnelly for being an endless source of help and support.

Professor Eileen Drew has been an exceptional supervisor. She has been a mentor, a friend and a highly committed supervisor. I would also like to thank my second supervisor Professor Carl Vogel, who gave me invaluable guidance and support.

This PhD work was only possible due to the Trinity Postgraduate Research Studentship. I would like to thank Trinity College Dublin for providing this support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

"Prejudice is often more powerful than fact" (Cameron, 1992, p.5s)

1.1 Background

Throughout the world there are vast differences in the distribution of political power among men and women. The global average for the percentage of women in national parliaments is 21 percent (IPU, 2013). In areas such as education, health and the economy global gender gaps are narrowing (WEF, 2013). However, despite increases in the numbers of women in politics over the past decade, the rate of increase has slowed, prompting some researchers to suggest that a plateau has been reached (Lavery, 2013).

Ireland has one of the worst records of women's political participation in the EU. It is ranked 24th in the EU and 91st in the world in terms of the number of women elected to lower houses of parliament (IPU, 2013). Presently 16 percent of elected members of the Irish parliament or Dáil are women. In the past decade the representation of women in parliaments in Europe, excluding Nordic countries, has grown by 7.3 percent. However, in Ireland this figure grew by only 2.4 percent (IPU, 2013).

The causes of such persistently low levels of women in politics are complex and wide-ranging. Sex role socialisation has been identified as resulting in lower levels of political ambition and confidence among women (Fox and Lawless, 2004). Despite evidence to the contrary, party leaders in the UK were found to believe that the likelihood of electorial success for female candidates is lower than male candidates (Sanbonmatsu, 2006). Gender bias in media coverage of female politicians has also been found to reduce the chances of electorial success for women (Heldman et al., 2005) and deter women from entering politics (Fox and Lawless, 2004).

Despite Ireland's low level of participation of women in politics, virtually no research has examined gender bias in Irish media (Galligan, 1998). To date, only one study by Ahmad et al. (2011) systematically examined gender bias in Irish political news. Brandenburg (2005) attributed this, in part, to a general consensus that the Irish media are predominantly unbiased. This thesis addresses the lack of analysis of gender bias in the Irish media.

1.2 Analysing Gender Bias in Media Content

Bias in the media has been defined by McQuail (1992, p.191) as "a systematic tendency to favour (in outcome) one side or position over another". In research exploring the coverage of politicians in the media, gender bias is understood as an expression of preference for male politicians, reflecting an underlying belief that politics is, and should be, primarily a male concern (Norris, 1997; Ross et al., 2013).

Research since the 1980s has analysed media content for evidence of gender bias, particularly focusing on election coverage (Banwart et al., 2003; Bystrom et al., 2001; Gidengil and Everitt, 2003; Kahn, 1996; Miller and Peake, 2013; Niven and Zilber, 2001; Norris, 1997; Trimble et al., 2013). These studies investigating gender bias in media predominantly rely on content analysis of newspaper articles or transcripts from television programmes (Macnamara, 2003). Content analysis is a methodology which allows for the analysis of meaning in text and is being increasingly used in social science research (Macnamara, 2003; Neuendorf, 2011; Riffe and Freitag, 1997), driven by the increasing availability of media texts in digital form.

The most common method of content analysis involves a quantitative analysis of recorded communication measuring variables in text (Neuendorf, 2002). Variables are predefined by the researcher and the text is typically manually coded according to a set of guidelines. Computer software is often used to support this research process (de Graaf and van der Vossen, 2013). Content analysis has frequently been criticised for requiring subjective judgement on the part of the researcher. A second criticism relates to the labour intensiveness of the coding process. This can limit the volume of text that can be analysed affecting the generalisability of the research (Lombard et al., 2002; Neuendorf, 2011; Riffe and Freitag, 1997).

New methods of analysing content are being developed that address many of the criticisms of content analysis, including those mentioned above, drawing upon computational sciences, artificial intelligence and computational linguistics to analyse text. The opportunities presented by using these techniques to analyse large sets of data and make inferences about human behaviour, has prompted researchers to define a new discipline called computational social science (Lazer et al., 2009). Some examples of research which has exploited this include the extraction of opinions from media text to predict the outcomes of elections (Chiang and Knight, 2011), whether a political party is in government or opposition (Hirst et al., 2014), sentiment in parliamentary proceedings (Grijzenhout et al., 2014) and the direction of market prices (Busse and Clifton Green, 2002; Tetlock, 2007).

In gender research, despite the advantages that computational approaches to analysing content present, most content analysis is still conducted manually (Neuendorf, 2011). The use of computational methods to examine large quantities of text is known as Corpus Linguistics. In an analysis of 63 articles published in the journal *Language and Gender* between 2007 and 2012, Baker (2014) found that only 4 papers used a corpus linguistics approach. This thesis addresses the issue identified by (Baker, 2014) concerning the lack of research in the discipline of gender and language which uses Corpus Linguistic approach.

This thesis uses text-analysis techniques employing machine learning and natural language processing to analyse gender bias in the coverage of politicians in Irish newspapers. The machine learning algorithm identified differences in the coverage of male and female politicians based on information extracted from the content of newspaper articles. Using natural language processing along with machine learning to differentiate between categories of texts is commonly termed text classification.

The newspaper articles analysed included 15 years of newspaper coverage of ministerial level Irish politicians between 1997 and 2011 as well as the coverage of the 2011 Irish Presidential election. A range of natural language processing techniques were applied to extract information from these texts and machine learning algorithms were then used to build models which would automatically identify whether articles featured male or female politicians. These models were then analysed to assess whether the differences in the coverage featuring male and female politicians that were identified by the machine learning algorithms were attributable to gender bias.

Studies concerning the subject of gender and language which use text classification techniques have predominantly focused on identifying whether there are differences in the speech or writing of men and women (Argamon et al., 2003; Boulis and Ostendorf, 2005; Corney et al., 2002; Otterbacher, 2010). These studies are positioned within the disciplines of computational linguistics and aim to develop classifiers that can differentiate between communication authored by men and women. This thesis uses the techniques developed in such studies but focuses less on achieving high classification accuracy and more on interpreting the patterns in the texts identified by the classifiers. Using text classification to explore differences between corpora in this way has been used in political science to examine ideological differences evident in political texts such as speeches (Diermeier et al., 2012; Yu et al., 2008). This thesis explores how text classification can be used in gender and language research by examining the nature of differences in media coverage of male and female politicians in Ireland to identify those that are a result of gender bias.

1.3 Research Objectives

The research has two main objectives. First it seeks to identify whether gender bias exists in the coverage of politicians in the Irish media. The second aim is to explore how machine learning and natural language processing can be used to establish the existence of gender bias in media content. These objectives form the basis for two sets of research questions presented below.

In order to analyse media content for gender bias, it is important first to identify which approaches to natural language processing and machine learning are useful in identifying gender bias in text. A broad range of approaches where implemented, based on those that were shown to be successful in related studies. Based on results of these, those applicable to the subject area of gender bias were identified.

There are three key parts of the automatic text classification process, for which there are a range of options. Numerous machine learning algorithms can be used to build classifiers based on information extracted from texts. The information extracted depends on the subject being studied. For example, extracting sentiment from texts is useful in predicting market prices (Busse and Clifton Green, 2002; Tetlock, 2007). Machine learning algorithms require that texts are represented in numerical form and the methodology used to do this can also affect the results of the classification. Addressing the first research question concerning how text classification techniques can best be used to explore gender bias in text involves identifying the best approach in these three aspects of the text classification process.

Models were developed by the machine learning algorithms to identify and classify articles according to whether they feature male or female politicians. These models are based on patterns in the texts uncovered by the algorithms which differentiate between coverage featuring male and female politicians. The differences identified by the models are then analysed and those which suggest they may be attributable to gender bias are identified. These patterns are then analysed further to examine whether they are a result of gender bias.

The research questions to be examined by this study are:

1. How can techniques from text classification be used to explore differences in the coverage of male and female politicians in order to identify gender bias?
   (a) What machine learning algorithm is suitable for identifying gender bias?
   (b) What approach to feature extraction is most informative in identifying gender bias in text?
   (c) Which is the optimal approach to representing features in text classification experiments to identify gender bias?
2. What differences in Irish newspaper coverage of male and female politicians indicate gender bias?

1.4 Contribution of the Research

This thesis addresses a gap in the research on women's participation in politics in Ireland. Given that the representation of women in Irish politics is low in comparison with other countries (IPU, 2013), it is important that the causes of this are investigated. Gender bias in media coverage of politicians is one of many known causes of lower levels of participation of women in politics (Fox and Lawless, 2004; Heldman et al., 2005). However, in Ireland, little research has examined whether gender bias exists in Irish coverage of politicians (Galligan, 1998). This research addresses that issue.

This research introduces a new method of analysing media content to identify gender bias in text. Studies in the discipline of gender and language have been slow to utilise computational approaches to analyzing content (Baker, 2014; Neuendorf, 2011). This thesis addresses the issue by identifying how natural language processing and machine learning can best be utilized to analyse texts and automatically identify gender bias.

1.5 Thesis Structure

This section presents an overview of how this thesis is structured. Chapter two provides a review of literature related to this thesis. It begins by presenting academic literature outlining how gender bias is manifested in language. Research which documents evidence of gender bias in media coverage of female politicians is then reviewed. Studies which use text classification techniques are also reviewed in order to identify the best approach to detecting differences in the coverage of male and female politicians.

The research methodology is outlined in chapter three. The philosophical stance of the research is explained and issues emerging from the adoption of computational approaches to content analysis are discussed. A description of the data used in the research is also presented.

Chapter four presents the main research findings. The results of a broad range of experiments, using text classification to analyse gender bias in text are presented and evaluated to identify the most useful methods of detecting gender bias. Analysis of the differences identified by the machine learning algorithm in order to identify those that are a result of gender bias is presented.

The final chapter provides a synthesis of the outcomes of the analysis to extract conclusions in relation to the question of whether there is gender bias in the

Irish media's coverage of politicians and how automatic text classification can be used to identify it.

Chapter 2

Literature Review

2.1 Introduction

This thesis encompasses research from a broad range of disciplines including gender and language, political science research on the representation of politicians in the media and automated text classification. This chapter reviews the relevant literature in each of these areas. Feminist and linguistic theories pertaining to the relationship between gender, language and society which form the theoretical foundation of the research theme are discussed. Research which identifies how gender bias can be manifested in the way language is used is reviewed. Studies of the coverage of male and female politicians in the media are critiqued and the main themes differentiating coverage of male and female politicians are highlighted. As this research introduces text classification techniques to the study of gender bias in language, it is important to review the latest research which uses these techniques to analyse language.

2.2 Theoretical Foundations

This section introduces the feminist and linguistic theories pertaining to gender and language that form the foundation of this research. It describes feminist perspectives on the relationship between language, gender ideology and society. It illustrates how research in linguistics has incorporated feminist perspectives, forming the discipline of feminist linguistics. The thesis builds upon approaches to analysing language developed in feminist linguistics, particularly focusing on feminist stylistics.

2.2.1 Feminist Theory on Language and Gender

*"There is no neutral discourse: whenever we speak we have to choose between different systems of meaning, different sets of values" (Coates, 2004*a*, p.302)*

Language and its relationship with gender in society has always been at the centre of feminist theory. Women's freedom to express themselves through speech and writing was a crucial part of the feminist movement in the early 1900's (Stanton, 1990; Woolf, 2008). Feminist research from the 1960's analysed how women were often represented as passive, emotional and irrational in literature (Millett, 1970) and how the media presented idealised portrayals of femininity (Friedan, 1963).

Influenced by poststructuralist thought, feminist research in the later part of the 20th century questioned the role of language in the perpetuation of gender ideologies in society that had no basis in the biological difference between the sexes (Butler, 1990). Spender (1980) and Kramarae (1981) argued that language itself was created by the male dominant group and embodied a male world view. These works identified ways in which gender ideology is embedded in language and how this can influence people's conceptions of women and expectations of behaviour associated with gender.

How language influences the behaviour of men and women was empirically demonstrated in a series of experiments in the 1970s. Bem and Bem (1973) showed how language in job advertisements affected the gender profile of subsequent applicants. Martyna (1978) showed how generic pronouns are not perceived by women as inclusive of men and women. The findings of such research led to the development of guidelines to avoid the use of gender biased or sexist language (Casey and Kate, 1980; Swift and Miller, 1981). For example, the publisher McGraw-Hill adopted editorial guidelines to avoid sexist language (Moulton et al., 1978).

The feminist standpoint that differences between genders in society are to a large extent a result of a gender ideology embedded in language and that gender ideology could be transformed by reforming language (Ehrlich and King, 1994), forms the central motivation for this research. In relation to differences in behaviour of male and female politicians, language used in the media has been shown to have an effect on their behaviour and outcomes of elections (Fox and Lawless, 2004; Heldman et al., 2005). Hence, the objective of this research is to determine whether gender ideology that is negative towards female politicians is evident in the media. Where it is evident, identifying the nature of it allows for language reform to be undertaken in order to address potential negative outcomes for women in politics. This aligns with the political goals of the strand of feminist research concerned with addressing gender-based inequalities in society though the way in which language is used (Litosseliti, 2006).

2.2.2 Linguistic Research in Language and Gender

While feminist theory proposed that language and how people use it is imbued with gender ideologies that disadvantage women in society, research in linguistics in the early 20th century adopted an opposing viewpoint. It inferred that the differences in how men and women communicate are based on biological differences between the sexes. For example, Jespersen (1922) proposed that women had smaller vocabularies, used hyperbole and had a tendency not to finish sentences, which was seen as a result of their "natural" compulsion to speak before they thought.

> "The two languages are two strata of the same language, one higher, more solemn, stiff and archaic, and another lower, more natural and familiar, and this easy, or perhaps we should say slipshod, style is the only one recognised for ordinary women." (Jespersen, 1922, p.242)

Trudgill (1971) found that women were inclined to pronounce their words more correctly than men, reflecting the fact that they were more status-conscious than men. The emergence of sociolinguistics in the 1960s, questioned linguistic theories that attributed differences in how men and women communicate to biological differences between the sexes. Instead, it viewed language as both reflecting and shaping norms pertaining to gender in society. Sociolinguists viewed language as being shaped by an ideological framework (Burton, 1982). Experience was therefore seen as being expressed and perceived through pre-defined conceptual categories embedded in language.

Lakoff (1975) was one of the first linguists to explore the relationship between gender, language and society from a feminist viewpoint. Lakoff (1975) identified a series of differences in the way men and women used language. Some of the differences identified are as follows:

Use of hedges

Super-politeness

Use of empty adjectives

Apologies

Less frequent speech

Tag questions

Direct quotations

Overuse of qualifiers

Lack of a sense of humour

Use more intensifiers

Lakoff (1975) explained the differences in speech as the result of the socialisation of women into subordinate positions in society and the cultural context of the communication. Lakoff's (1975) interpretation of the differences were critiqued for characterising women's speech as lacking in comparison with that of men (Cameron, 1992; Coates, 2004*a*) and being largely intuitive and anecdotal (Wardhaugh, 1998). Cameron (1992) criticised the methods of linguistics for perpetuating a gender ideology that saw women as inferior by using practices such as categorising feminine words as [-male].

2.2.3 Feminist Linguistics

In addressing the need for a more critical approach to the broader social context of communication linguistic research adopted the goals of feminist research in gender and language and formed the discipline of feminist linguistics (Litosseliti, 2006). Research in feminist linguistics aims to identify how language creates and sustains gender divisions and inequalities in society. Analysis of gender and language moved beyond identifying and analysing differences in how men and women communicate, and gender bias in language as an abstract system (Sunderland and Litosseliti, 2002). Research became more concerned with analysing how gender ideologies influence the choices people make in representing men and women in language. This thesis builds upon this area of research.

Research in feminist linguistics which analysed gender-related meanings imparted through language use was influenced by critical linguistics, founded by Halliday (1973) and Fowler et al. (1979), which introduced aspects of critical theory to linguistic analysis. Halliday (1973) highlighted the choices people make in generating language and how these choices convey meaning. He considered the act of choosing specific words to convey meaning a form of social action. Fowler et al. (1979, p.196) analysed how language encodes ideology concerning power in order "to establish and maintain [people] in economically convenient roles and statuses". The approach Halliday (1973) outlined involved systematically identifying the choices people make in generating language in order to analyse the ideological implications of this. Halliday's (1973) approach allowed for consideration of context and subjectivity, while also utilising an empirical approach to language analysis.

Feminist stylistics is a systematic approach to analysing how women are represented in language which builds upon Halliday's (1973) critical approach to language analysis. It involves an empirical analysis of the language in a text to identify patterns of representation that are indicative of gender bias (Page, 2010). Mills (2002)'s objective was to identify features of newspaper and literature text which indicated gender bias. To achieve this Mills (2002) developed a framework for analysing stylistic attributes of texts for gender bias.

The linguistic features of this gender bias were isolated to develop a toolkit by which gender bias could be identified in media texts. Mills's (2002) framework was the first to systematically analyse language for gender bias. However, given that an objective of the stylistic approach to language analysis is to systematise the analysis of language, this approach is more suitable to computational implementation than one which relies on researcher interpretation as evidenced by the recent growth in the area of computational stylistics (Argamon, Whitelaw, Chase, Hota, Garg and Levitan, 2007).

2.3 Identifying Gender Bias in Language

This section outlines research in or closely related to the discipline of feminist linguistics that identified patterns in how language is used, in relation to women, which could be attributed to gender bias. The categorisation of the literature is based on the feminist stylistic framework for identifying gender bias in text (Mills, 2002; Page, 2010). These approaches are particularly relevant to this thesis in that they outline ways in which text can be systematically analysed for gender bias. The themes indicating gender bias identified in the literature inform both the approach to feature extraction in the text classification process (Section 3.8.3 for description of this process) and the interpretation of the findings of this research. These are presented in Table 2.1 and described subsequently.

Theme
Use of gender specific terms
Describing women as girls
Use of honorific titles revealing marital status
Naming order - placing males before females
Negative Portrayals of Women
Use of stereotypes in descriptions of women
Use of Metaphors presenting negative portrayals of women
Narrative structure implying negative assumptions regarding women
Frequency of mentions of women

Table 2.1: Attributes of Language Use Indicating Gender Bias

2.3.1 Gender specific terms

Researchers have pointed to differences in how categories of men and women are named as evidence of gender bias (Holmes and Sigley, 2002; Mills, 2002; Romaine, 1999; Sigley and Holmes, 2002). Romaine (1999) highlights how the term 'career woman' is often used to describe professional women but 'career men' is not used to describe professional men. Similarly 'family man' is a term used to describe men but 'family woman' is not used in relation to women. Such expressions "count as two strikes against women. On the one hand, they suggest that as women, females can't be real professionals, and on the other, they suggest that as professionals, females can't be real women" (Romaine, 1999, p.131). Mills (2002) also identified the use of terms such as 'single mum', 'working mother', 'career woman' and 'unmarried mother' in the media revealing social preconceptions of women. Occupational terms used in relation to women were found to be often premodified by a gender specification such as 'female lawyer' and 'woman judge', identifying their existence as counter to societal expectations (Sigley and Holmes, 2002).

Gender neutral terms for positions and occupations such as chair, chairperson or police officer are used increasingly to refer to both men and women (Baker, 2008; Holmes and Sigley, 2002; Pauwels, 2000; Romaine, 2001). This was illustrated in Romaine's (2001) analysis of the analysis of the British National Corpus (1995), which contained the gender neutral terms spokesperson(s) and chairperson(s) where these did not occur at all in the corpora from the 1960s and 1970s (Table 2.2). However, in an analysis of US newspapers, Ehrlich and King (1994) found that this increase is due to gender neutral terms being consistently used to refer to women but not to men. "Rather than ridding the language of a masculine generic, then, the introduction of neutral generic forms such as chairperson or chair has led to a gender-based distinction between forms such as chairperson or chair (used to designate females) vs. chairman (used to designate males)" (Ehrlich and King, 1994, p.63).

Term	**Frequency**
chairman/men	12,052
chairwoman/women	71
chairperson/s	166
madam chairman	37
chairlady	1
spokesman/men	4,233
spokeswoman/women	618
spokesperson/s	276

Table 2.2: Word Frequencies of Different Forms of the Positions of Chair and Spokesperson in British National Corpus -1995 (Romaine 2001)

Use of diminutives such as '-ess' and '-ette' to indicate female occupational roles has also been found to trivialise women's professional status assuming that they are dependent on, or derived from that of, men (Lei, 2006). Uses of such terms include poetess, usherette and sculptress. However, use of such terms has been found to be in decline since the 1960s (Sigley and Holmes, 2002).

Another gender bias in language which is in decline is the use of androcentric terms such as 'he', 'him', 'man' and 'mankind' to refer to both men and women (Spender, 1980). While research has shown this is in decline (Baker, 2008; Holmes and Sigley, 2002), there is a persistent trend that in discussions of groups where there is an expectation that the individuals in question are more likely be of a particular gender, that gender will be used to refer to both men and women in the group (Sunderland and Litosseliti, 2002). For example in reference to a group of fire-fighters individuals are more likely to be referred to in male terms. "Might man-words die but with an accompanying increase in the use of sex-specific man-words in, say, examples and verbal illustrations - perhaps very stereotyped ones? " (Sunderland and Litosseliti, 2002, p.520).

2.3.2 Describing Women as Girls

Women are described as girls more often than men are described as boys (Romaine, 1999). In an analysis of the use of the terms girl(s) and boy(s) in a corpus of text of British, American and New Zealand English, (Sigley and Holmes, 2002) found that the term girl is 3 times more likely than the term boy to refer to an adult. They found that women were described as girls in in order to characterise them as immature, innocent, of youthful appearance, subordinate status, emotionally weak or financially dependent. Using 'girl' in conjunction with occupations also reduced the status of the jobs (Sigley and Holmes, 2002). Baker (2010) found that the terms 'boy' and 'girl' occurred with equal frequency in an analysis of examples of British English texts including literature and media content from 2006. However, the term 'girl' referred to women 52 percent of the time while 'boy' referred to a man 28 percent of the time. Baker (2010) also found that 'girl' was used in more disparaging and sexual contexts than 'boy'.

2.3.3 Honorific Titles

Honorific titles such as 'Miss.' and 'Mrs.' reflect marital status of women but the male equivalent does not reflect the marital status of men showing how women are portrayed in terms of their relationships to others (Mills, 2002; Swift and Miller, 1981). In the 1970s the term 'Ms.' was introduced as an equivalent for 'Mr.' to address this asymmetry in the meaning of the titles for both genders.

Baker (2010) found that the uses of 'Ms.' increased from 2.7 percent in 1991 to 10.9 percent in 2006 in an analysis of a corpus of British English. A study of business language from business related internet sites by Fuertes-Olivera (2007) found that 90.7 percent of addresses to women used 'Ms.' as the title. The internet sites included in the study were from in Britain, USA, Pakistan, Netherlands, Belgium, Switzerland and Hong Kong. While there may be differences in the frequency of 'Ms.' across countries, this was not examined in the paper. Further studies have shown that in New Zealand newspapers, while there is an increase in the use of the term 'Ms.', it is being used to replace 'Miss.' but not 'Mrs.' (Holmes, 1994). As will be discussed in the next section, this form of gender bias has also been identified in relation to female politicians, where their marital status is considered more important than that of men.

2.3.4 Naming Order

It is a convention in English when naming pairs of men and women, to name the male first (eg. son and daughter, husband and wife, Mr. and Mrs.) (Mills, 2002). This practice demonstrates a bias which places more importance on men (Mills, 2002; Mollin, 2012; Vefali and Erdentug, 2010). "Women are more often named in second position, and this is perceived as indicating a power-related social order" (Motschenbacher, 2013, p.216). This practice of naming the most powerful of a pair first is evidenced by the following terms: 'master/servant', 'teacher/pupil' and 'doctor/nurse' (Mollin, 2012).

Motschenbacher (2013) conducted a comprehensive study of the ordering of personal binomials in the British National Corpus. Examples of word pairs studied included 'man/woman', 'girl/boy', nobility titles such as 'lady/gentleman', 'princess/prince', kingship terms such as 'wife/husband', occupations such as 'actress/actor' and pronouns such as ' he/she'. While there were variances in the order of naming the pairs (Motschenbacher, 2013) found that gender was the most important influencing factor regarding which of the pair of terms was named first.

2.3.5 Negative Portrayals of Women

In an analysis of adjectives used to describe men and women in British newspapers, Caldas-Coulthard and Moon (2010) found that men were more frequently described in terms of their behaviour while women were described in terms of their appearance and sexuality. In an analysis of the context of the use of the term 'girl', research has shown that girls and boys are represented differently with girls being more objectified (Taylor, 2013) and portrayed in more negative contexts (Baker, 2010). The negative connotations associated with the portrayal of women and girls was also shown by Romaine (2000) in the analysis of mentions of the terms in the British National Corpus (Table 2.3).

Descriptions	woman	girl	man	boy
blonde	25	28	1	1
frigid	2	0	0	0
honest	11	2	68	1
hysterical	14	1	0	0
intelligent	17	9	44	3
loose	3	2	1	1
neurotic	2	2	2	0
silly	16	35	0	10
ugly	6	4	0	0

Table 2.3: Descriptions of Men, Women, Girls and Boys in the British National Corpus (Romaine, 2000, p.110)

2.3.6 Use of Stereotypes in Portrayals of Women

The use of stereotypes in the portrayal of women has been studied particularly in relation to the media. Stereotypes have been identified in the portrayal of women related to sexuality (Masser et al., 2010) and beauty (Frith et al., 2005). Pearce (2008, p.19), in a study of how men and women were represented in the British National Corpus analysed the collocates of man and woman. Collocates refer to words that appear in a corpus alongside a given word with statistically significant frequency. He found striking stereotypical portrayals of men and women. An example of the stereotypical portrayal of women in relation to personality is presented in Table 2.4. The adjectives in bold are those which were used exclusively in reference to men or women.

	Man	Woman
+ extraversion	eminent, garrulous, **gregarious**, influential, powerful vivacious	**bossy, chattering, gossiping**, promiscuous, spirited,
− extraversion	ascetic, **cautious**, humble, quiet, reserved, sensitive, shy, unassuming,	**submissive**
+ agreeableness	**affable**, **amiable,amiable-looking**, **avuncular**, charming, considerate, content, contented, courteous, **funniest**, funny, generous, **good-natured**,happier, happiest, happy, happy, jolly, **jovial**, kind, kindest, kindly, **likeable**, merry, **mild-mannered**,nice, nicest, **personable**, polite	glad, grateful
− agreeableness	arrogant, **cruel**, cruel, dangerous, **dour**, embittered, evil, hateful, impossible, **indifferent**, **insensitive**, **insufferable**, nasty, proud, sinful, unwilling, violent	**bitchy**
+ conscientiousness	**braver**, conscientious, earnest, faithful, generous, good, **humane**, **law-worthy**, loyal, **patient**, prudent, reasonable, **sincere**, thoughtful, **tolerant**, **trusted**, **trustworthy**, **truthful**, upright, **upstanding**	
− conscientiousness		
+ neuroticism	**anxious, insane**, mad, scared, sensitive, upset	dissatisfied, distraught, **hysterical**, mad, neurotic, silly, **weeping**
− neuroticism	sane	satisfied
+ openness to experience	**astute**, brilliant, clever, gifted, learned, rational, reasonable, **scholarly**, **self-educated**, shrewd, thoughtful, wise, wiser	resourceful, strong-minded
− openness to experience	**ignorant**, retarded	daft, dependent, dumb

Table 2.4: Personality Adjectives Most Associated with Men and Women; Adjectives Occurring Exclusively With Men or Women are in Boldface (Pearce, 2008, p.14)

(Mills, 2002) found differences in verbs used by men and women which impart agency to men and passivity to women. This is supported by Pearce (2008) who highlighted the stereotypical portrayals of women in relation to the actions they performed and those that they were subject to.

2.3.7 Metaphor

In their work on metaphor in language, Lakoff and Johnson (1980, p.159) show how metaphors in discourse "play a central role in the construction of social and political reality" and develop cognitive metaphor analysis in order to systematically identify the concepts underlying the use of metaphor in discourse. Metaphor has also been found to reflect underlying societal and political ideologies Goatly (2007). In an endeavour to systematise the identification of metaphors in text, Steen (1999) outlined five steps that could be applied to linguistic features of a text to identify whether their use was metaphorical or not. (Semino, 2008) also explored how metaphor is manifested in language, and how it is used to convey meaning in discourse.

The use of metaphor has been identified as one way in which societal gender ideology is manifested through language (Hines, 1999; Koller, 2004; Mills, 2002). How metaphor is used in political discourse has also been found to affect women's political engagement with politics (Ahrens, 2009; Radic-Bojanc and Silaski, 2012).

Research on the kind of metaphors used to portray men and women has identified a gender bias whereby those metaphors used to portray women are "more prolific and more derogatory than those used exclusively for men" (Holmes, 2003, 115). In an analysis of a corpus of feature articles from business magazines, (Koller, 2004) identified the centrality of metaphors of war in business discourse. Business women were also portrayed as 'warriors' or 'fighters' more often than men. These discursive practices are attributed by Koller (2004) to hegemonic masculinity reinforcing the subordinate status of women.

Mills (2002) noted that by virtue of the fact that they are so commonly used and accepted by society, commonly used metaphors are imbued with the perception of inherent wisdom and truth. For instance, she pointed out that male sexuality is often spoken of in terms of predatory behaviour and women's sexuality is spoken of in terms of heat or lack thereof.

2.3.8 Narrative Structure

Structural aspects of texts can play an important role in imparting an ideological stance concerning gender (Mills, 2002). Mills (2002) generates a set of questions about a text to ascertain whether a text is gender biased. These questions don't engage with the content of the narrative itself and so don't involve narrative analysis, but focus more on analysing certain attributes of the narrative style. The following is a sample of some of these questions:

- Are the sentences short or long? Are they composed of subordinate clauses or co-ordinate clauses? How are they linked?
- Are the verbs used concerned with action, with doing, or with reporting feelings or emotions?
- Is the narration first person or third person, and is that narration from the point of view of a character within the text or is it narrated by a voice external to the text?
- Does the text address you as a male or female? What sort of male or female? White or black? Straight or gay?

- Does the text assume that you will agree with certain of its statements? Are these statements about gender?

The benefits of the approach set out by Mills (2002) is that the answers to many of these questions could be extracted computationally from the texts. For example the length of sentences can easily be calculated and included as features of texts for machine learning experiments.

2.3.9 Frequency of Mentions of Women

Gender bias in the media expressed through absence of women was analysed by Romaine (2001) who examined the entire British National Corpus and found that 'Mr.' occurs more often than 'Mrs.', 'Miss.' and 'Ms.' combined. Using the British National Corpus, Pearce (2008) found that the frequency by which women were mentioned in the text more than doubled since the 1960s resulting in women being mentioned more often than men. However, mentions of individual men, as distinct from mentions of men as a general category, occurred twice as often as mentions of individual women. In an analysis of business literature, Fuertes-Olivera (2007) also found that mentions of men occurred 10 times more often than mentions of women and that of the total mentions of terms of terms of address (including Mr., Ms., Mrs., and Miss.), 93.5 percent were occurrences of 'Mr.'.

Ali et al. (2010) also used numbers of references to women in the media as a measure of gender bias in a large-scale study of 3.5 million articles from British newspapers. Automated methods were devised to identify the gender of subjects referenced in newspaper articles. It was found that men were referenced in 49 percent of top stories while women were referenced in 18 percent. The largest gender divide was found in the sport sections followed by business and politics.

These studies show how relatively straight-forward frequency counts of the mentions of women can be a powerful indicator of gender bias. However, these techniques do not account for the context in which women are discussed.

2.3.10 Summary

The literature presented in this section set out ways in which gender bias is evident in how women are represented in language. How language is used in various forms including speech, media coverage and literature is analysed. The kinds of bias range from how women are described and named to the kinds of discourse they are associated as well as their absence from certain discourses. These attributes of how women are portrayed in language were used in this research to inform both the interpretation of the results of the research and the kinds of features of text that are analysed for evidence of gender bias. The next section reviews literature specifically related to the coverage of female politicians in the media. Many of the same kinds of bias are evident in the portrayal of female politicians.

2.4 Gender Bias in Media Coverage of Politicians

This section reviews academic research that analysed media coverage of politicians. It presents research which identified gender bias in the coverage of politicians in the media and the effect this can have on politicians. A focus of this review is on identifying how gender bias is manifest in media coverage of politicians since these findings inform what features of texts are analysed in this corpus. How bias is manifested in particular attributes of language such as certain verb types, adjectives or other linguistic phenomena is highlighted as these informed the features of the text that are analysed in this research.

Forms of bias in news media were categorised by D'Alessio and Allen (2000) as coverage, agenda and statement bias. The first refers to the volume of coverage afforded to a particular subject. If a media publication is biased towards a particular party, ideology or category of politician, this will be reflected in the volume of coverage devoted to it. Agenda bias refers to the policy focus of the media. Preference for one political party for example could be demonstrated by selection of stories that accord with that party's agenda. Statement bias relates to explicit preference expressed in media discussion about politicians or political parties. In a study of gender bias in the representation of female politicians, Gidengil and Everitt (1999) outlined a similar typology of bias. Invisibility/visibility referred to the quantity of coverage afforded to male and female politicians. The kind of coverage indicated the substance of the coverage and gendered mediation concerned the gendered framing of news stories. These typologies of bias in the media are used to categorise the findings of research on gender bias in the media.

2.4.1 Quantity of Coverage

The first category Gidengil and Everitt (1999) outlined was visibility/invisibility relating to the quantity of coverage afforded to male and female politicians in the mass media. Gender bias was attributed to a text where there were significant differences in the amount of coverage afforded to male and female politicians that could not be explained by factors other than the politician's gender. Research in the 1980s and the early 1990s noted significant differences in the volume of coverage between male and female politicians (Heldman et al., 2005; Kahn and Goldenberg, 1991). However research focusing on more recent political campaigns has shown that this difference has diminished and in some cases been eliminated (Bystrom et al., 2001; Miller et al., 2010).

Kahn and Goldenberg (1991) in a content analysis of US newspaper coverage of US Senate elections between 1982 and 1986 found that male candidates received significantly more coverage than female candidates. Heldman et al.'s (2005) analysis of newspaper coverage of Elizabeth Dole's campaign for the Republican Party presidential nomination, showed that she received significantly less coverage than the male candidates. This was surprising since she was second in the opinion polls at the time. Their analysis of the coverage led them to assert that had her press coverage been less gender biased she may have been able to stay in the election longer and that the differences in coverage she received were "seemingly inexplicable except in terms of gender bias" (Heldman et al., 2005, p.331).

Norris (1997) found decreasing or no difference in the volume of coverage attributed to male and female politicians and proposed that differences still exist in the coverage of male and female politicians but that these concern the quality and framing of the stories. Bystrom et al. (2001) conducted a content analysis of 707 articles from newspapers in the US which featured coverage of Senate and Gubernatorial elections. Findings of the study indicated that female candidates featured in more of the articles (96.6 percent) than male candidates did (74.8 percent). Hillary Clinton, in her 2008 campaign for the democratic presidential nomination, received an equal amount or more coverage than the male nominees (Miller et al., 2010). Later studies on gender bias in the media have focused more on analysing the content and framing of political news stories rather than the quantity of the coverage.

2.4.2 Substance of Coverage

The second category of bias outlined by Gidengil and Everitt (1999) relates to the substance of media coverage coverage. While the volume of coverage of female politicians has increased relative to that of male politicians, the quality of that coverage is often found to be more favourable to male politicians. These differences are summarised in Table 2.5. As a primary motivation for this research relates to improving the approach to conducting analysis of gender bias in the media, the review of literature in this area describes the approaches taken to analyse the newspaper content in these studies.

Family Relationships and Roles

Family relationships and roles within the family are recurring themes in analysis of media coverage of female politicians. The prevalence of this issue in relation to female politicians has been found to be a result of stereotypical views of women (White, 2012). In a study of gender in the media it was found that only 1 percent of the coverage of male politicians and 17 percent of coverage of female politicians mentioned their family status (Spears et al., 2000).

Analysis of newspaper coverage of Sara Palin during the 2008 campaign for the Republican Party Vice-President nomination, demonstrates this emphasis on family status. Coverage of Palin featured comments on her status as a mother four times as often as her counterpart Biden and her marital status was mentioned twice as often (Miller and Peake, 2013). The research was conducted using content analysis of a sample of 2,592 articles from 17

Theme	Summary of Findings
Family relationships and roles	Female politician's family and personal relationships emphasised more than male politicians
Personal issues over policy	More policy coverage related to male politicians. Coverage of female policy is more personal or focused on health and social issues
Emotional Control	Female politicians portrayed less in control of emotion
Style of Politics	Female politician's speech and debate portrayed less positively than male politicians
Stereotypes	Negative stereotypes used in descriptions of female politicians
Personalised Coverage	Greater focus on female politician's appearance, style and personal information than male politicians
Focus on Gender	Unnecessary focus on gender in coverage of female politicians
Masculine Narrative/Methaphors	Framing narratives using stereotypically male analogies or metaphors such as war or combative sport

Table 2.5: Summary of Categories of Gender Bias in Coverage of Female Politicians

newspaper sources during the 2008 campaign. As part of the content analysis process, a coding guide was developed to identify certain features of the newspaper articles. A team of ten researchers then manually coded the articles to extract features indicating whether the candidate was mentioned in the headline, the tone/sentiment expressed towards the candidate in the headline and the article (negative/neutral/positive), number of sentences devoted to them, mentions of gender, clothing/appearance, and family status. This study demonstrates the cost of the manual coding of data for content analysis studies. Also, while extensive training was given to the coders, estimates of the tone of articles does involve subjective judgement on the part of the researcher. Inter-coder reliability tests were conducted and a moderate to high accuracy was found in coding. This shows that if codes are defined to include information on latent meaning in text, this can reduce the reliability of content analysis.

A trend has emerged in political discourse in the media which is described as 'celebrity politics', where the personal life of politicians is coming under increased scrutiny (Van Zoonen, 2006). Female politicians report that their personal lives, particularly in relation to their families if they have children, is commented on more than that of their male counterparts (Brikse, 2004). An analysis of magazine articles showed that coverage of women's families and children in particular is negative, suggesting a conflict and incompatibility between their public and private duties (Van Zoonen, 2006). In contrast, for male politicians, having children represents reliability and integrity.

Attributing a focus on a politician's family life to media bias can be problematic given that some female politicians who have children emphasise their role in their family themselves, seeing it as an advantage in politics (Ross and Sreberny, 2000). A politician stressing family status may be justified given the evidence that the media portrays female politicians more positively when they conform to gendered norms of behaviour (Scharrer, 2002). This is demonstrated in a study of newspaper coverage of Palin's election campaign (Miller and Peake, 2013). Throughout her election campaign Palin emphasised her personal role in her family. Analysis of the tone of these articles showed that the most positive articles were those that focused on her family. The articles relating to Palin in terms of policy issues were the most negative in tone. However, this difference in tone may be due, not to gender bias in the media, but to particular issues with Palin's policies.

Use of Stereotypes

Research on gender bias in the media cites the use of gendered stereotypes in the portrayal of female politicians (Carroll and Schreiber, 1997; Kahn, 1996; Robinson and Saint-Jean, 1991; Ross, 2000; Tremblay and Belanger, 1997; Trimble et al., 2013). It also found that more informal naming conventions are used to refer to female politicians (Uscinski and Goren, 2011) aligning with Van Zoonen's (2006) finding that political coverage is more personal for women.

In an analysis by Tremblay and Belanger (1997) of 351 cartoon images from Canadian newspapers in 1993, featuring two female politicians, half of the images represented women as cinderellas, witches or victims of violence. In contrast, where stereotypes were invoked to represent male politicians, they were represented primarily as supermen, businessmen or athletes.

Emotional Control

In an analysis of how politicians' emotional responses to the murder of two Swedish politicians were portrayed in Finnish newspapers, Pantti (2005) detected gender bias in the portrayal of female politicians as being less in control of their emotions. This involved a qualitative analysis of 202 newspaper articles over a period of 10 days during September 2003. It was found that there was little reporting of male politicians' emotional reaction. Many of them refused to comment to the press. One politician was described as being "deeply shocked". Another politician was quoted as describing the events as "horrible". This contrasted with the many descriptions in the articles of women crying. They were depicted as trying to control their emotions but failing and subsequently breaking into tears. While the sample size of this study is small, this work identifies a new theme in the study of how gender bias in manifested in media coverage of politicians.

Personalised Coverage

Trimble et al. (2013), in a longitudinal study of newspaper content covering Canadian national elections from 1975 to 2012, found that coverage of female candidates accorded significantly more attention to their sexuality, looks, and marital situation than was the case for male candidates. However, they did not find any disproportionate focus on personal issues such as upbringing, age, children or gender. Trimble et al. (2013)'s study involved a content analysis of 2463 articles from Canadian national newspapers featuring a sample of candidates in Canadian national elections. The articles were manually coded according to the number of references in an article to the following: gender markers, sexuality, age, appearance, marital situation, children (or childlessness) and upbringing. The features extracted from the texts were based on a review of previous findings on gender bias in the coverage of politicians. These measures were then combined to produce a 'personalisation index' which indicated that there was consistently more personalisation in the coverage of female politicians than male politicians.

Coverage is more likely to focus on gender-based evaluations of the style and appearance of female politicians than of male politicians. In an analysis of leading newspapers in France, Italy, Spain and the UK, Garcia-Blanco and Wahl-Jorgensen (2011) concluded that female politicians were judged by their appearance more than their male peers. The sample of articles from the newspapers focused on the 15 days after Zapatero, the Spanish Prime-Minister, appointed a new cabinet the majority of whom were women. Articles relating to the topic of women in politics or which mentioned female ministers were identified. This comprised a sample of 164 articles. Their analysis involved both quantitative content analysis and qualitative analysis of the newspaper text. In relation to comments about style and appearance, Garcia-Blanco and Wahl-Jorgensen (2011) concluded that "stories focusing on women's physical appearance and celebrating traditional gender roles, reinforcing the construction of women politicians as falling outside of the norm. This makes it almost impossible to discursively construct women 'just' as politicians, assessing their suitability on the basis of their education, previous experience, and their political performance." The study finds references to the physical appearance or style of the female ministers in 15 percent of the newspaper articles studied. However, the authors gathered no comparable data on male politicians, making it difficult to draw conclusions regarding gender bias on the part of the newspaper.

This focus on the personal style of politicians is supported by a study by Sreberny-Mohammadi and Ross (1996) which interviewed British MPs about their opinions on the coverage they receive in the press:

> "I don't know whether it is deliberate or it's so ingrained, but a woman's appearance is always commented on, her age is always commented on, her style of dress is always commented on. That never happens to male politicians, ever, unless they have made a particular point about their style but then they are presented as extreme, exceptions that prove the rule. Women are never the right age. We're too young, we're too old. We're too thin, we're too fat. We wear too much make-up, we don't wear enough. We're too flashy in our dress, we don't take enough care. There isn't a thing we can do that's right." (Sreberny-Mohammadi and Ross, 1996, p.108)

Style of Politics

In a study of whether there are differences in how newspapers report speech of male or female politicians, (Gidengil and Everitt, 2003) extracted the verbs used to describe the action of a politician speaking in broadcast media. They used transcripts of television news coverage of the 1993 Canadian national election. From this, a sample of 956 instances of reported speech were extracted. The verbs were categorised as to whether they were positive, negative or neutral. For example 'admit' was considered negative given that it is commonly used in the context of politicians confessing to or revealing negative information. The verbs were also classified according to whether they were expressive or aggressive. They found that verbs used to describe women's speech were twice as likely to appear on the list of the most negative speech verbs compared with those used to describe men's speech. Aggressive verbs were also applied much more frequently to the women (13.6 percent) than to the men (7.9 percent).

This was supported by an experiment evaluating people's perception of the verbs.

Differences in how men and women's speech is described was also reported by British MPs themselves who stated

> "I don't know how many times I've been described as having my claws out, instead of saying here's a woman being robust, which is what they would say about men" (Sreberny-Mohammadi and Ross, 1996, p.111)

Personal Issues over Policy

Research of campaigns in the 1980s found that regardless of a politician's role in political life, men were more likely to be discussed in terms of policy issues (Kahn, 1996; Kahn and Goldenberg, 1991). In a content analysis of newspaper coverage of 26 US Senate campaigns in 1984 and 1986, Kahn and Goldenberg (1991) found that female candidates consistently received less attention related to policy issues in the press than their male counterparts. Recent studies have identified difference in the kinds of policy issues related to male and female politicians with women being more likely to be associated with health and social issues (Brikse, 2004; Carroll and Schreiber, 1997). Carroll and Schreiber (1997) examined coverage of members of Congress in the US national media from 1993 to 1995. They found an excessive focus on issues particularly pertaining to women.

> "What is missing from general press coverage on women in Congress is any sense that women are important players on legislation other than women's health, abortion, and a handful of

other related concerns" (Carroll and Schreiber, 1997, p.145).

These findings were contradicted by Bystrom et al. (2001), who in a content analysis of coverage of US Senate and Gubernatorial primary campaigns found few gender differences in the policy focus of articles. Women's issues were found to be only marginally more associated with female candidates but the difference was not statistically significant (15.7 percent female vs. 10.2 percent male) (Bystrom et al., 2001).

However, a gender bias evidenced through the media's positive coverage of female politicians when they discuss more traditionally feminine issues was cited by Scharrer (2002). In this study, stories featuring Hillary Clinton when she was First Lady which involved issues that were political rather than those traditionally associated with First Lady roles, were more negative. This aligns with research on Palin's campaign which showed that coverage was more positive when the issues pertained to family and private life.

Masculine Narrative/Methaphors

Gendered mediation was first outlined by Ross and Sreberny-Mohammadi (1997) to reference how conventional news uses stereotypically masculine narratives to frame news stories, thereby alienating female politicians. Much recent work on the representation of women in politics analyses how political news stories are framed and how this effects female politicians (Burns et al., 2013; Trimble et al., 2013). Framing political discourse around domains not usually associated with women, such as war, alienates them from politics and portrays women's engagement with politics counter to gender stereotypes (Adcock, 2010). Behaviour that is counter to gender stereotypes is often viewed more negatively in the media than behaviour that aligns with gendered stereotypes (Burns et al., 2013; Eagly et al., 1992; Scharrer, 2002). Framing the discourse of politics around domains such as war has therefore been identified as evidence of a gender bias.

In an analysis of the portrayal of political events through metaphors of war, Gidengil and Everitt (1999) studied video recordings of Canadian leaders' debates and compared them with media coverage of those debates. While in the video recordings, female candidates were less aggressive than the male politicians during the debate, in the media coverage they were portrayed as more aggressive. Gidengil and Everitt (1999) concluded from this that framing the discourse of politics around war imagery served to highlight counter-stereotypical gender behaviour in women.

Focus on Gender

Norris (1997) identified an unnecessary focus on gender in stories about female politicians which disadvantaged them. Through analysing broadcast and newspaper coverage of political events, Norris (1997) demonstrated how gender was often used to frame the narrative around certain events. This is manifested through a female politician's gender being pointed out and thereby her engagement in politics being conveyed as unusual and/or unexpected. Female politician's actions on women's issues were highlighted to the neglect of other issues. Such focus on the gender of a female politician, serves to reinforce the concept of politics as normally a male domain (Norris, 1997).

Coverage of Palin's 2008 campaign illustrates this focus on a female politician's gender. A content analysis of newspaper coverage of her campaign showed that Palin's gender was mentioned four times more than that of her male running mate, Biden (Miller et al., 2010). In a content analysis of 6,600 articles from leading US newspapers in the 2008 democratic nominee elections, articles were 9 times more likely to mention Clinton's gender than Obama's. This mentioning of Clinton's gender tended to be accompanied by questioning of her likelihood to succeed. Clinton's electability was questioned in 1.8 percent of the articles which did not mention her gender. However, her electability was questioned in 8.8 percent of the articles that did mention her gender. From this Miller et al. (2010) concluded that coverage of Clinton's campaign implicitly cued readers to think that she might not be electable because of her gender.

Summary

Research on the media's representation of women in political life has consistently found evidence of gender bias. While many of the kinds of bias identified are interrelated, common themes emerged in this review of the literature. These themes broadly include an over-emphasis on the politician's gender, a focus on their personal life, negative stereotypical descriptions of them and a gender-based association with policy issues.

The methods employed by many of these studies involved content analysis. As will be discussed in Chapter 5, these methods are comparable to a corpus-based approach to analysing texts where the researcher examines the occurrences of certain predefined attributes of the texts. These attributes of texts informed what kinds of features were examined in this research and also how the patterns identified by the text classifier were interpreted. The next section of this thesis focuses on how machine learning can be used to analyse texts for evidence of the kinds of gender bias that this section outlined and also uncover new ways in which gender bias is manifested.

2.5 Text Classification in Gender and Language Research

An objective of this thesis concerns how text classification techniques can be used to examine differences in newspaper coverage of male and female politicians and identify gender bias. This section reviews literature critiquing existing approaches to analysing gender bias in text and in doing so explains the rationale for exploring a new approach to identifying gender bias in text. Literature which uses text classification to address questions related to this research is reviewed, followed by a discussion of how these approaches could be used to analyse gender bias in media content.

Neuman (2007, p.273) proposed that "qualitative content analysis is not highly respected by most positivist researchers. Nonetheless, feminist researchers and others adopting more critical and interpretative approaches favour it". A reliance on small text samples has led to questioning the scientific validity of the findings of feminist qualitative research. "Qualitative content analysis relies heavily on researcher 'readings' and interpretation of media texts. This intensive and time-consuming focus is one of the reasons that much qualitative content analysis has involved small samples of media content and been criticised by some researchers as unscientific and unreliable." (Macnamara, 2003, p.5)

The majority of studies on gender bias in the media representation of politicians has used quantitative approaches to analysing media content. Quantitative approaches to text analysis have been criticised for "not being able to capture the context within which a media text becomes meaningful" (Newbold et al., 2002, p.84). Computational methods of analysing language have the potential to address some of these criticisms.

The increasing amount of media data becoming available digitally has presented new opportunities for using computational approaches to analysing language. The volume of data available makes it theoretically possible to automatically identify patterns in texts and make inferences from this. The ability to harness the analytical capacity of computers to analyse language is resulting in the expansion of existing research paradigms. While it had been a widely accepted tenet of traditional research methodologies that no data gathering approach could be general, realistic and precise, the availability of computational approaches to gathering data makes this possible. However, as demonstrated subsequently in this thesis, these approaches do not replace the importance of human analysis and interpretation of text but rather, they can act as a valuable tool to process and highlight patterns in large volumes of data that can subsequently be interpreted by the researcher.

Wiedemann (2013) described using machine learning to identify patterns in text as an approach where computer-assisted text analysis grows from word counting to a tool enabling the researcher to answer their questions in a controlled way guided by theoretical or empirical foundations (Wiedemann, 2013). Scharkow (2013) argues that text classification is most analogous to traditional approaches to quantitative content analysis.

> "Supervised text classification, which uses superficial statistical algorithms from machine learning, has the potential to become a standard method for quantitative content analysis. By using manually coded material for training, supervised classification seamlessly bridges the gap between traditional thematic and automatic content analysis. Unlike other automatic approaches, supervised classification does not require a completely different way of conducting content analyses" (Scharkow, 2013, p. 2).

Wiedemann (2013) describes text classification as having the the potential to bridge qualitative and quantitative analysis by supplementing qualitative research and thereby addressing many of the criticisms of qualitative research in terms of scientific reliability and validity. This thesis therefore explores the validity of using text classification in identifying gender bias in text. By bridging qualitative and quantitative research, this method of text analysis to address gender bias addresses many existing critiques of feminist studies of language.

Detailed descriptions of the text classification process is presented in Chapter 4. The process involves using a machine learning algorithm to identify patterns that differentiate between categories of documents. A machine learning algorithm requires that text is represented numerically by an array of vectors. How the documents are represented determines the kinds of patterns that can be identified by the algorithm. Less processing of the data will allow features to emerge from the data that are unexpected by the researcher. However, if particular linguistic phenomena are being investigated, the documents can be processed to represent only those features of interest to the researcher. The next section reviews literature that explore a range of such approaches to automatically classifying documents.

2.6 Text Classification Research

This section reviews text classification research. While there are no existing studies which use text classification to identify gender bias in media content, there are studies that address similar research questions. This research builds upon the findings and methodologies developed in these studies.

This review highlights the findings of the research concerning what attributes or features of the texts were found to be useful in classifying documents. What features are extracted from texts for analysis using text classification are often guided by theoretical frameworks related to the research questions being addressed.

To date two broad approaches to analysing patterns in data have been identified. These are data/corpus-driven or ata/corpus-based approaches (Baker, 2014; Flaounas et al., 2010). An example of a corpus-based approach is Argamon, Whitelaw, Chase, Hota, Garg and Levitan's (2007) automatic classification of texts according to style. This study analysed specific linguistic features of the texts that were assumed to constitute writing style. Other research adopts a data-driven approach where little pre-processing is applied to the text and the patterns emerge through a large-scale analysis of unfiltered features of texts (Baker, 2010). Since this thesis explores both approaches to extracting features, in order to identify which is the most suitable for identifying gender bias, findings from previous research which used both a data-driven and data-based approach are reviewed.

The findings of the research reviewed in this section are concerned with the accuracies gained in automatically classifying documents. These indicate the usefulness of particular features in differentiating documents according to different categories. Some of the studies also analyse which individual features are associated with which category. For example, in an analysis of political ideology in texts, Diermeier et al. (2012) found that it is possible to classify text according to ideology based on the words they use. However, Diermeier et al. (2012) also analysed what kind of words are associated with which ideology. As this approach is pertinent to this thesis, the findings of such feature analysis are also reviewed.

Text classification studies have addressed a range of issues such as an author's gender, political ideology, subjectivity and the author's personality type. While these are not directly related to gender bias, the authors do use text classification to explore attributes of the texts and it is therefore important to review the findings of this research in order to identify the approaches to text classification that can be applied to the subject of gender bias in text.

2.6.1 Automatically Identifying an Author's Gender

Research in text classification has identified features of texts which differentiate between texts written by women and men. Baker (2014) outlined how this research focuses on identifying differences in how men and women speak or write and can often be used to reinforce a 'difference mindset'. In contrast, research in the area of Gender and Language has shown that the similarities in how men and women use language outweigh the differences Hyde (2005).

Some the text classification research which examined differences in male and female use of language extracted linguistic features from the text which captured the style of writing and tested whether men and women wrote using different styles (Corney et al., 2002; Hota et al., 2006; Koppel et al., 2002; Sabin et al., 2008). Other approaches used little pre-processing and from this the machine learning algorithm highlighted patterns differentiating between those written by men and women that were unexpected by the researchers (Argamon et al., 2009; Boulis and Ostendorf, 2005; Nowson and Oberlander, 2006; Opsomer et al., 2008). Yet more research adopted both approaches testing whether particular linguistic features correlated with one gender over another and also conducted analysis on the entire unprocessed data in order to allow the classifier to identify patterns differentiating the texts that were unexpected by the authors

(Argamon, Koppel, Pennebaker and Schler, 2007; Kucukyilmaz et al., 2006; Otterbacher, 2010).

Koppel et al. (2002) investigated whether the writing style of a piece literature was indicative of an author's gender. Using a sample of 566 texts from the British National Corpus with an average length of 34,000 words, they developed a model to predict an author's gender based on features extracted from the texts which were designed to represent the writing style of each text. Style was captured by representing the texts as frequencies of function words. Based on previous work in the area of stylometry which identified attributes of writing which characterise a writer's style, they selected function words as features as they are generally content-independent words. Koppel et al.'s (2002) experiments predicted the gender just using function words to an accuracy of approximately 80 percent. Hota et al. (2006) achieved the same level of accuracy in a similar experiment using function words to represent writing style. Hota et al. (2006) found that using function words as features in representing the documents was as useful as more detailed linguistic features such as character based, syntactic and structural features.

A corpus-driven approach to text classification can be used to identify patterns in texts previously unexpected by researchers and these patterns can subsequently generate new research questions which can often be analysed using more qualitative methods of content analysis (Heyer et al., 2006). This is demonstrated by Argamon et al. (2009), who uncovered patterns of difference in the writings of male and female authors in French historical literature not previously identified by existing theories of gender and language. The features that were extracted from the documents included words, lemmas of words, parts of speech and part of speech categories as features. The accuracies of the predictive models generated by the classifiers are presented in Table 2.6.

The classifier accuracies indicate the extent to which each approach to representing the documents (feature sets) predict the author's gender. The model was not tested on a more contemporary corpus so it is unknown how generalisable their findings are to male and female authors.

Word	Lemma	Part of Speech	Part of Speech Group
85.7%	85.9%	74.4%	74.2%

Table 2.6: Classifier Accuracy in Author Gender Prediction (Argamon et al. 2009)

Along with evaluating which features best indicate an author's gender, Argamon et al. (2009) examined the individual features of French historical literature that were most associated with either gender. This showed that personal pronouns and words conveying negative sentiment were more associated with female writing than male writing. Male writers also used more determiners and numerators. In writing on religious topics, men used religious terminology while women used more secular language.

Gender differences are evident in both the writing style and content of blogs (Argamon, Koppel, Pennebaker and Schler, 2007). Analysis of style based features showed that female writers are more associated with the use of personal pronouns, conjunctions and auxiliary verbs. Articles and propositions were used significantly more frequently by male bloggers. Argamon, Koppel, Pennebaker and Schler (2007) also used content words as features which revealed that male bloggers were more likely to write on topics such as religion, politics, business and the internet. Female bloggers wrote more on topics related to fun, romance and household related topics. They also ran classification experiments on the age profile of the authors and found that different age groups were more associated with certain topics. The highest accuracy results were obtained by using function words as features. Nowson and Oberlander (2006) also predicted the gender of bloggers to a high accuracy of 93 percent by using bigrams (sequences of two adjacent words) and trigrams (sequences of three adjacent words) as features.

Blogs were analysed by Corney et al. (2002) using content-neutral features including vocabulary measures and style markers. The authors also used structural document features which included information about the documents and a list of characteristics of words which they cited as gender-preferential. These were defined as words ending with ible, ful, able and apology words. The rationale used to generate the list of gender predicting words is unexplained. However, they are possibly derived from work such as that of Lakoff (1975) which associated women with an increased use of apology words. The full list of feature sets are listed in Table 2.7. Along with Koppel et al. (2002), Corney et al. (2002) found that function words best predicted the blogger's gender. The gender-preferential words only increased the results marginally. These results are presented in Table 2.8. Table 2.7 shows the list of all feature sets used in Corney et al. (2002)'s research. It is one of the most comprehensive set of features used in such studies of text classification. However, while it is clear from this study that some of the features can indicate the gender of an author, further detail on the meaning of this requires a deeper examination of what individual features correlate with which gender.

Feature Types
Document Based
Number of blank lines/total number of lines
Average sentence length (number of words)
Word-based
Average word length
Vocabulary richness
Number of function words
Number of short words
Vocabulary measuress
Character-based
Number of characters in words
Number of alphabetic characters
Number of upper-case characters in words
Number of digit characters in words
Number of white-space characters
Number of spaces
Number of spaces/Number white-space chars
Number of tab spaces
Number of tab spaces/Number white-space chars
Number of punctuation characters
Function Words
Function word frequency distribution
Other
Word length frequency distribution
Structural
Reply status
Has a greeting acknowledgement
Uses a farewell acknowledgement
Contains signature text
Number of attachments
Position of re-quoted text within e-mail body
HTML tag frequency distribution
Gender-Preferential
Number of words ending with able
Number of words ending with al
Number of words ending with ful
Number of words ending with ible
Number of words ending with ic
Number of words ending with ive
Number of words ending with less
Number of words ending with ly
Number of words ending with ous
Number of sorry words
Number of words starting with apolog

Table 2.7: Features Used In Classifying Male and Female Blogs (Corney et al. 2002)

Feature Set Type	Accuracy %
All attributes	70.2
Character-based attributes removed	70.0
Word-based attributes removed	69.6
Word length distribution removed	67.4
Structural attributes removed	68.1
Function words removed	64.0

Table 2.8: Effect of Feature Type on Classification Accuracy Results (Corney et al. 2002)

The value of examining the most useful features in differentiating categories of texts within a feature set is demonstrated by Boulis and Ostendorf (2005). In this study gender differences in conversational speech of men and women were explored by analysing transcripts of telephone conversations. The features extracted from the texts included words (unigrams) and two word sequences (bigrams). Table 2.9 shows the accuracies obtained using these two difference approaches to feature extracting and a number of different machine learning algorithms.

Machine Learning Algorithm	Unigrams	Bigrams
Cosine	76.3%	86.5%
Naïve Bayes	83.0%	89.2%
MaxEnt	85.6%	90.%3
SVM	88.6%	92.5%

Table 2.9: Accuracies for Speaker Gender Prediction (Boulis and Ostendorf 2005)

Boulis and Ostendorf (2005) also examined what features the machine learning algorithms used in the research found most indicative of an author's gender. These features are termed discriminative features. Table 2.10 shows the most discriminative features used. Of the 2000 most discriminative features, swear words were heavily associated with male bloggers. Family relationships such as the following appeared in the top 2000 words associated with female bloggers: children, grandchild, child, grandchildren, childhood, childbirth, kids, grandkids, son, grandson, daughter, granddaughter, boyfriend, marriage, mother, grandmother. Boulis and Ostendorf (2005) also found that the accuracy of the models increased when tested against same-gender conversations. This implied that linguistic conventions among participants in a conversation converged in cross-gender conversations aligning with research in gender and language exploring differences in linguistic conventions in same gender conversations (Coates, 2003, 2004*b*; Pichler, 2009). .

Male	Female
dude	husband
shit	husband's
fucking	refunding
wife	goodness
wives	boyfriend
matt	coupons
steve	crafts
bass	linda
ben	gosh
fuck	cute

Table 2.10: Words Discriminating Between Male and Female Bloggers in Order of Importance (Boulis and Ostendorf 2005)

The gender of participants in online Turkish text messages was predicted to an accuracy of 82.4 percent by Kucukyilmaz et al. (2006) using a dataset of 250,000 short online messages with an average length of 6.2 words. Their research used the terms in the messages along with style markers as features in their experiments. These style based features along with how they were measured are shown in Table 2.11. They found that for smaller samples of the dataset, terms were the best indicator of the gender of the author. However when the dataset was enlarged, style was a better indicator.

Feature	Description	Possible feature values
message length	average message length	low, medium, high
word length	average word length	low, medium, high
stopword usage	frequency of stopwords	low, medium, high
stopwords	a list of 78 stopwords	exists, not exists
smiley usage	frequency of smileys	low, medium, high
smileys	a list of 79 smileys	exists, not exists
punctuation usage	frequency of punctuation marks	low, medium, high
punctuation	marks a list of 37 punctuation marks	exists, not exists
vocabulary	richness number of distinct words	poor, average, rich
character usage	frequency of each character	low, medium, high

Table 2.11: Style Based Features (Kucukyilmaz et al. 2006)

In exploring gender related differences in online movie reviews by men and women, Otterbacher (2010) discovered that the best indicator of an author's gender was the perceived value of the review by the reader. This demonstrates how findings in text classification studies can uncover unexpected patterns in data which can subsequently form the basis of new research topics. The features Otterbacher (2010) extracted from the movie reviews included style markers, content words and structural information about the documents such as the age of the review and user feedback. Their experiments using style and content as features aligned with findings of previous work such as that by Koppel et al. (2002) in which style was a better indicator of the gender of the author than other linguistic features. However, the best indicator of an author's gender was the feedback readers gave on the perceived utility of the reviews. The content related words were not as accurate as style based features in identifying gender. They found that female reviewers were more likely to discuss people, family and relationships and to use pronouns more frequently.

Sabin et al. (2008) analysed gender differences in both speech and writing in an experiment which involved gathering emails, essays and phone interview transcripts from 12 male and 12 female students. The authors extracted features from the data based on dictionaries from the Linguistic Inquiry and Word Count (Pennebaker et al., 2001) which they expected to be gender related. Examples of the word categories they used were words concerning emotions, social issues, leisure, family, finance, grooming and swear words. While most experiments excluding those on email messages yielded approximately 80 percent accuracy, function words were the most indicative of gender of the author or speaker. However, the authors did not present a breakdown of which features sets were indicative of which gender. This demonstrates how predictive accuracy alone is of limited value without an examination of the features which are most valuable in the text classification experiments.

Children's speech has also been successfully classified according to gender. Opsomer et al. (2008) yielded 70.5 percent classification accuracy using words as features extracted from transcripts of conversations from 12 children. There were a total of 806 conversations recorded. The features extracted from the transcripts were simply the words that the children used. While accuracies of 70.5 percent were obtained in the text classification experiments, which words were most important is not analysed. This research confirms that male and female speech in children can be differentiated to some degree of accuracy based on the words they use, but the nature of these differences are not explored.

2.6.2 Political Ideology Classification

Research using text classification in political science research has gained popularity (Grimmer and Stewart, 2013). Identification of gender bias in text is comparable to automatically identifying political ideology in text as both concern identifying ideologies expressed in texts. The former concerns ideologies pertaining to gender while the later concerns political ideology. In studies which use text classification to analyse texts to address political science research questions, analysis of the individual features (discriminative features) associated with each category is important.

Diermeier et al. (2012) used text classification to explore differences in speeches by liberal and conservative politicians in the US Senate. The data analysed was composed of transcripts of 400 Senatorial speeches from 1989 to 2008. The features extracted from the speeches included words occurring in the texts, stems of those words, nouns, verbs and adjectives. Table 2.12 shows the predictive accuracy gained from the classifiers in discriminating between speeches by liberals and conservative politicians.

Feature Representation	**word**	**stem**	**noun**	verb	**adj**	**adv**
Boolean	88.9%	88.0%	89.5%	82.9%	89.2%	77.2%
Normalised frequency	88.6%	89.7%	84.6%	66.7%	84.6%	61.8%
tf*idf	95.4%	94.0%	95.4%	93.4%	93.4%	88.9%

Table 2.12: Classifying Speech According to Political Ideology (Diermeier et al. 2012)

Their analysis of the most valuable features in the text classification experiments were those concerned with cultural issues. Economic issues were less indicative of whether the speaker was a conservative or liberal politician. Table 2.13 presents the most discriminative features identified by the text classifier when all of the words in the speeches were included as features in order of importance. From this the researchers identified patterns where ideologies were expressed not through talking about the same issues differently as was expected by the researchers, but about different issues. For example conservatives spoke about issues such as marriage, cloning, abortion and homosexuality. Gun control was the only cultural issue that democrats were associated with.

Liberal		Conservative	
FAS: -199.49	SBA: -113.1	habeas: 193.55	homosexual: 103.07
Ethanol: -198.92	Nursing: -109.38	CFTC: 187.16	everglades: 102.87
Wealthiest: -159.74	Providence: -108.73	surtax: 151.81	tower: 101.67
Collider: -142.28	Arctic: -108.3	marriage: 145.79	tripartisan: 101.23
WIC: -140.14	Orange: -107.98	cloning: 141.71	PRC: 102.9
ILO: -139.89	Glaxo: -107.81	tritium: 133.49	scouts: 97.55
Handgun: -129.01	Libraries: -107.7	ranchers: 132.95	nashua: 99.32
Lobbyists: -128.95	Disabilities: -106.44	BTU: 121.92	ballistic: 97.22
Enron: -127.71	Prescription: -106.31	grazing: 121.59	salting: 94.28
Fishery: -127.3	NIH: -105.52	unfunded: 120.82	abortion: 91.94
Hydrogen: -122.59	Lobbying: -105.35	catfish: 120.82	NTSB: 93.81
Souter: -121.4	NRA: -105.2	IRS: 114.91	Haiti: 97.28
PTSD: -119.87	Trident: -104.15	unborn: 111.88	PAC: 92.85
Gun: -119.52	RNC: -103.46	Taiwan: 111.13	taxing: 90.39
Firestone: -117.9	Lobbyist: -99.38	PLO: 106.56	nonseverability: 89.26
Lakes: -114.84	Homelessness: -95.68	EMS: 103.99	embryonic: 88.83
Souter: -121.4	NRA: -105.2	IRS: 114.91	Haiti: 97.28
PTSD: -119.87	Trident: -104.15	unborn: 111.88	PAC: 92.85
Gun: -119.52	RNC: -103.46	Taiwan: 111.13	taxing: 90.39
Firestone: -117.9	Lobbyist: -99.38	PLO: 106.56	nonseverability: 89.26
Lakes: -114.84	Homelessness: -95.68	EMS: 103.99	embryonic: 88.83

Table 2.13: Weighting of Features in Classifying Speech According to Political Ideology (Diermeier et al. 2012)

In a similar attempt to classify political speeches according to their ideology, Yu et al. (2008) analysed 2005 US Senate and House speeches using text classification. They took membership of the Republican or Democratic party as indicative of ideology. They used words as features, excluding rare and common words. Table 2.14 shows the words most indicative of political ideology. They found better predictive accuracy using speech data from the US House speeches rather than the US Senate. This suggests that the US House was more partisan in terms of political ideology than the US Senate. This showed that Republicans were associated with discourses on the economy, abortion, tax and terrorism while democrats were more focused on social welfare, health-care and children.

Republican	Democrat
economy	cuts
commend	republican
reforms	opposition
bringing	care
thank	new
understanding	cut
jobs	budget
gentleman	majority
worked	programs
assets	iraq
area	debt
hard	middle
times	health
chairman	substitute
embryo	children
urge	oppose
areas	values
passage	community
growing	fails
dollars	administration
committee	diabetes
stop	women
certainly	benefit
government	proposed
terrorists	failed
growth	medical
terror	child
issue	question
small	bush
tough	republicans

Table 2.14: Most Discriminative Features in Classifying US Liberal and Conservative Speeches in Order of Importance(Yu et al. 2008)

Jiang and Argamon (2008) conducted research to automatically identify conservative and liberal blogs. The data was comprised of frontpage text from 1054 liberal blogs and 793 conservative blogs. The front pages all contained at least 140 words. From this they extracted sentences that contained subjective statements. In one experiment, words were extracted from these sentences and used as features and in another, phrases involving a noun with either a verb or an adjective were used as features. Accuracies gained by using these features were between 80 and 87 percent.

Koppel et al. (2009) achieved almost perfect accuracy in experiments to classify Islamic religious texts according to their source organisation or ideological stance. The features used were simply the 1000 most frequently used words. In the experiment to predict the organisation that wrote the documents the corpus used consisted of 552 documents labeled according to whether it as written by the Hamaz, Al Qaeda, Muslim Brotherhood or Hizballah. The documents were sourced by what the authors describe as a wide variety of public sources. Using 10-fold cross validation 100 percent accuracy was achieved. As part of the same research a different corpus of 1,485 documents demonstrating various ideological streams associated with Salafi-jihadi, Muslim Brotherhood, mainstream Islam and Wahhabi was examined. Using the most frequent 1000 words in 10-fold classification experiments classified all but two of the documents correctly. Using only function words as features an accuracy of 73 percent was obtained. This research demonstrates how simple features that require little expert intervention can be extremely effective in identifying political ideology of the authors of a text.

2.6.3 Subjectivity Classification

There has been a considerable volume of research on the identification and evaluation of subjective clauses in texts. The goal of identifying opinion in text is comparable to identifying gender bias in text. For this reason, the approaches to text classification used in subjectivity classification research has informed the design of this thesis.

Different approaches to representing documents have been explored in subjectivity classification research. Pang and Lee (2008) classified movie reviews into either positive or negative categories based on single word features. Kim and Hovy (2006) found that including information about the place of the words in sentences improved results compared with just using the words as features. Dave et al. (2003) found that bigrams (two word sequences) and trigrams (three word sequences) improved the classification accuracy of polarity of product reviews.

Subjectivity classification also involves much more complex feature extraction methods than other types of text classification. For example Riloff et al. (2006) gained high classification accuracy by using features which used a subsumption hierarchy. This involved identifying phrases that can be grouped together to form higher level categories of terms which carry a similar sentiment.

Part-of-speech tags are often very important in sentiment classification. The presence of adjectives in a text for example can often indicate that sentiment is being expressed in a text. One such study was that undertaken by Riloff et al. (2006), who compiled a lexicon of sentiment bearing terms based on the appraisal theory formulated by Martin and White (2005). Using this lexicon, combined with term frequencies, obtained an accuracy of 90.2 percent in classifying texts according to the sentiment expressed in them.

2.6.4 Personality Type Classification

Luyckx and Daelemans (2008) utilised syntactic features to generate a model to predict the personality type of the author. The personality types of the authors were determined using an online version of the Myers-Brigs Type Indicator test. The participants were then asked to write on a topic given to them by the researchers. Based on work by Stamatatos et al. (2001), the authors took syntactic features as indicators of style. The corpus was parsed using a shallow parser which tokenised, tagged and parsed sentences into subject and object clauses. The findings of this research suggested a link between writing style and personality type. This demonstrates how text classification techniques can be used to highlight unexpected patterns and suggest research questions which could be examined by further research.

2.7 Conclusion

This chapter reviewed literature from the broad range of disciplines related to this thesis. The theoretical foundation motivating this thesis was explained and key research in linguistics and feminist theory on language was reviewed. Research in feminist linguistics was described showing how feminist theories on how women are represented in language were tested empirically using methods from linguistics. The review of these studies presented the main ways in which language has been used to negatively portray women.

Research bridging feminist linguistics and political science which analysed how female politicians are represented in the media was then reviewed. Many of these studies used content analysis methods of analysing texts. The focus of the review of this literature was on highlighting what kinds of attributes of the content were analysed as these informed the design of this research and interpretation of the results. The kinds of bias that were found in media coverage of female politicians also mirrored the findings from feminist linguistics of bias in how the media represent women more generally.

The final section of this thesis explored how computational approaches, specifically text classification techniques could be used to identify differences in how male and female politicians are represented in the media. Research which motivated the decision to explore a computational approach to text analysis in this research was presented. Studies which have addressed similar research questions using automatic text classification techniques were then reviewed. The main focus of this was on identifying how texts were represented so that a machine learning algorithm could identify patterns in them.

The next chapter presents the methodology used in this research. This explains in detail the methods of automatic text classification using machine learning referred to in this chapter. How these methods were used in this research to analyse text within a feminist theoretical framework, building upon findings of previous studies on the portrayal of women in language is described.

Chapter 3

Methodology

3.1 Introduction

This chapter sets out the methodology used in this thesis. Chapter 2 considered how computational approaches to analysing text could address some critiques of traditional approaches to analysing text (Wiedemann, 2013). This chapter explains how techniques of text classification can achieve this and outlines the methodological decisions in this research and the rationale behind them. The computational methods used and the design of the research is also presented.

Creswell (2014) outlined three main components of an approach to research methodology including philosophical viewpoint, design and methods. This framework is used to structure the discussion of the methodology of this research. However, since computational approaches to data collection are not incorporated into Creswell's framework, some methodological decisions extend beyond the framework outlined by Creswell (2014).

Throughout this thesis gender is used to refer to cultural and social attitudes that together shape and sanction "feminine" and "masculine" behaviors, products, technologies, environments, and knowledge (Schiebinger et al., 2011-2013). According to Schiebinger et al. (2011-2013), gender does not necessarily match sex. For this reason, and in accordance with norms in related academic literature, gender rather than sex as a biological quality is referred to throughout this thesis to denote whether individuals are aligned to the cultural or social category of male or female.

3.2 Ontology and Epistemology

This research is grounded in a feminist poststructuralist ontology and epistemology, which asserts that societal concepts of gender originate largely from language and how it is used. Feminist poststructuralism is concerned with how language, subjectivity, society and power contribute to the construction of concepts of gender and associated societal gender dynamics (Weedon, 1997). Research in this area seeks to understand how language and discourse generate social constructions of gender that disadvantage women. These disadvantages are particularly related to issues of power. Poststructural approaches to feminist research analyse how language and discourse create and reinforce gendered power dynamics within society and in doing so identify strategies for change (Bristor and Fischer, 1993).

In line with a feminist poststructuralist ontology, the assumption in this research is that gendered power dynamics are created and sustained to a large extent by societal discourse. Access to political power and social acceptance of its possession is therefore largely determined by societal political discourse. A significant proportion of this political discourse takes place in the mass media.

This research analyses such political discourse to identify whether the ways in which language is used in the media might contribute to gendered societal conceptions of political power.

3.3 Research Philosophy

Feminist poststructuralism is a research philosophy with associated methods of language analysis such as deconstruction (Derrida, 1976) and discourse analysis (Foucault, 1970). However, post structuralism, as a research paradigm, does not prescribe principles that guide the research design process. Therefore, holding a feminist poststructuralist ontology informing the research question and analysis is compatible with a research process informed by philosophies such as positivism, post-positivism, interpretivism or pragmatism.

Feminist research, regardless of the multiplicity of viewpoints within this category, is commonly characterised as rejecting a post positivist approach to research (Creswell, 2014, p.64, Prasad, 2005, p.173). However, while this may have held true during the emergence of feminism within academia in the 1970s and 1980s, many examples have emerged from a feminist ontological perspective that work within positivist and post-positivist research paradigms (Denzin and Lincoln, 1994). Research is being carried out which not only addresses gender in empirical research but also embraces feminist ontological and epistemological perspectives, while remaining aligned with the established methodologies, often post-positivist, of the discipline within which the researcher is working.

The research philosophy that guided the design of this research is pragmatism. Creswell (2014, p.11) described pragmatism as an approach that determines what and how to do research, based on the intended consequences.

Pragmatism is associated with a mixed methods approach to research. This is the approach taken here.

3.4 Research Strategy

New computational approaches to traditionally qualitative research topics are challenging the traditional dichotomy of qualitative and quantitative research as described in chapter 2. This is particularly applicable in social science where quantitative computational approaches to research are being applied within a qualitatively oriented methodological framework. Qualitatively oriented social scientists, particularly poststructuralists, commonly use methods such as close reading and qualitative analysis on small samples of textual data. Computational approaches to language research have been associated with purely quantitative strategies of processing text. However, advances in methods of processing and understanding natural language are increasingly enabling meaning to be considered as part of the computational process thus dissolving the binary distinction between quantitative and qualitative approaches to text analysis.

Using a computational approach to addressing the research question in this thesis enabled the research to mix both qualitative and quantitative approaches to research. However, this was not by utilizing a mixed methods approach in the sequential sense outlined by Creswell (2014, p.11) but by breaking down the distinction between qualitative and quantitative data and analysis. Creswell (2014, p.11) outlines a typology of mixed methods research. However, his typologies assume a distinction between sources of qualitative and quantitative data and the means by which they are gathered. Binary distinctions between qualitative and quantitative approaches are implicit in other work such as Denzin

(2009, p.11) when he asserts that "the bias inherent in any particular data source, investigators, and particularly method will be cancelled out when used in conjunction with other data sources, investigators, and methods......the result will be a convergence upon the truth about some social phenomenon". This research adopts the strategy discussed by Caracelli and Greene (1993) of transforming qualitative data into quantitative data. In this way this research blends the two approaches to form one research process.

In this research, the primary data, which is a corpus of articles from Irish newspapers, are transformed into quantitative representations of the text. The quantitative representations involved representing text as an array of vectors. Each vector corresponds to a particular feature of the text. Multiple methods of quantifying the text features were used, such as a binary representation indicating the presence of a word in a body of text or a measure of the frequency of the word. Beyond representations of words in the text, conceptual features of the text were represented. For example, the sentiment embedded in a text was measured and represented quantitatively. This incorporation of context in this quantitative representation allowed for analysis of text computationally incorporating information that would previously only have been feasible in qualitative analysis. Both inductive and deductive reasoning were used in analysing this data. Use of machine learning to determine rules from examples is an inductive process. Scharkow (2013) noted that a deductive approach is utilised when dictionaries are used to extract features for analysis from the text.

Hansen et al. (1998) and Harrington et al. (2008) outlined the compatibility and complementarity of quantitative analysis of text with qualitative approaches. This dissertation adopts the stance that Hansen et al. (1998) outlined along with others such as Curran (2002) and Gauntlett (2008) where qualitative and quantitative approaches are viewed not as opposing but as complementary and valuable aspects of the study of gender and language. This perspective is captured well by Hansen et al. (1998), who posited that "rather than emphasizing its alleged incompatibility with other more qualitative approaches (such as semiotics, structuralist analysis, discourse analysis) we wish to stress that content analysis is and should be enriched by the theoretical framework offered by other more qualitative approaches, while bringing to these a methodological rigour, prescriptions for use, and systematicity rarely found in many of the more qualitative approaches" (Hansen et al., 1998, p.81). Harrington et al. (2008, p.101) however, pointed out how quantitative analysis must "take care of how raw data of counted tokens is translated into meaningful patterns". This thesis addresses this by analysing the quantitative findings of the research using a qualitative analysis of the occurrence of linguistic features in the text.

3.5 Research Methods

3.5.1 Content Analysis and Text Classification Techniques

Text classification refers to the task of automatically classifying documents based on certain features extracted from the text. Most classification studies involve binary categories. The process involves gathering text documents as training data, extracting features and representing them numerically. Machine learning algorithms then either categorise texts according to given categories (supervised) or group the texts according to categories which are identified by the machine learning algorithm (semi-supervised). The supervised approach involves labelling the documents according to predefined categories. The algorithm generates a predictive model which, when applied to new data, predicts the category of the text. This predictive model is tested and the level of accuracy of the prediction is evaluated. The supervised machine learning approach was adopted for use in this research.

Deciding how to represent documents in text classification studies often involves preliminary qualitative analysis of the text. Extracting features from documents is comparable to coding text in content analysis. Representations of the texts can involve simple features such as word counts to more abstract representations of the meaning in the language used.

Extracting narrowly defined features from texts and analysing whether they are associated with membership of a category of texts constitutes what Baker (2010) described as a corpus-based approach to analysing texts. An approach to text classification where a broad range of features are extracted in order that the machine learning algorithm can highlight patterns that may be unexpected by the researcher is a corpus-driven approach (Baker, 2010). This research utilised both of these approaches to feature extraction. Features are extracted to test findings of previous studies of gender bias in media coverage of politicians. An unfiltered approach was also applied where all of the words used are analysed to allow a machine learning algorithm uncover new patterns of difference in how male and female politicians are featured.

Heyer et al. (2006) describes how the analysis of data is often an iterative process whereby more qualitative data can be added to the analysis. Significantly, Heyer et al. (2006) points out how highly accurate classification results are not always what is of interest to researchers in the social sciences. The results are used instead as part of an overall, often qualitative, analysis of the texts.

The following is an overview of each step in the classification process with a focus on its application in analysing media content:

Text Pre-Processing: Text classification requires that documents are transformed into a numerical representation of the text. This involves identifying characteristics or features of the documents and representing those numerically. Each document appears then as an array of features with the presence or absence of these features represented numerically. What kind of features of the texts are represented is the focus of much research in text classification studies (Whitelaw et al., 2005). Extracting features of texts and representing them numerically is analogous to coding in content analysis.

Classification Algorithms: Supervised machine learning algorithms build models that predict the category of a document based on examples provided by the researcher. The overarching aim of using machine learning algorithms in content analysis is to identify patterns in text documents which correlate with certain attributes of the text. There are practical applications for developing models that can automatically and accurately classify documents into predefined categories. However within the context of media content analysis, the value is in uncovering those features of the text that are indicative of the document belonging to a certain category. This will be illustrated further in the studies reviewed in the next section.

Classifier Evaluation: Evaluation of the performance of a text classification involves measures of the accuracy gained when the predictive model is applied to new documents. This is indicative of the strength of the correlation between the features extracted from the documents and the category to which it can be assigned. Evaluation of the classifiers in relation to a social science research question also involves a qualitative analysis of what features were used to identify the category that the document belongs to. As Heyer et al. (2006) posits, high predictive accuracy does not always yield the most insightful results in the context of research in social science or media content analysis. Therefore, while high accuracy indicates the strength of the correlation between certain features of a text and its overall category, gaining high accuracy is not always the primary goal of text classification in such studies.

3.5.2 Content Analysis in Gender Research

In a review of research on gender and language, Neuendorf (2011) identified a lack of scientific rigour in many studies which used content analysis methods. To address this she developed a framework for methodological standards in content analysis in gender research. This framework informed the design decisions of this research. Neuendorf (2011) reviewed 133 research papers and identified little evidence of reliability assessment. While the framework does not include the text classification methods used in this research, the overall approach to ensuring scientific rigour in content analysis is revelent. The goal of data mining techniques when applied to social science questions is not always to gain the highest accuracy in classification experiments (Wiedemann, 2013). Text classification in the social sciences can therefore require specific evaluation strategies. For this reason, the framework outlined by Neuendorf (2011) was used in this research to evaluate the methods in order to ensure the scientific rigour in the research design. The following outlines these methodological considerations and how they were applied in this study:

Framework for Methodological Standards

Theoretical and conceptual backing: Neuendorf (2011) posited that content analysis must be guided by a theoretical framework. This research is grounded in post-structuralist thesis to guide the method feminist theories of gender and language which guided the framing of the research question and informs the interpretation of the results of the research.

Scope of the Research: The research utilised what Neuendorf (2011) described as an "integrative" approach which emphasises the relationship between the dependent variable, which is the gender of the politician, and a range of independent variables, from both the content itself and information external to the content such as meta-information about the articles including the section, date, and placement in the newspaper.

Review of past research: Neuendorf (2011) emphasised the importance of reviewing past research and building upon these methodological approaches. While this study does introduce a new methodological approach to gender and language research, it does so by building upon previous research, particularly in the feature extraction process. Where dictionaries are used for information extraction purposes they rely on pre-existing lexicons used successfully in similar research so that results can be compared with existing literature thereby strengthening the reliability and validity of the feature selection process.

Content population: Neuendorf (2011) points out the distinction between the availability approach and the exposure approach to population definition in content analysis. The availability approach concerns the production of the media in question and the exposure approach concerns the effect on the consumers. This research focuses on the production side, investigating whether the gender of the politician affects what is written in newspapers. Therefore, the population of this research is all print media featuring Irish politicians.

Immersion in the message pool: Researchers should take a practical approach and "seek additional clues from a thorough examination of the pool of messages constituting the defined population" (Neuendorf, 2011). This will typically result in the emergence of key variables that might otherwise not have been detected. A guiding principle of the research design was the applicability of the methods to new corpora. Given this, although a deep knowledge of the population is called upon especially in interpreting the findings, the research attempts to avoid using specific knowledge about the samples in the feature extraction stage so as to maintain the generalisability of the methods to new corpora.

Human coding vs. computer coding: It is important to be aware of the advantages and disadvantages of hand coding or automatically coding texts in order to make the right decision between both these approaches (Neuendorf, 2011). Given that all text analysis in this study is automated, it is important then to consider what is being lost in terms of a human understanding of meaning and context. This factor is considered during the interpretation of the findings of the research.

Guidelines for Methodological Decisions

Neuendorf (2011) provided guidelines for methodological decisions in content analysis for gender research and these were used in this thesis. The following describes how this thesis addressed each of these concerns:

Unitising: The units of analysis in this research are definitive and are extracted using automated methods. The population consists of articles downloaded online. From this, some sub-samples were created filtering types within this population such as instances where the politician is mentioned in the headline. Each of these are defined explicitly and require no human intervention. Issues of reliability did occur when politicians' names were misidentified in searches. To ensure reliability, articles were manually checked by the researcher.

Sampling: Defining the population and samples are crucial decisions in research. Neuendorf (2011) critiqued many existing content analysis studies in gender research for lacking population and sampling methodologies, leaving the generalisability of findings undefined. Sampling issues are addressed in this chapter.

Measurement: Gender analysis focuses on latent meaning in language, which is normally the focus of qualitative research. However, Neuendorf (2011) argues that with thorough definitions in the content analysis codebook, based on reliable qualitative research, measurement of latent content though content analysis is possible. In this research latent meaning such as sentiment of text is operationalised using clear and established methodologies.

Training: This guideline emphasises the importance of training coders in manually coding content. While this research used automated approaches to analysing content, the process of extracting features from text, which is analogous to coding in content analysis, was tested in this research to ensure reliability.

Reliability: The principle of reliability in content analysis refers to the degree to which human coders agree on how codes are applied to texts. This is comparable to the selection of features form the text. The features selected in this study are computationally processed. While this does eliminate the inter-coder reliability issues, it introduces news issues of reliability when automated methods do not take context into consideration and meaning is erroneously attributed to text or text is tagged or parsed incorrectly. For this reason the features used in the text classification of this research remained as close to the actual text as possible so as to maximise reliability. Consistency was also ensured by using the same language processing algorithms such as tagging and stemming algorithms throughout each experiment so where errors such as tagging do occur, they are consistent throughout the experiments. Multiple machine learning methods are also used and the results compared. Gaining a consistent pattern among several different machine learning algorithms strengthens the reliability of results.

3.6 Population and Sample Selection

The population is coverage of Irish politicians in prominent Irish newspapers. No sampling frame for this population was available. The sampling strategy involved both purposive and convenience sampling. The politicians included in the research were purposively selected to best address the objectives of the research and then media content which featured these politicians was then selected based on the resources available to the researcher.

3.6.1 Selection of Ministers

A set of ministerial politicians in Ireland was selected which would best facilitate the identification of gender bias in newspaper coverage of them. It was important that differences between the politicians, other than gender, were minimised. This allowed for variables other than gender to remain as consistent as possible among the political subjects to be compared. While it is not possible to control for all the differences between politicians, the selection of the sample in this research controlled as much as possible for differences in political ideology and cabinet role.

The politicians selected were cabinet level politicians in government between 1997 and 2011. During this time, Fianna Fail remained the largest political party in government. This political continuity ensured that media content for a substantial length of time could be analysed while keeping variations in factors such as the political ideology of cabinet ministers to a minimum. Female cabinet ministers who held political office during that time were identified. Then male politicians who at some point between 1997 and 2011 held ministerial posts that the female ministers had held were identified. This strategy of selecting female cabinet politicians and then comparable male politicians within the same timeframe minimised differences, other than gender, among the set of politicians which would explain differences in media coverage of them. For example news coverage of a Minister for Finance was not included as no women held that post during the 15 years in question. Table 3.1 details the politicians selected and table 3.2 shows the combined duration spent by these politicians in Parliament broken down by gender. The following is an overview of how the sets of male and female politicians were matched:

Date in Position: Politicians were selected only during the 15 years from 1997 to 2011, when Fianna Fail were the main party in government.

Political Role: All female politicians were selected between 1997 and 2011. Male politicians who had held roles that the female politicians had also held during that time were also selected as part of the sample for this research. However, timeframes where they held roles that female politicians had never held were excluded. There were no roles that female ministers held between 1997 and 2011 that male ministers didn't also hold.

Political Party: The majority of the sample of politicians were from the main Fianna Fail political party. However, one female politician, Mary Harney was leader of the smaller coalition party in government, The Progressive Democrats (PDs). There were no male ministers from this party. Among the male politicians selected for this research was also a Green Party TD. The Green Party were also a minor coalition party. There was no equivalent female minister from the Green Party.

Male	**Female**
Batt O'Keeffe	Mary Coughlan
Brendan Smithe	Mary Hanafin
Brian Cowen	Mary Harney
Dermot Ahern	Mary O'Rourke
Éamon Ó Cuív	Síle de Valera
Eamon Ryan	
Frank Fahey	
Jim McDaid	
Joe Walsh	
John O'Donoghue	
Martin Cullen	
Michael Woods	
Micheál Martin	
Noel Dempsey	
Pat Carey	
Seamus Brennan	

Table 3.1: Sample of Irish Cabinet Ministers 1996-2011

Gender	Years
Male	84.1
Female	38.6

Table 3.2: Total Number of Years Male and Female MPs have been in Parliament

Table 3.4 details the names, roles and dates in office of the politicians included in this research. The dataset spanned the 28th, 29th and 30th governments of Ireland. Two of the female cabinet ministers, deputies Mary Harney and Mary Coughlan held the role of deputy prime minister. No male politician held this portfolio during this time. This imbalance may be addressed by the fact that among the male politicians was a future Taoiseach and party leader. In some cases there are title changes in the cabinet portfolio names and the composition of portfolios are changed. In such instances the closest match was identified.

Politician	Political Portfolio	Duration of Cabinet Post From	To
	28th Dáil		
Brian Cowen	Health and children	26/06/1997	26/01/2000
Dermot Ahern	Social, community and family affairs	26/06/1997	05/06/2002
Frank Fahey	Marine and natural resources	27/01/2000	06/06/2002
Jim McDaid	Tourism, sport and recreation	26/06/1997	05/06/2002
Joe Walsh	Agriculture, food and rural development	26/06/1997	05/06/2002
Mary Harney	Enterprise, trade and employment	26/06/1997	05/06/2002
Mary O'Rourke	Public enterprise	26/06/1997	05/06/2002
Michael Woods	Marine and natural resources	26/06/1997	27/01/2000
	Education and science	27/01/2000	05/06/2002
Micheál Martin	Health and children	27/01/2000	06/06/2002
	Education and science	26/06/1997	27/01/2000
Síle de Valera	Arts, heritage, gaeltacht and the islands	26/06/1997	05/06/2002
	29th Dáil		
Dermot Ahern	Communications, marine and natural resources	06/06/2002	13/09/2004
Éamon Ó Cuív	Community, rural and gaeltacht affairs	06/06/2002	13/06/2007
Joe Walsh	Agriculture and food	06/06/2002	29/09/2004
John O'Donoghue	Arts, sport and tourism	06/06/2002	13/06/2007
Martin Cullen	Transport	29/09/2004	13/06/2007
Mary Coughlan	Social and family affairs	06/06/2002	29/09/2004
	Agriculture and food	29/09/2004	13/06/2007
Mary Hanifin	Education and science	01/02/2000	06/06/2002
Mary Harney	Enterprise, trade and employment	06/06/2002	29/09/2004
	Health	29/09/2004	14/06/2007
Micheál Martin	Enterprise, trade and innovation	29/09/2004	14/06/2007
	Health and children	06/06/2002	29/09/2004
Noel Dempsey	Communications, marine and natural resources	29/09/2004	14/06/2007
	Education and science	06/06/2002	29/09/2004
Seamus Brennan	Social and family affairs	29/09/2004	14/06/2007
	Transport	06/06/2002	29/09/2004
	30th Dáil		
Batt O'Keeffe	Education and science	07/05/2008	22/03/2010
	Enterprise, trade and innovation	23/03/2010	20/01/2011
Brendan Smith	Agriculture, fisheries and food	07/05/2008	08/03/2011
Eamon Ryan	Communications, energy and natural resources	14/06/2007	23/01/2011
Éamon Ó Cuív	Social protection	23/06/2010	08/03/2011
	Community, rural and gaeltacht affairs	14/06/2007	22/03/2010
Martin Cullen	Arts, sport and tourism	07/06/2007	23/03/2010
	Social and family affairs	14/06/2007	06/05/2008
Mary Coughlan	Enterprise, trade and employment	07/05/2008	22/03/2010
	Education and skills	23/03/2010	08/03/2011
	Agriculture fisheries and food	14/06/2007	06/05/2008
Mary Hanafin	Social and family affairs	07/05/2008	22/03/2010
	Education and science	14/06/2007	06/05/2008
	Tourism, culture and sport	23/03/2010	08/03/2011
Mary Harney	Health and children	14/06/2007	20/01/2011
Micheál Martin	Enterprise, trade and employment	14/06/2007	06/05/2008
Noel Dempsey	Transport	14/06/2007	20/01/2011
Pat Carey	Community, equality and gaeltacht affairs	23/03/2010	08/03/2011
Seamus Brennan	Arts sports and tourism	14/06/2007	06/05/2008

Table 3.4: Politicians Included in Research, Cabinet Posts and Duration Held

Minister	Years in Office
Batt O'Keeffe	2.7
Brendan Smith	2.8
Brian Cowen	2.6
Dermot Ahern	7.2
Eamon Ó Cuiv	8.8
Eamon Ryan	3.6
Frank Fahey	2.4
Jim McDaid	4.9
Joe Walsh	7.3
John O'Donoghue	5.0
Martin Cullen	5.5
Mary Coughlan	8.8
Mary Hanafin	6.4
Mary Harney	13.6
Mary O'Rourke	4.9
Michael Woods	4.9
Micheál Martin	10.9
Noel Dempsey	8.6
Pat Carey	1.0
Síle de Valera	4.9
Seamus Brennan	5.9

Table 3.5: Number of Years Ministers Spent in Office

3.6.2 Irish Presidential Election Candidates 2011

A second corpus was developed to explore whether findings were consistent for election candidates and elected politicians. This corpus was composed of newspaper coverage of candidates for the Irish Presidential election in 2011. The campaign officially began on the last day nominations closed, on the 28th of September. The election was held on the 27th October 2011. The candidates in the election included two female candidates, Dana Rosemary Scallon and Mary Davis. The male candidates were Senator David Norris, Micheal D Higgins, Martin McGuinness, Gay Mitchell and Seán Gallagher.

3.6.3 Media Content Sampling Strategy

The availability of online sources of content for analysis can often dictate the sampling methodology for content analysis. Stempel and Stewart (2000) posited that this approach leads to convenience rather than representative sampling. However, while this is true at the moment due to the limited availability of computer readable archives of newspapers online, as this availability increases, it should be possible to collect more representative samples. This will broaden the range of sampling strategies available to researchers.

The sample of media content gathered was limited by the data available. For the time spanning the 1997 to 2011, text from the Irish Times was available in computer readable form. Text from the Irish Independent and Sunday Independent was also available for the 30th and 31st Dáil terms. As can be seen in Irish newspaper readership figures detailed in table 3.6, these are the largest selling newspapers in Ireland. The Sunday Independent is the Sunday edition of the Irish Independent and was therefore included in the corpus as the one publication.

A broader range of newspapers was available for the presidential election in 2011. Local papers were also included in this corpus. The newspaper titles are detailed in table 3.7

	Combined Readership	Print Readership	Digital Readership
Daily Titles			
Irish Independent	649,000	570,000	119,000
The Irish Times	390,000	337,000	106,000
Irish Examiner	224,000	21,0001	20,000
Irish Daily Star	391,000	387,000	8,000
Irish Daily Mirror	274,000	269,000	9,000
Irish Sun (Mon-Sat)	344,000	337,000	14,000
Irish Daily Mail	230,000	206,000	31,000
The Herald	315,000	308,000	10,000
Sunday Titles			
Sunday Independent	938,000	903,000	71,000
Sunday World	785,000	780,000	12,000
Sunday Business Post	147,000	140,000	14,000
The Sunday Times	409,000	388,000	50,000
Irish Sunday Mirror	181,000	180,000	3,000
Irish Mail on Sunday	354,000	351,000	11,000
Irish Sun (Sunday)	251,000	249,000	9,000

**(Source: National Newspapers of Ireland)*

Table 3.6: Irish Newspaper Readership

Media Sources
The Irish Times
Irish Independent
The Irish News
RTE News online
Sunday Independent
The Mirror
BreakingNews.ie
Irish Examiner
Kerryman
The Argus
Western People
Belfast Telegraph
Sunday Mirror
Waterford News and Star
Sunday Life
Wexford People
Wicklow People
Carlow People
Carlow Nationalist
Enniscorthy Guardian
Fingal Independent
Gorey Guardian
New Ross Standard
Sunday Business Post
Drogheda Independent
Corkman
Wexford Echo
Bray People
Business and Finance Magazine
Business World (Digest)

Table 3.7: Media Sources for 2011 Irish Presidential Election Coverage

3.7 The Corpus

Newspaper content was gathered from the online news aggregator, Lexis Nexis, which allowed articles to be downloaded in html form at a rate of 500 articles per search. Media content was downloaded which featured the politicians and election candidates in the articles. The criteria in gathering the data were that the articles would feature the individuals in question. The research question in this thesis concerned coverage of a minister. The articles examined did not need to be about the politician. It was important that all coverage, where the politicians were featured, was identified. However, in some articles politicians were named incidentally at the end of articles but did not feature prominently in them. This produced a large collection of articles where the coverage incidentally mentioned a politician but could not be described as coverage of a politician. To address this, various methods of refining the text were explored (Figure 3.1).

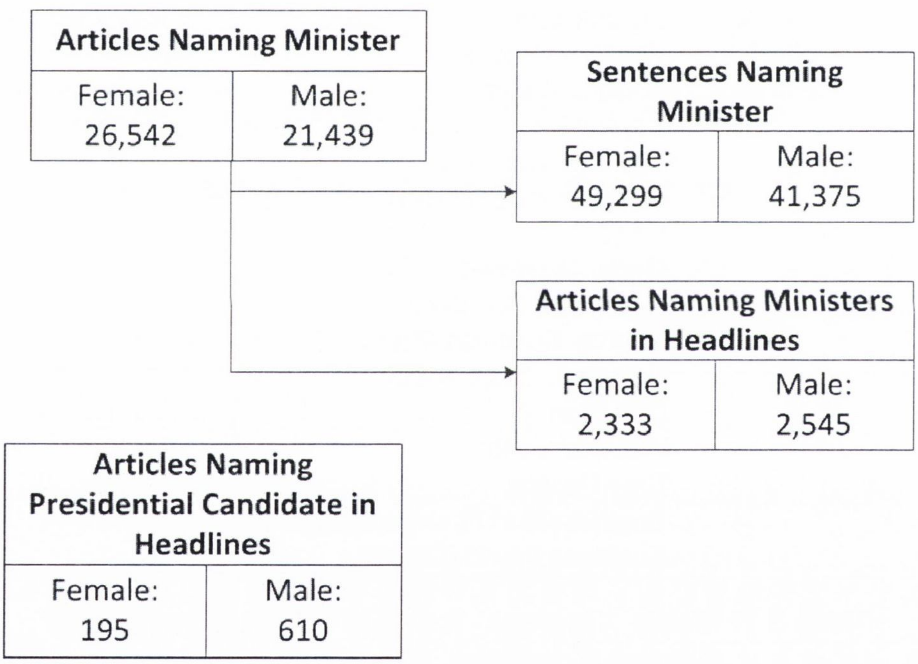

Figure 3.1: Newspaper Content Corpus

Minister Named in Article: This corpus was created by searching Lexis Nexis for articles which named ministers listed in Table 3.4 while they held cabinet posts. This resulted in a large dataset which included some articles naming ministers incidentally. In some examples multiple ministers were mentioned at the end of articles but had little relationship to the subject of the article. This necessitated refinement of the corpus to isolate coverage of the ministers in question. Sentences were extracted from this corpus which named the ministers. This reduced the size of the corpus while also ensuring the media content was relevant to the research question. Table 3.8 details the number of articles were one of the sample of cabinet ministers was mentioned in an article, during the time they held the portfolios as outlined in table 3.4.

Politician	Number of Articles
Batt O'Keeffe	1315
Brendan Smith	1014
Brian Cowen	968
Dermot Ahern	765
Éamon Ó Cuív	1980
Eamon Ryan	2322
Frank Fahey	102
Jim McDaid	427
Joe Walsh	496
John O'Donoghue	1234
Martin Cullen	2243
Mary Coughlan	4128
Mary Hanafin	2040
Mary Harney	15055
Mary O'Rourke	3611
Michael Woods	342
Micheál Martin	2512
Noel Dempsey	3996
Pat Carey	300
Síle de Valera	1708
Seamus Brennan	1423

Table 3.8: Newspaper Articles Mentioning One or More Ministers

Minister Named in Headline: Another corpus was created by extracting articles which named one of the politicians in the headline. Because the ministers were named in the headlines of the articles, the content of the articles was more likely to be about the minister or they were to feature prominently in the stories. This resulted in a corpus which included the full text of articles which prominently featured a cabinet minister. It also produced a corpus which was of a size that could be manually checked for accuracy. For this reason, this corpus was used as the primary corpus upon which to refine the research design. The distribution of articles among the sample of politicians is given in Table 3.9.

Politician	Number of Articles
Batt O'Keeffe	230
Brendan Smith	46
Brian Cowen	114
Dermot Ahern	67
Éamon Ó Cuív	169
Eamon Ryan	170
Frank Fahey	56
Jim McDaid	102
Joe Walsh	201
John O'Donoghue	63
Martin Cullen	231
Mary Coughlan	234
Mary Hanafin	331
Mary Harney	1490
Mary O'Rourke	185
Michael Woods	124
Micheál Martin	433
Noel Dempsey	352
Pat Carey	10
Síle de Valera	93
Seamus Brennan	177

Table 3.9: Newspaper Articles Mentioning Minister in Headline

Presidential Candidate Named in Headline: The corpus pertaining to the 2011 Irish Presidential Election was created based on a search which identified articles where one of the candidates was named in the headline. This method was selected due to the success of this approach in experiments on the Irish Parliamentary corpus which also identified articles where ministers were named in the articles. Given the short duration of the election, media coverage was expanded to include local and online media as restricting to national newspapers significantly reduced the corpus size. The distribution of election coverage is presented in Table 3.10

Politician	Number of Articles
Dana Rosemary Scallon	156
David Norris	110
Gay Mitchell	61
Mary Davis	42
Martin McGuinness	191
Michael D. Higgins	92
Sean Gallagher	156

Table 3.10: Presidential Candidate Mentioned in Headline of Articles

3.8 Text Classification Experiments

3.8.1 Text Classification Experiment Design

The text classification process was explained in section 3.5.1. How this process was applied in this research is illustrated by Figure 3.2. Once the features were extracted from the documents and represented as an array of vectors, a machine learning algorithm developed predictive models indicating what features of the texts indicated the gender of the politician featured in the article.

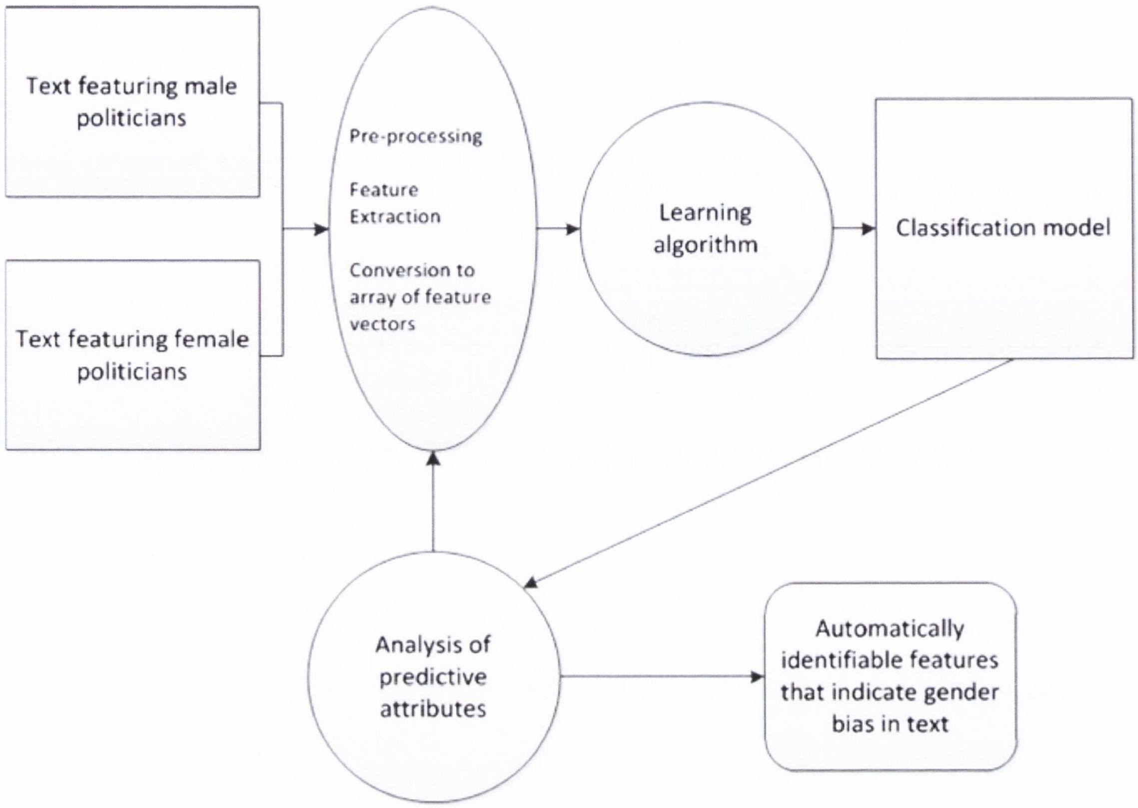

Figure 3.2: System Architecture: Text classification

The rationale for using the gender of the politicians featured in the articles as the category which the classifier was to predict was that this would highlight differences in the newspaper coverage of male and female politicians. Greater differences in coverage, depending on the features used, could then indicate greater differences in the way newspapers cover male and female candidates. There may be circumstantial differences in the coverage of male and female politicians which could not be attributed to gender bias. These differences are considered during the feature extraction process and the final analysis of the discriminative features uncovered in the text classification process.

The texts were preprocessed to remove some attributes that indicated the gender of the politician or features of the text such as dates and numbers which could be used by the machine learning algorithm to identify the gender of a politician featured in an article, but which would have no meaning in relation to the topic of gender bias. These attributes included a politician's name, political party name, dates, political and professional titles, pronouns and numbers. Details of the pre-processing of data in this research is described in section 3.8.2.

Various features were then extracted from the articles. These included linguistic features of the text in the articles along with structural features of the articles. The feature set design was based on previous work in text classification studies and work on gender bias in the coverage of politicians in the media. Section 3.8.4 details the features extracted from the text and the rationale behind them.

One the features to be extracted from the articles were identified, these were represented in a form that could be used by the machine learning algorithms in Weka. Weka uses the attribute-relation file format (ARFF). An illustrative example from Witten et al. (2011, p.53.) is presented below.

```
@RELATION iris
@ATTRIBUTE sepallength  NUMERIC
@ATTRIBUTE sepalwidth   NUMERIC
@ATTRIBUTE petallength  NUMERIC
@ATTRIBUTE petalwidth   NUMERIC
@ATTRIBUTE class        {setosa,versicolor,virginica}

@DATA
5.1,3.5,1.4,0.2,Iris-setosa
4.9,3.0,1.4,0.2,Iris-setosa
4.7,3.2,1.3,0.2,Iris-setosa
4.6,3.1,1.5,0.2,Iris-setosa
5.0,3.6,1.4,0.2,Iris-setosa
```

```
5.4,3.9,1.7,0.4,Iris-setosa
4.6,3.4,1.4,0.3,Iris-setosa
5.0,3.4,1.5,0.2,Iris-setosa
4.4,2.9,1.4,0.2,Iris-setosa
4.9,3.1,1.5,0.1,Iris-setosa
```

The entire corpus of articles was represented as a set of instances. Each article corresponded to one instance. Each instance included an identifier as to whether the politician featured in the article was male or female, along with values representing attributes or features of the text as an array. The methodology used for transforming the features of newspaper articles into an array of vectors is described in section 3.8.5. Using 10-fold cross validation, the machine learning algorithms developed models to predict the gender of the politician featured in the articles based on attributes of the articles. Details of each step in the machine learning process are detailed in the following sections.

3.8.2 Data Pre-Processing

Once the newspaper articles were downloaded from Lexis Nexis, they were converted from html to text documents. Duplicates were then removed by automatically comparing the texts of the articles in the search results from each politician and upon finding a match, deleting one of them. Duplicate articles occur in Lexis Nexis when newspapers issue reprints of articles.

Since text classification was used to identify features that would indicate gender bias by identifying differences in news coverage of male and female politicians, it was important that those features which would reveal the gender of a politician but could not be attributed to gender bias, were removed. For example, a classification model could gain high accuracy from predicting the gender of a politician from the pronouns used to refer to them. However, use of these pronouns would not indicate gender bias. Such indicators were therefore removed so the classifier would not use these features to predict the gender of a politician. These kinds of words were replaced in the text with neutral terms and the rationale for this is detailed below. Removing identifying words was an iterative process based on analysing the results of initial experiments. Once classifiers were run, the discriminative factors used were analysed and those features of texts which were allowing a classifier to identify the gender of a politician but based on factors other than gender-bias, were removed.

Politician's Name: The names of the politician were removed from the content of the articles. This was in order that the classifiers did not have information on which individual was featured in the articles. The names were replaced in the text by a generic reference to a politician's name. Whether the first name and/or surname were used to refer to the politician was included in the mark-up of the text.

Party Name: Given that there were some minority coalition parties in government, namely The Progressive Democrats and The Green Party, on initial runs of the classifier, the party the politician was associated with was used to identify the gender of a politician. Mentions of different political parties were therefore updated with a generic name for a political party.

Dates: At various points in time, certain stories associated with certain politicians come to prominence. For this reason, initial experiments showed that the classifier used information about when the story occurred to indicate the probability that the story was about a male or female politician. Given that such information was not indicative of gender bias, individual dates were removed and replaced with a generic word.

Political and Professional Titles: There were variances in the sample for this research in the political titles held by some politicians. Mary Harney and Mary Coughlan were both Deputy Prime Ministers at some point between 1997 and 2011. David Norris held the title of Senator which differentiated him from other electorial candidates in the Presidential Election. Dr. was also used as part of naming one of the cabinet ministers in the sample. Given that this information was used by the classifier to indicate the gender of the subject, having the same effect as including their first name. Titles like this were treated as part of the naming of the politicians. Replacing the individual titles with generic descriptions of titles allowed for gender differences to be identified without being linked with specific individuals.

Pronouns: Gender specific pronouns in the text allowed the classifier to achieve almost 100 percent accuracy in predicting the gender of a subject. However, this language is not indicative of gender bias and so they were removed from the text.

Numbers: Numbers were often used by classifiers to indicate the gender of a politician. These often referred to budgetary issues. Since there was no latent meaning in such numbers regarding gender bias, these were removed and replaced with a word which indicated a number.

Meta-information about the articles in the corpus was extracted automatically from the text. This aided the organisation of the articles into smaller corpora. Some of this information was also used as features in the text classification experiments. Some of the meta information was included as part of the documents downloaded from Lexis Nexis and other information was inferred.

Names of the individual politicians were anonymised in the text and were instead added to the filename. The dates of the articles were extracted to identify which parliament term the article was written in. Other meta-data extracted included newspaper section and article length. All metadata that was part of the document content downloaded from Lexis Nexis was then removed from the text content itself.

All newspaper articles were classified according to whether the politicians featured in them were male or female. There were instances where more than one politician was featured in an article. In these cases the same article was entered into the corpus multiple times with labelling of each corresponding to one of the politicians.

Due to the fact that the names of the politicians are not all unique in Ireland, articles were returned which featured individuals other than the politicians in question. To filter these, manual checks were done on the corpora to ensure the articles featured the correct individuals. For the two smaller corpora which involved checking that the politician's name was in the headline, it was possible to manually check every article.

3.8.3 Sub-Corpora

In order to explore different aspects of the newspaper coverage of ministers, sub corpora were generated from the corpus containing coverage of cabinet ministers where these ministers were named in the headline. The criteria according to which they were categorised included the section of the newspapers, the Dáil (Irish National Parliament) term and the ministerial role of the cabinet minister mentioned. Table 3.11 shows the sub-corpora that were generated.

Sub-Corpus	Criteria
Paper Section	Opinion and Analysis
	Front Page
	Letters
Ministerial Role	Agriculture
	Communications
	Community, Gaeltacht Affairs
	Eudcation
	Enterprise
	Health
	Natural Resources
	Social and Family Affairs
	Sport
	Transport
Dáil Term	28th Dáil
	29th Dáil
	30th Dáil

Table 3.11: Description of Sub-Corpora

The style, tone and topics of articles can vary according to the section they appear in. For this reason, three sections of a newspaper were identified and articles which featured in these sections were extracted. The front page was considered a section in its own right because of the importance attached to appearing in an article on that page. The opinion and analysis section was also selected as the style and tone can be quite different to other sections of the paper. Finally, the letters section provides the most contrast to news articles as they are authored by readers and so not subjected to as strict editorial guidelines..

The distribution of articles which feature politicians while they held certain political roles is varied. To control for such differences, sub-corpora were created which grouped articles according to the cabinet post of the politician.

A longitudinal analysis of trends in gender bias was done by isolating articles corresponding to the three Dáil terms. These where the 28th, 29th and 30th

Irish parliaments. This facilitated analysis of any changing trends in language in coverage of male and female politicians from 1996 to 2011.

3.8.4 Feature Extraction

Determining what features to extract from the newspaper articles was informed by findings from previous studies in text classification and also from previous research on gender bias in language. As no previous research has directly addressed the issue of gender bias using text classification, this research initially explored a wide range of approaches to feature extraction. Several experiments were conducted to identify the best features and as the research progressed, the feature sets extracted from the texts were refined.

A guiding principle in selection of the features was to ensure that the features remained as close to the text as possible. This was done for two reasons. Firstly, inclusion of subjective readings of texts as features could leave the analysis open to criticisms of researcher bias. Secondly, as explained by Stamatatos (2009), the more detailed the text analysis that is required in extracting features, the less accurate (and the more noisy) the resulting representation of the text.

Stamatatos (2009) presented a classification of the kinds of features used in text classification studies including include lexical, character-based, syntactic, semantic and topic specific text attributes. This framework was used in this research as a basis upon which to develop the list of feature types used in this research (Table 3.12).

Features
Word unigrams
Word unigrams (stemmed)
Word bigrams
Word trigrams
Vocabulary Richness Measures
Parts-of-Speech tags (unigrams and bigrams)
Sentiment in articles
Stop words
Descriptive words associated with politicians
Actions associated with politicians
Specialised dictionaries
Gender of article's author
Section of newspaper
Page number
Article word count

Table 3.12: Features Extracted from Newspaper Articles

Unigrams: Using words that appear in a text as features is the most common approach to representing texts text classification studies (Sebastiani, 2002). However, it does not include contextual information such as the placement of the words or their relationship to other words in the document. This approach considers text as a collection of words and ignores context, word order and meaning.

Stemming words in text has been found to reduce dimensionality and is regularly used in text classification. This involves identifying the root of words and using these as features instead of the full version of the words. In this way, different words with related meaning can be grouped and considered one word for the purposes of text classification. For example, the words teacher and teaching could be represented as a single feature such as teach. This approach to representing text was explored in this research using the Snowball stemmer (Porter, 2001).

N-Grams: Representing texts as a collection of word sequences is a way to capture some of the context within which words occurs. It can capture sequences of terms with specific meaning. The value of n greatly increases the dimensionality of the data however. The most common values of n in text classification studies is 2 (bi-grams) and 3 (tri-grams). These are the values used in this research.

Vocabulary Richness Measures: TTR (type-token ratio) and Hapax Legomena were used as measures of the complexity of vocabulary. TTR is the number of unique tokens divided by the total number of tokens. Hapax Legomena refers to the number of words occurring only once. These features uncover differences in the complexity of the language used.

Sentence Length: Words in articles are tokenised and then broken into sentences. The length of these sentences is calculated and the text is represented as an array of vectors where each vector is a word count of each sentence.

Parts of Speech Unigrams and Bi-grams: Parts-of-speech have been successfully used as features in text classification studies. In this research words in the documents were tagged using a part of speech tagger (Bird, 2006). Two word sequences of tags were then identified to form parts-of-speech bigrams.

Sentiment: The approach to evaluating sentiment used in this research was that developed by Argamon, Whitelaw, Chase, Hota, Garg and Levitan (2007). This approach involved extracting words that appear in a sentiment lexicon and updating them with the sentiment value in the lexicon. Words that are not sentiment bearing were removed. This lexicon was designed to capture the sentiment expressed in text. The lexicon used is detailed in table B.1.

Stop words: As discussed in Chapter 2, many approaches to text classification that use this approach exclude function words. These are words which link concepts in a sentence. In topic oriented text classification studies these are often removed to reduce the dimensionality of the feature set. In classification based on writing style however, stop words have been proven to be most useful (Argamon and Levitan, 2005). The concept of stop words was introduced by Luhn (1958) who posited that "the frequency of word occurrence in an article furnishes a useful measurement of word significance". In this research, stop words are are extracted by identifying the most commonly used words in the text. The method for identifying stop words was outlined by Manning et al. (2008, p.27) as follows: "sort the terms by collection frequency (the total number of times each term appears in the document collection), and then to take the most frequent terms, often hand-filtered for their semantic content relative to the domain of the documents being indexed, as a stop list, the members of which are then discarded during indexing". This is the approach adopted in this research. In some experiments stop words were excluded from the list of feature words and in others only stop words are are included as features.

Verbs and Adjectives Associated with Politicians: Adjectives were used as features in this research to examine differences in how male and female politicians were described. Two approaches to extracting the features were explored. One method involved extracting all adjectives from each document. The second approach extracted only adjectives from sentences mentioning the politicians. The rationale for extracting adjectives as features is based on previous research which detailed differences in how male and female politicians are described in the media

(Mills, 2002; Sreberny-Mohammadi and Ross, 1996).

Specialised Dictionaries: Developing lexicons and using these to extract features from text is an approach that is particularly common in sentiment analysis studies (Whitelaw et al., 2005). To date, there is no existing lexicon that is specifically designed to identify gender bias. This research used lexicons which have been successfully used in related studies (Ahmad et al., 2011). This approach involved using word categories from The Harvard General Inquirer system (Stone et al., 1966).

The word categories which were used as lexicons were those pertaining to power, politics and action. Features are extracted based on the political lexicon from The Harvard General Inquirer to explore whether male or female politicians are more associated with political or policy issues. Previous literature suggested that male politicians are discussed more in relation to political or policy matters (Brikse, 2004; Carroll and Schreiber, 1997; Devitt, 1999; Kahn, 1994, 1996). Power words were used as features to explore whether there is a difference in the extent to which politicians are portrayed as having the ability to influence government policy (Ahmad et al., 2011; Gidengil and Everitt, 2003; Ross and Sreberny, 2000). Action words from the General Inquirer were used as features in order to test whether there were differences in transitivity as outlined by Mills (2002).

Gender of the Author: The gender of the article's author was extracted and used as features to explore whether the gender of journalists affect the coverage of male and female politicians. As many of the names in the Irish national newspapers are in the Irish language, the names were extracted and a lexicon was created manually to identify the gender of the authors. For some articles, there were multiple authors. In these cases the article was tagged with an identifier representing the multiple genders and since the main author normally comes first, the ordering of the individuals was represented in the feature.

Editorial Decisions: The placement of an article and its' word-count convey a level of ascribed importance of a topic in newspapers. Front-page articles, for example, reflect what the newspapers deem of most importance. Longer articles are also an indicator of the prominence a newspaper is giving to a story. Differences in the levels of prominence a newspaper affords articles featuring male or female politicians could therefore indicate gender bias. To capture editorial decisions about the importance of a story, the length of articles and the sections of the newspapers were extracted from the texts and represented as features. This also highlights whether there are differences in what supplemental sections of the newspapers politicians appear in. For example, if women appeared significantly more often in a style section of a newspaper this could be evidence of the kind of gender bias described by Ross and Sreberny (2000) and Devitt (1999) where a female politician's dress and personal style is a focus of attention more then for their male counterparts.

3.8.5 Text Representation

Text documents were transformed into an abstract representation of the articles by extracting certain features of the texts and representing them as an array of vectors. Three commonly used methods of representing features of the media content numerically were explored in this research to determine which was the optimal approach for detecting gender bias. A minimum frequency threshold is often used to reduce the dimensionality of the data for machine learning algorithms. Diermeier et al. (2012) for example used a minimum frequency of 50 in their study to classify texts according to political ideology. Various options for setting a minimum frequency are explored in this research. The following is a description of these methods of representing features of a text:

Bag of Words: A bag-of-words (BOW) representation of features is an array consisting of a count of the frequency of occurrence of each feature. This is the most widely used approach to representing texts (Diermeier et al., 2012).

Boolean: A boolean representation denotes whether a feature occurs in a text or not. Information on the frequency of the occurrence is not included.

Tf-idf: It is often useful to reduce the effect in a corpus of words that occur frequently throughout the entire corpus. This is done by a range of approaches categorised as Tf-idf weighting (Wu et al., 2008). The approach used in this research is tf-idf = tf (term frequency) x idf (inverse document frequency) (Witten et al., 2011, p581).

3.8.6 Classification Algorithm

Multiple machine learning(ML) algorithms were tested to identify the best performing algorithm. These included a J48 decision tree learner, a support vector machine and Naïve Bayes algorithm. The Weka toolkit was used for classification experiments (Holmes et al., 1994). As iterations of experiments were completed, the methods used in subsequent experiments were refined. The following is an overview of the classification algorithms used:

J48 Decision Tree: A decision tree learner generates from examples, a predictive model which identifies the class that texts belong to. The model is composed of a tree with nodes corresponding to features of the text. The branches emerging from these nodes correspond to options regarding how these nodes appear in the text. Sequences of nodes denote a relationship between features in the text in predicting the class.

Naïve Bayes: Naïve bayes classifiers calculate, for each feature of a text, the probability that that feature is associated with a particular class. Unlike the J48 decision tree, each feature is viewed independently.

Support Vector Machine (SVM): Support vector machines generate a representation of the features of the texts and generate two separate hyper-planes where the margin between the two is maximised. Support vector machines have been identified as one of the most effective text classification and feature selection methods (Diermeier et al., 2012).

Machine learning algorithms assume an even split between classes. In most of the experiments in this research, this even split did occur. However, in the sub-corpora, in some experiments there was an imbalance in the data set. There are many approaches to dealing with such imbalance of which under-sampling is a popular method (Akbani et al., 2004). This involves randomly sampling from the majority class, the same number of instances as in the minority class. The disadvantage of this approach is that potentially valuable training data is discarded. However it has been utilised effectively in previous studies (Price et al., 2003) and was therefore was selected as the approach used in this research.

3.8.7 Evaluation of Classification Algorithms

Stratified 10-fold cross-validation was used to evaluate the predictive models generated by the classifiers. In this method, a dataset is randomly separated into 10 datasets of equal size and class distributions. For each of the folds, the classifiers learn from all but one of the 10 datasets and the model is tested on the remaining. Overall accuracy or cross validation is found by averaging the results of the tests on each of the 10 datasets. This approach to evaluating text classifiers has been widely used in similar research (Durant and Smith, 2007; Pollak et al., 2013; Somasundaran and Wiebe, 2010; Wilson et al., 2004).The best classifier of those tested in this research is determined by the highest predictive accuracy.

3.9 Evaluation of Results

Evaluation of the results of this research was conducted in two phases corresponding to the two research questions of this thesis. The first phase of evaluation was focused on the validity of the method of using automatic text classification in identifying gender bias in text. Central to this approach was an evaluation based on the interpretability of the machine learning experiments. The second research question concerned evaluation of the patterns of difference identified by the classification algorithms and verification of whether the differences identified constituted gender bias.

3.9.1 Interpretability

Interpretability, along with prediction accuracy has been defined as one of the fundamental performance criteria for machine learning (Ruping, 2005). Interpretability refers to the understandability of why the model is accurate and how the accuracy is deduced from the data. It is a criterion that is particularly important when data mining is being used to address social science questions were predictive accuracy of a model is not the sole goal of the research.

Interpretability is managed in machine learning through feature selection, instance selection and analysis of the models generated by the machine learning algorithms. Using dictionary based approaches to feature selection from the articles enables the interpretability of the result to be maximised. This approach is explored in this research using specialised dictionaries. Interpretation of the results is also influenced by the instances that are selected for training the models. To do this, sub-corpora were created and models generated from these instances. Finally, the support vector machine learning algorithm was analysed to identify what features were associated with which category.

3.9.2 Evaluation of Methodological Approach

The findings of the first phase of evaluation addressed the first set of research questions in this thesis. These concerned exploring how text classification could be used to identify gender bias in text. Table 3.13 presents a summary of each research question and the corresponding evaluation method used to address them.

Research Question	Topic	Evaluation Method
Q1(a)	Machine Learning Algorithm	Highest predictive accuracy using 10-fold classification
(b)	Feature Representation	Highest predictive accuracy and high ranking of features suggesting gender bias
(c)	Feature type	Most interpretable features in relation to identifying gender bias

Table 3.13: Evaluation of Text Classification Results

1. How can techniques from text classification be used to explore differences in the coverage of male and female politicians in order to identify gender bias?
    (a) What machine learning algorithm is suitable for identifying gender bias?
    (b) What approach to feature extraction is most informative in identifying gender bias in text?
    (c) Which is the optimal approach to representing features in text classification experiments to identify gender bias?
2. What differences in Irish newspaper coverage of male and female politicians indicate gender bias?

Addressing research question 1(a) involved identifying which of the three machine learning algorithms tested yielded the highest predictive accuracy in terms of identifying whether an article featured a male or female politician. The predictive accuracy was ascertained using 10-fold cross validation as outlined in Section 3.8.7.

Research question 1(b) concerned the best method of representing features. This was assessed by identifying the representation which yielded highest predictive accuracy and also that which best identified features that indicated gender bias.

Addressing question 1(c) involved evaluation of the discriminative features identified by the classification algorithms. Using some features may produce high predictive accuracy but the features may not be interpretable in the context of the research topic. For instance, some feature types may distinguish between articles featuring male or female politicians. However, those features may not be useful in deciphering whether or not gender bias exists in those articles. The output of this evaluation is a list of feature types which are useful in identifying gender bias.

3.9.3 Evaluation of Features for Evidence of Gender Bias

The second phase of evaluation addressed the second research question in this thesis. In this, the discriminative features in the feature sets identified as suggesting gender bias are examined. Table 3.14 presents an overview of the second research question and and the method used to address it.

Research Question	Topic	Evaluation Method
Q2	Features suggesting gender bias	Interpretation of discriminative features based on previous theories on gender bias, researcher knowledge of political context and analysis of the context in which the terms were used in the corpus

Table 3.14: Evaluation of Discriminative Features

Once the feature types which proved to be interpretable regarding identification of gender bias were identified (Question 1(c)), the discriminative features of these types were then examined to extract those which suggest gender bias.

The first step in this process was to analyse what features were associated with each category following an approach outlined by Diermeier et al. (2012). As the support vector machine learning algorithm was the best performing algorithm, these were the results used to analyse the features of texts which discriminated between articles featuring male and female politicians.

The features were ordered according to the weighting of each attribute assigned by the linear SVM (Chang and Lin, 2008; Guyon et al., 2002). The support vector machine learning algorithm ranks each feature according to its association with one category of texts. These can be seen in the support vectors that are generated by the learning algorithm. The highest ranking of the features are those which best discriminate between the texts. Some of the features are given positive values and others are negative values. These polarities correspond to one of the categories. Diermeier et al. (2012) for example examines the top 20 features. This research begins by examining the top 100 features.

Table 3.15 shows an extract of ordered features according to each text category. The features marked in bold were those identified as indicating gender bias. These were further analysed to explore the context in which the words occurred in the text. This approach to analysing the features involves researcher interpretation of the patterns identified by the classifiers based on previous academic literature and also contextual knowledge of political events. This involved identifying features which corresponded to themes identified in previous academic literature as indicating gender bias. These themes are outlined in Table 2.5 and Table 2.1. The justification for selectively analysing features which point to gender bias, as decided by researcher interpretation, rather than analysing each feature is that a goal of this research is to identify whether there is evidence of gender bias. Analysing the use of other features which do not indicate gender bias does not address the research question as to whether there is evidence of gender bias. In taking this approach, some features which do occur as a result of gender bias may be missed. However, analysing some features and identifying whether these are evidence of gender bias still addresses the research question regarding whether some features of newspaper coverage of male and female politicians could be attributed to gender bias. Showing that some other features don't indicate gender bias does not address the research question. For this reason, this approach of selecting features to explore further, based in researcher interpretation of features that seem to indicate gender bias, was deemed suitable.

Some features, which do not constitute gender bias but may have a cancelling effect on those that do, could remain unexamined. To address this issue, the researcher interpretation also included some counter examples to features identified as potentially indicating gender bias. For example, from the list below childcare was associated with female politicians, which based on previous findings, points towards gender bias existing in the association of issues with politicians of either gender. However, 'children' was associated with male politicians so a similar association could be made for. For this reason, both words were identified and the context in which they were used was analysed.

Female	Male
hse *(health service executive)*	transport
...	
define	natural
first-name-last-name	for
....	
childcare	community
...	
argued	**children**
...	

Table 3.15: Corpus Naming Ministers in Headlines - SVM Classifier Unigram Features

Concordance Analysis

Where analysis of the discriminative features was not sufficient to draw conclusions regarding gender bias, these were analysed further to verify how they were used in the corpus. The method used to analysed the context in which the features were used was concordance analysis (Pollach, 2012). Concordance analysis involves searching the corpus for uses of a word and extracting it along with a certain number of words that appear on each side of the word in the corpus. This process results in a list of extracts of the use of the words and the context in which they were used. These were then analysed to ascertain whether certain discriminative features were a result of gender bias in the coverage of female politicians.

3.10 Conclusion

This chapter presented the methodology for this research. The methodology involved using a computational approach to analysing text in order to identify gender bias in the newspaper coverage of Irish cabinet ministers and Presidential election candidates. The computational technique used in this research was text classification. This process was described in this chapter and how it is applied in this research to analysed the language used in newspaper articles was presented.

Due to the fact that to date, no previous academic literature has used text classification to identify gender bias in the coverage of female politicians, a broad scope of approaches to text classification was explored based on methods outlined in studies on related subjects. These studies were outlined in Chapter 2. As the experiments were conducted and results were evaluated the scope of these experiments were refined.

The attributes of newspaper coverage which were identified by the classification algorithms as distinguishing between articles featuring male or female politicians were interpreted to identify whether they could be attributed to gender bias. Some discriminative features were identified as gender bias based on a combination of findings from previous literature and researcher knowledge of the political context at the time the newspaper articles were written. Verification of whether some discriminative features constituted gender bias required further analysis of how they were used in the text. This was done using analysis of concordance extracts capturing the context within which the terms were used in the newspapers.

Chapter 4

Analysis and Findings

4.1 Introduction

This chapter presents the main results and findings of the research. The use of automatic text classification in identifying gender bias in text is evaluated. Differences between the coverage of male and female politicians are then examined to identify evidence of gender bias.

The research questions addressed are set out below. The main objective of the first research question concerns the methodological approaches used. Addressing this involves evaluating how accurately machine learning algorithms can identify the gender of a politician in an article and what attributes of articles are used to do this. The approaches which best highlight gender bias in the articles are identified. The second objective is to ascertain whether gender bias exists in the newspaper coverage of Irish political figures and if it does, what the nature of it is. This involves a qualitative analysis of the text attributes or features that machine learning algorithms have found to be associated with politicians of either gender. Both questions are addressed in discrete sections of this chapter. The following are the research questions addressed:

1. How can automatic text classification be used to explore differences in the coverage of male and female politicians in order to identify gender bias?

   (a) What machine learning algorithm is suitable for identifying gender bias?

   (b) Which is the optimal approach to representing features in text classification experiments to identify gender bias?

   (c) What approach to feature extraction is most informative in identifying gender bias in text?

2. What differences in Irish newspaper coverage of male and female politicians indicate gender bias?

Three corpora were analysed in this research. This chapter is structured to present the findings of each corpus separately. Figure 4.1 illustrates how the results address each research question. This process was repeated for each corpus. The concluding section of this chapter synthesises the results of all experiments to present the overall findings of this research in relation to each research question.

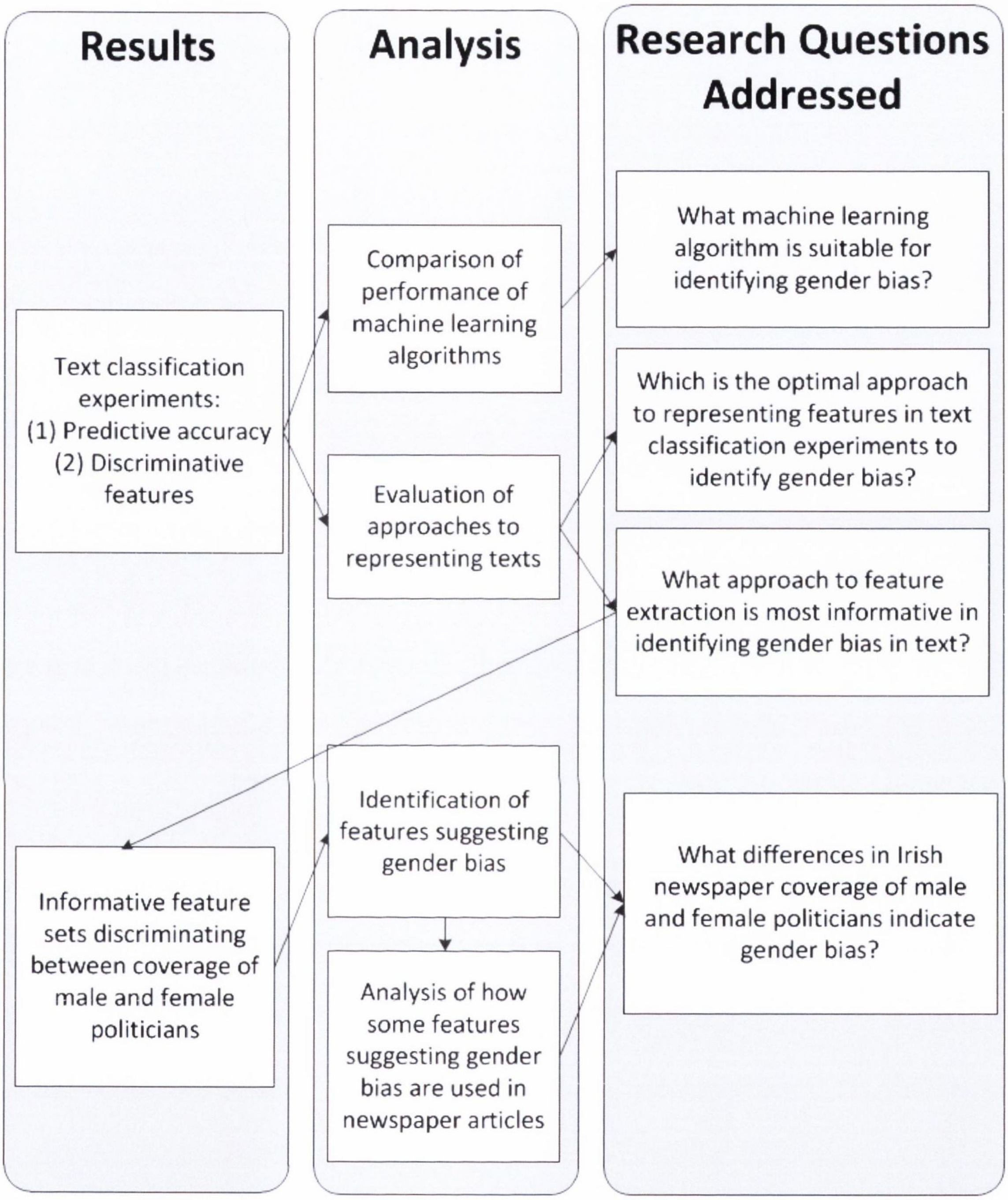

Figure 4.1: Process of Analysis

4.2 Corpus Analysis of Articles Naming Ministers in Headlines

The first corpus analysed was that based on the search criterion that a minister's name is mentioned in the headline of a newspaper article. The purpose of this was to ensure that the minister was central to the story in the article. The corpus was comprised of 4,878 articles. A full description of the corpus is presented in chapter 3, Section 3.7.

An extensive range of approaches to text classification were implemented in analysing the corpus (Section 3.8.3). These were evaluated and the most informative of them in terms of detecting gender bias were identified. The scope of subsequent experiments in this thesis were narrowed based on these findings. Sub-corpora were developed in order to isolate the effect of certain variables on the coverage of politicians. These variables included the newspaper sections the articles appeared in, when the articles were written and the cabinet positions the ministers held at the time.

4.2.1 Evaluation of Feature Extraction, Machine Learning Algorithms and Feature Representation

This section evaluates whether text classification can be effectively used to detect gender bias in text and how the techniques can best be implemented. Some key issues to address in using text classification techniques to analyse text are what machine learning algorithm to use, what features to extract and how to represent them. Table 4.1 presents an overview of the main findings in relation to each of these issues.

Question No.	Summary of Findings
Question 1(a)	**What machine learning algorithm is suitable for identifying gender bias?** The support vector machine learning algorithm (SVM) yielded the most accurate and interpretable results.
(b)	**Which is the optimal approach to representing features in text classification experiments to identify gender bias?** A boolean approach to representing features was the optimal approach in terms of predictive accuracy and interpretability of the results.
(c)	**What approach to feature extraction is most informative in identifying gender bias in text?** The following are the informative feature sets in terms of identifying gender bias in the text. The discriminative features in these feature sets are used for analysis in relation to the second research question of this thesis: Unigrams Adjectives Verbs General Inquirer lexicons Newspaper section

Table 4.1: Summary of Findings of Research Questions 1(a-c) from Corpus Analysis of Articles Naming Ministers in Headlines

The output of this phase of analysis is a list of features which yield classification results with good prediction accuracy and discriminative features which are meaningful in terms of their interpretability in relation to gender bias. These features are then analysed in depth to verify whether there is gender bias.

The following section present detailed analysis of the methods of text classification used in this research. The feature types extracted from the corpus are shown in Table 4.2. For each of these feature sets, different machine learning algorithms and methods of representing the texts were tested. The machine learning algorithms tested included a support vector machine, decision tree (J48) and Naïve Bayes algorithm. The methods of representing features included a boolean representation, word frequencies or tf-idf (term frequency/inverse document frequency). Details of these approaches are presented in Chapter 3. Analysis is structured according to each feature type.

No.	Feature Sets
1	Unigrams (single words)
2	Bigrams (two word sequences)
3	Vocabulary complexity
4	Parts of speech n-grams
5	Stop words
6	Sentiment
7	Adjectives
8	Verbs
9	General inquirer lexicons
10	Author's gender
11	Newspaper section

Table 4.2: Feature Sets Used in Text Classification

1. **Feature Set - Unigrams (single words)**:

   Table 4.3 and Table 4.4 show the prediction accuracy results for classification experiments using unigrams as features. This analysis involved a broad range of experiments including different approaches to representing features (boolean, tf-idf, bag-of-words), minimum frequencies and machine learning algorithms.

Feature Sets		SVM	NB	J48
Min freq 1	Bag-of-Words	74.3	68.8	74.0
	Boolean	75.7	70.0	75.7
	Tf-idf	76.0	68.1	73.5
Min freq 2	Bag-of-Words	74.1	68.8	74.0
	Boolean	75.2	70.3	75.7
	Tf-idf	75.9	68.1	73.4
Min freq 3	Bag-of-Words	73.6	68.8	74.0
	Boolean	75.0	70.2	75.7
	Tf-idf	75.5	68.7	74.0
Min freq 4	Bag-of-Words	73.4	68.8	73.9
	Boolean	75.0	70.2	75.7
	Tf-idf	75.6	68.4	73.5
Min freq 5	Bag-of-Words	73.1	68.8	73.9
	Boolean	75.0	70.2	75.7
	Tf-idf	75.2	68.7	73.8
Min freq 10	Bag-of-Words	71.1	68.9	73.9
	Boolean	74.5	70.1	75.6
	Tf-idf	74.1	74.1	74.7
Average accuracy		74.6	69.5	74.5
Median accuracy		75.0	68.8	74.0

Table 4.3: Classifier Prediction Accuracy Using Unigrams as Features

Feature Sets		SVM	NB	J48
Stemmed	Bag-of-Words	74.1	69.4	72.2
	Boolean	74.0	70.8	73.9
	Tf-idf	74.7	66.9	73.1
Stop words removed	Bag-of-Words	73.7	69.4	74.4
	Boolean	75.0	71.7	76.1
	Tf-idf	76.6	68.6	74.1

Table 4.4: Classifier Prediction Accuracy for Unigrams with Stop Words Removed and Stemming Applied

These results show that the support vector machine learning algorithm produced were the most accurate in terms of identifying the gender of the minister featured in an article. A boolean representation of the features with a minimum frequency of 2 also yielded the best results. Neither removing stop words, nor applying stemming, to unigrams improved the accuracy of the models or the quality of the discriminative featured identified. The following is a detailed description of the results:

Machine Learning Algorithm: The support vector machine (SVM) classification algorithm consistently gained higher prediction accuracy than both the J48 and Naïve Bayes classifier. The Naïve Bayes classifier produced the lowest accuracy. This accords with literature that identified SVMs as the most widely used in text classification tasks (Pilaszy, 2005). The SVM algorithm also yielded weighted features according to each class thereby facilitating analysis of the most important discriminative features.

Feature Representation: Unigrams were represented numerically using three different approaches. These were word frequency counts, a boolean representation indicating word presence and a term frequency-inverse document frequency (tf-idf) representation. The tf-idf representation yielded the highest prediction accuracies. However, they were only marginally better than those resulting from the boolean representation. Representing the features as word frequencies yielded the lowest classification accuracies.

As Wiedemann (2013) pointed out, in text classification in the social sciences, the text classification approach that produces the highest accuracy is not always the best in terms of the research goals. Although in this research the tf-idf representation yielded marginally higher accuracy than the boolean representation, the Boolean representation ranked features suggesting gender bias as more important. This can be seen in Appendix A.5 and Appendix A.6. The fact that in the Boolean representation, these features are identified as more important distinguishing between articles featuring male and female politicians, suggests that it is the optimal representation for features in terms of the research objectives.

The optimal minimum frequency of features in the corpus was 2. The reason for this was that incidental features of texts such as spelling errors were included as weighted features in Weka. Weka outputs weights based on the full data set. For example 'duboin' which is a misspelling of the Irish city 'Dublin' was included as a feature indicative of texts featuring female politicians as can be seen in Appendix A.2. Setting the minimum frequency to 2 had the benefit of removing incidental misspellings of words allowing for optimal analysis of the weighting applied by the support vector machine learning algorithms.

Feature Extraction: Stemming the unigrams marginally reduced the predictive accuracy of the classification model. Removing stop words also saw reduced accuracy of the model except when a tf-idf representation of the features was used (Table 4.4). Examination of the discriminative features identified by the support vector machines showed that the discriminative features were no more relevant to the topic of gender bias than the other more accurate approaches.

The discriminative features identified by the classifier contained unigram (single-word) features which, based on findings from previous literature, suggest the existence of gender bias. For example, words pertaining to 'family' and 'childcare' were correlated with coverage of female ministers. Given this, it was concluded that the approach of using single words as features yielded informative and interpretable results.

The list of discriminative features focused on were those generated using the machine learning algorithm (SVM) and the feature representation (binary) method which was found to yield the most meaningful results.

2. **Feature Set - Bigrams (two word sequences)**:

Results of classification experiments using bigrams as features showed a similar pattern to those gained using unigrams as features (Table 4.5). The support vector machine learning algorithm yielded high accuracy rates. A boolean representation with a minimum frequency of 2 yielded the best results based on both predictive accuracy and the prioritisation of discriminative features that suggested gender bias. Representing text as bigrams significantly increases the dimensionality of the vector space. Due to this the J48 machine learning algorithm was not used in these experiments.

Feature Sets		**SVM**	**NB**
Min freq 1	Bag-of-Words	58.6	55.1
	Boolean	58.6	54.7
	Tf-idf	59.3	51.6
Min freq 2	Bag-of-Words	82.5	69.3
	Boolean	83.6	73.2
	Tf-idf	84.4	69.3

Table 4.5: Classifier Prediction Accuracy Using Bigrams as Features

The discriminative features identified by the classifier were composed of sequences of words involving the same features as those identified using unigrams as features. For example, the word transport was identified as one of the most discriminative features in a classification experiment using unigrams as features. Using bigrams as features, the corresponding discriminative features were 'for transport' and 'transport minister'. Bigrams did provide extra contextual information about the features. However, there was a significant increase in the computational cost of running the experiments due to the size of the dataset. It was also necessary to analyse more discriminative features to identify the same patterns as was evident in the experiments using unigrams as features.

In the classification experiments which used bigrams as features, the negative outcomes outweighed the advantages gained by inclusion of contextual information in the features. A better approach was to use unigrams as features to identify patterns in the data and explore contextual information of those patterns further by extracting concordance lines of the words from the original corpus.

3. **Feature Set - Vocabulary Measures**:

The prediction accuracies gained from using vocabulary measures as features for classification experiments were all below 55 percent (Table 4.6). This indicated that there were no differences in the complexity of vocabulary of articles featuring male and female ministers. Measures included in these tests included world lengths, hapax legomenon, TTR (type-token ratio) and the length of the headlines. While vocabulary measures were successful differentiators between male and female writing styles (Corney et al., 2002; Kucukyilmaz et al., 2006), these results show that vocabulary complexity does not change when male or female ministers are featured in media articles.

The patterns relating to the performance of the machine learning algorithms were consistent with the results from previous experiments.

The SVM classification algorithm yielded the highest overall accuracies. The boolean and tf-idf representation of features yielded almost the same levels of predictive accuracy.

Feature Sets	SVM	NB	J48
Bag-of-Words	53.4	53.4	52.1
Boolean	54.0	50.6	53.6
Tf-idf	54.1	52.1	54.5

Table 4.6: Classifier Accuracies Using Vocabulary Measures as Features

4. **Feature Set - Parts of Speech N-Grams**:

No significant gender difference was detected in experiments using parts of speech as features. Parts of speech unigrams and bigrams were tested in these experiments. Results are presented in Table 4.7. The highest predictive accuracy of 55.7 percent was produced by the SMV classifier using a boolean representation of trigrams as features. This contrasts with the 75.4 percent accuracy achieved by Argamon et al. (2009) using parts of speech tags as features in predicting an author's gender. The low predictive accuracy gained from using parts of speech as features indicates that there are no significant style differences when a male or female politician is featured in an article.

The SVM classification algorithm and representing the features using a boolean approach yielded the best results in all of the experiments. This is consistent with results from previous experiments.

Feature Sets		SVM	NB	J48
	Bag-of-Words	52.0	49.4	52.2
POS Tags	Boolean	52.4	51.1	52.2
	Tf-idf	51.7	49.3	52.2
	Bag-of-Words	52.6	54.3	50.8
POS Bigrams	Boolean	53.8	53.2	50.9
	Tf-idf	52.9	52.9	50.1
	Bag-of-Words	55.2	54.2	51.6
POS Trigrams	Boolean	55.7	53.6	50.9
	Tf-idf	55.6	52.7	52.5

Table 4.7: Classifier Accuracies Using Parts-of-Speech N-Grams as Features

5. **Feature Set - Stop words**:

Similar to the other measures of style such as parts-of-speech, little difference was found between articles featuring male and female politicians when stop words were used as features (Table 4.8). Since stop words have little meaning in themselves, the features alone are not interpretable in relation to gender bias (Appendix A.4). Given the over 70 percent accuracy in other classification studies detecting gender-linked differences in style using stop words as features (Argamon, Koppel, Pennebaker and Schler, 2007), it was concluded that the results of this research indicate no differences in writing style when the subject is a male or female politician.

Feature Sets		SVM	NB	J48
	Bag-of-Words	56.1	52.8	50.7
Stop words	Boolean	51.3	52.6	51.3
	Tf-idf	53.6	53.5	51.8

Table 4.8: Classifier Accuracies Using Stop Words as Features

6. **Feature Set - Sentiment**:

No difference was observed in overall sentiment of the articles featuring male or female politicians (Table 4.9). The same result was found in examining the sentiment of sentences which named the politicians. While more complex approaches to evaluating sentiment in text could be explored, these results do not suggest differences in the overall sentiment of the articles or sentences featuring male or female politicians. These results suggest that there is no difference in the sentiment being expressed towards male or female politicians so there is no indication of overt gender bias.

Feature Sets		**SVM**	**NB**	**J48**
	Bag-of-Words	52.1	52.3	51.0
Sentiment	Boolean	51.9	51.7	51.5
	Tf-idf	50.6	52.3	52.2

Table 4.9: Classifier Accuracies Using Sentiment Indicators as Features

7. **Feature Set - Adjectives**:

Table 4.10 shows the prediction accuracy achieved using adjectives as features. The highest accuracy was produced using a boolean representation of all adjectives in an article using the SVM classifier. As can be seen from the discriminative features identified by the classifier (Appendix A.7) some terms were incorrectly identified as adjectives by the tagger demonstrating Stamatatos's (2009) observation that the more text analysis involved in extracting features from text, the less accurate the results.

Feature Sets		SVM	NB	J48
Adjectives in articles	Bag-of-Words	62.1	60.7	63.6
	Boolean	65.5	62.5	63.2
	Tf-idf	65.1	61.8	62.3
Adjectives in sentences naming ministers (including instances with no adjectives)	Bag-of-Words	58.3	56.9	57.5
	Boolean	58.4	58.1	56.9
	Tf-idf	59.0	55.5	55.6
Adjectives in sentences naming ministers (excluding instances with no adjectives) (baseline 52 %)	Bag-of-Words	59.2	56.9	56.6
	Boolean	59.0	58.2	56.9
	Tf-idf	58.9	55.1	55.3

Table 4.10: Classifier Accuracies Using Adjectives as Features

The discriminative features identified by this experiment highlighted adjectives that were highly interpretable in relation to gender bias. For example the words 'drink' and 'beat' were identified as words that appeared in articles featuring male politicians. Adjectives describing personal appearance were notably absent, contrasting with previous research showing an excessive focus on female politicians' appearance (Trimble et al., 2013). This demonstrates the benefits of using certain word types to target the analysis of the text.

8. **Feature Set - Verbs**:

The predictive capacity of verbs in classifying texts according to the gender of the minister featured in the articles all remained close to 55 percent (Table 4.11). Highest accuracy was gained by analysing all verbs in an article. However, the features that were more likely to pertain to the ministers themselves yielded accuracies that were close to the baseline. Only one experiment suggested that further investigation may be valuable. This was the boolean representation which identified verbs from sentences featuring ministers. The SVM classifier yielded the most accurate results and the Boolean representation of features was most accurate in predicting the gender of the minister featured in the articles.

Feature Sets		**SVM**	**NB**	**J48**
Verbs in articles	Bag-of-Words	57.4	55.6	53.4
	Boolean	57.4	54.6	52.6
	Tf-idf	57.6	54.6	54.0
Verbs in sentences featuring politicians	Bag-of-Words	53.9	52.4	50.8
	Boolean	56.2	52.8	51.8
	Tf-idf	54.8	51.8	51.9

Table 4.11: Classifier Accuracies Using Verbs as Features

9. **Feature Set - General Inquirer Lexicons**:

Experiments which used the General Inquirer Lexicons to extract features from the newspaper articles demonstrate the advantages of using domain specific lexicons in text classification. Analysis of the discriminative features showed how this approach allowed features that were highly relevant to the topic of this research to be identified in the texts and analysed.

There was variance in the overall accuracies gained from each experiment (Table 4.12). Isolating sentences that named the politician did not dramatically improve the overall accuracy of the experiments. However, it was assumed that the results would be more likely to pertain directly to the politician in question if the words were extracted from the sentences naming the ministers. This increased relevance of the discriminative features compensated for a lower accuracy in overall classification.

The benefits of using a domain specific lexicon to extract features from text are that it focuses the analysis on attributes that are particularly relevant to the topic. However, dictionary based lexicons do not take account of the context in which a word is used. For example the words 'shell' and 'enterprise' are in the lexicon of action words but their use in the corpus refers to a policy issue and the name of a ministerial role respectively.

Consistent with the findings from previous experiments, the SVM classifier produced the highest overall accuracy along with the tf-idft and boolean word frequency.

Feature Set	**Extracted from**	**Representation**	**SVM**	**NB**	**J48**
GI Action Words	All Articles	Bag-of-Words	63.3	61.2	62.0
		Boolean	64.5	62.4	62.3
		Tf-idf	64.1	54.9	61.8
	Sentences naming minister	Bag-of-Words	63.7	60.0	62.1
		Boolean	64.4	60.0	61.4
		Tf-idf	64.0	55.4	61.1
GI Power Words	All Articles	Bag-of-Words	58.5	55.9	54.1
		Boolean	59.2	58.1	55.9
		Tf-idf	58.8	56.6	56.1
	Sentences naming minister	Bag-of-Words	58.8	56.5	55.0
		Boolean	58.5	58.5	56.0
		Tf-idf	59.3	56.1	56.3
GI Political Words	All Articles	Bag-of-Words	52.7	49.7	53.1
		Boolean	53.1	53.5	53.3
		Tf-idf	53.8	50.8	53.4
	Sentences naming minister	Bag-of-Words	54.4	54.4	55.8
		Boolean	57.2	56.7	55.9
		Tf-idf	58.1	57.3	56.4

Table 4.12: Classifier Accuracies Using General Inquirer Lexicons as Features

10. **Feature Set - Author's Gender**:

The gender of the author did not influence whether the subject of the article is male or female (Table 4.13). The baseline for the classification experiments was 56.2 percent. As some articles did not list an author these articles were excluded. The fact that the predictive accuracy of the classifier was almost equivalent to the baseline indicates that gender was not a factor influencing the proportion of articles featuring male or female ministers.

Feature Set	SVM	NB	J48	Baseline
Author Gender	56.6	55.9	56.1	56.2

Table 4.13: Classifier Accuracies Using Author's Gender as Feature

11. **Feature Set - Newspaper Section**:

To explore whether gender bias is evident in the placement of articles within the newspapers, sections of newspapers were identified and used as features for text classification. The models yielded almost 60 percent accuracy with little difference in the performance of the classifiers. While the accuracy is not particularly high, they are at least 5 percent over the baseline and the discriminative features are interpretable. The discriminative features are therefore analysed in the next section for evidence of gender bias.

Feature Set	SVM	NB	J48
Section	58.7	58.8	58.7

Table 4.14: Classifier Accuracies Using Newspaper Section as Feature

4.2.2 Gender-based Difference in Quantity of Coverage

Female ministers were mentioned in headlines of articles between 1997 and 2011 almost twice as often as male ministers. Given that previous research found that the frequency of mentions of women in texts relative to that of men can reflect gender bias (Bystrom et al., 2001; Fuertes-Olivera, 2007; Miller et al., 2010; Pearce, 2008), the results of this study suggest a gender bias against male politicians (Figure 4.2).

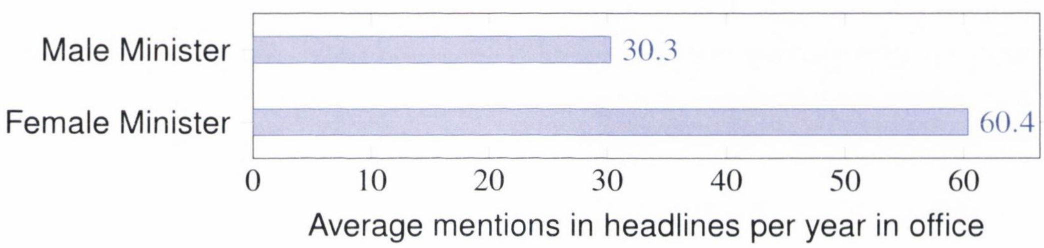

Figure 4.2: Average Number of Mentions in Newspaper Headlines Per Year in Office

Of the total of 4,878 articles in the corpus there was an almost even division between articles featuring female (2,333) and male (2,545) ministers. However, in the time period that articles in this corpus were written, the female ministers held office for a total of 38.6 years and comparable male ministers held office for 84.1 years. When this was taken into consideration, coverage of female ministers was almost twice that of male ministers.

Dail No.	Gender	Years in Office	Publications Available
28th Dáil	M	29.7	Irish Times
	F	14.8	Irish Times
29th Dáil	M	32.4	Irish Times and Independent
	F	12.7	Irish Times and Independent
30th Dáil	M	22.0	Irish Times and Independent
	F	11.1	Irish Times and Independent

Table 4.15: Number of Years Ministers Spent in Office for Each Dáil Term

Newspaper articles from the Irish Independent were only available for the 29th and 30th Dáil terms (Table 4.15). To address this, average coverage was calculated for each newspaper based on the Dáil terms the newspapers corresponded to (Table 4.16). This showed differences in the quantity of coverage afforded to male and female ministers in both publications with the largest disparity evident in the Irish Independent (Figure 4.3).

Publication	**Gender**	**No. Articles**	**Years in office**	**Articles per year**
Irish Times	M	1,894	38.6	49.0
Irish Times	F	2,189	84.1	26.0
Irish Independent	M	356	62.0	5.7
Irish Independent	F	439	27.6	16.0

Table 4.16: Total Time in Office and Number of Mentions in Newspaper Headlines

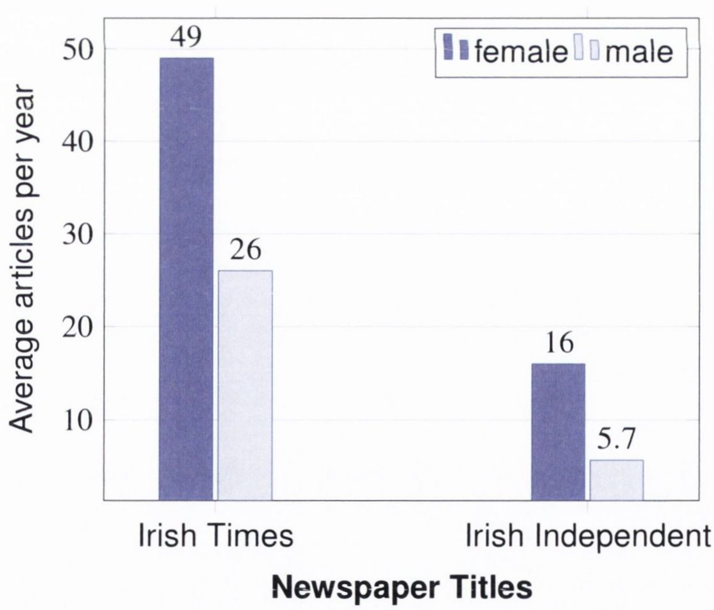

Figure 4.3: Average Number of Mentions of Ministers in Headlines in Irish Times and Irish Independent

Discussion of Findings

The extent of the difference in coverage of male and female ministers in Irish newspapers suggests that there is bias in favour of female politicians. However, it is also possible that while female ministers are named more often in headlines, this may not translate to more coverage overall. Male ministers may be more frequently featured in the content of articles. This was examined when the search criteria in developing a corpus were broadened to include mentions of ministers in the content of articles. However, this showed a similar pattern where female ministers received greater coverage than male ministers.

No previous studies of the volume of coverage afforded to male and female politicians has uncovered differences on the scale found in this research. However, the majority them focused on media coverage during election campaigns (Bystrom et al., 2001; Heldman et al., 2005; Miller et al., 2010). It is therefore possible that different patterns exist for ministerial level politicians. To examine whether this is the case, coverage of the Irish Presidential Elections in 2011 was examined and results are detailed further in this chapter. A further explanation is that female ministers are gender schema-inconsistent (Bem, 1981) and therefore attract more coverage. However, verification of this requires further research into the correlation between gender schema consistent or inconsistency and coverage in the media.

Studies of the quantity of coverage of politicians do not address whether the coverage is positive or negative. While female ministers may receive more coverage than their male counterparts, this may not be an advantage if that coverage is more negative. The next section addresses this issue, analysing the differences in the content of articles featuring male and female politicians.

4.2.3 Gender Bias in Content of Articles

This section presents an evaluation of the features of texts which discriminate between articles featuring male and female politicians and which are also interpretable in terms of gender bias. The objective of the analysis is to identify differences in coverage of male and female ministers that are attributable to gender bias. The differences analysed are those identified by text classification using the following feature sets:

1. Unigrams (single words)
2. Adjectives
3. Verbs
4. General Inquirer lexicon
5. Newspaper section

The discriminative features used by the classifiers for each of the above approaches to feature extraction were interpreted and those suggesting gender bias identified. The interpretation of the features was based on findings from previous research on this topic along with contextual knowledge of political events between 1997 and 2011. To verify the context of the words in the corpus, concordance lines of the words were extracted from the corpus and qualitatively analysed to access whether the patterns identified by the classifier constituted gender bias.

The discriminative features analysed in this research are presented below within the categorisation of gender bias identified in the review of related literature. These themes, along with an overview of the main findings, are outlined in Table 4.17.

Theme	Summary of Finding
Family Relationships	Female minister's spouse mentioned 4 times as often as male minister's spouse
	Gender-based differences in portrayal of spouses
	Male ministers associated with discussions of families in economic terms while female ministers are associated with discussions of family in relation to social policy or in personal terms
	Conflict between demands of job and family life emphasised in relation to female ministers while male minister's enjoyment of family life is focused on
	Discussions of work-life-balance policies are associated with female ministers
	Childcare policy primarily associated with female ministers
Focus on Gender	Female minister's gender mentioned while a male minister's gender not mentioned
	Female ministers portrayed as supporting gender equality issues while male ministers portrayed as under pressure to conform to them
Use of Stereotypes	No evidence of stereotypes concerning descriptions of physical appearance
	Association of political drinking culture with male ministers
	Sexualised identity created for female mininsters
Political Style	Female ministers portrayed as more accepting of government policies
	Association of female ministers with negative political communication style
Personalised Coverage	First name and surname used to refer to female ministers while surnames only are used more often to refer to male ministers
	Female ministers more likely to be described as 'formidable'
Masculine Narrative	Male ministers associated with terms associated with military, institutions of state and sport
Different Policy Focus	Stereotypical association of male and female ministers with policy issues
	More personalised coverage of female ministers and coverage of male ministers in relation to policy
	Female ministers featured more predominantly in highest profile newspaper sections

Table 4.17: Gender Bias in Content of Articles Naming Ministers in Headlines

Family Relationships and Roles

Terms pertaining to family relationships and roles were associated with female ministers (Table 4.18). This pattern emerged using single words as features (Appendix A.6, A.8 and A.17 for listing of all discriminative unigrams).

Female Minister	Male Minister
husband	households
childcare	children
family	
female	

Table 4.18: Discriminative Features Associated With Family

Mentions of a minister's spouse:

The word 'husband' was more likely to be used in articles featuring a female politician in the headline (Table 4.18). However, the term referring to the spouse of male ministers, 'wife', did not appear as a discriminative feature in any of the classification experiments. This imbalance in the references to spouses suggests a gender bias since more importance is placed on the marital status of women (Baker, 2010; Fuertes-Olivera, 2007; Holmes, 1994). This reinforces previous research which found that coverage of female politicians was more likely to focus on their personal family circumstances than male politicians (Brikse, 2004; Miller and Peake, 2013; Spears et al., 2000).

To further examine the context in which spouses of ministers were discussed, concordances of both words from the entire corpus were extracted. These were manually analysed to identify those instances that referred to a minister's spouse. Analysis of concordance lines of the word 'husband' in the corpus showed that there were 48 mentions of the term which referred to a minister's spouse. There were 27 mentions of the word 'wife' referring to a minister's spouse. Hence, for every year in office, a female minister's spouse is mentioned four times as often as that of a male minister's spouse (Table 4.19).

Gender	Mentions of Minister's Spouse	Years in Office	Mentions/Year
F	48	38.6	1.2
M	27	84.1	.3

Table 4.19: Mentions of Spouse Per Year in Office

There were also differences in the context in which spouse were mentioned. The context was examined by extracting concordance lines of words referring to the spouses of ministers. These were then grouped according to common themes in how they were used.

Analysis of the context of the occurrences of the terms 'wife' and 'husband' showed that the wives of minsters were often spoken of in the context of being 'at home'. Husbands of ministers however were not mentioned in this way. They were spoken of in discussions of family life and their relationships were portrayed as more equal partnerships. Some examples how the words 'wife' and 'husband' were used in relation to ministers are shown in Table 4.20. These show how wives of ministers were described as emotional, attractive and conducting supportive tasks such as driving the ministers during their campaigns. In contrast, husbands were characterised as sharing in decisions jointly, helping and supporting the ministers.

Portrayal of Spouse	Concordance Extracts
	Spouse of Male Ministers
At home	...government business in Dublin and only his wife, Mary, was at home. She became upset...
	...from my home where I have a wife and five children," he said.
	Micheal Martin at home with his wife Mary and children Micheal Aodh and...
Attractive	...all a politician could need, such as an attractive wife and two sweet young daughters.
Providing Assistance	...battled to regain his seat, with his wife driving him all over the country to...
	Spouse of Female Ministers
Joint decisions	...on the matter, discussed it with her husband and decided "that now is the right...
	...that she holds shares jointly with her husband Brian Geoghegan in the Bank of Ireland.
Family life	...have a normal family life with her husband, David Charlton, and their two children. "I...
	...uninterrupted day walking in Donegal with her husband and the "wains" (children) and, when the...
Support and help	...difficult and challenging. However, she had her husband there to support her and help her...

Table 4.20: Examples of Concordance Lines of Political Spouse

Differences in how spouses are referred to could be a result of the fact that husbands of ministers may have had careers outside the home while wives of ministers may not. However, using the test for gender bias described by Mills

(2002) involving swapping of the gender in one expression for another does suggest that if the female ministers were male, their spouses may have been spoken of in different terms (Table 4.21). The phrase 'at home' is reminiscent of the stereotype of women occupying the domestic sphere. Men might have been portrayed as angry instead of 'upset' and it seems unlikely that their attractiveness would be commented on in the media. This points to gender based differences in how families are portrayed.

Concordance Examples of Political Spouse
...government business in Dublin and only her husband John, was at home. He became upset...
...from my home where I have a husband and five children, 'he said. 'I have...
...at home with her husband John and children Micheal, Aodh and...
...politician could need, such as an attractive husband and two sweet young daughters....
...battled to regain her seat, with her husband driving her all over the country to...

Table 4.21: Concordance Examples of Political Spouse with Gender Swapped

Families vs. households:

The term 'household' is associated with male politicians. This term is commonly used in economic discourse referring to households as economic units. Concordance analysis showed that the term was used primarily to discuss public policy and economic issues pertaining to families.

This contrasts with the association of female politicians with the word 'family'. This term is more likely to be used in personal discussions of an individual's family or in relation to social policy. References to 'family', taking into account the years spent in office, were cited 2.5 times more often in articles featuring female ministers than in articles featuring male ministers (Table 4.22).

Gender	Mentions of Family	Years in Office	Mentions/year
F	440	38.6	11.4
M	380	84.1	4.5

Table 4.22: All Occurrences of the Word 'Family', Referring to Minister's Family, Per Year in Ministerial Office

When only extracts that referenced a minister's own family were analysed, the families of male ministers were mentioned more often per year in office than the families of female ministers. A likely explanation for this is that a greater proportion of male ministers had children and were married. Each of these concordance extracts are shown in Table 4.23.

A qualitative analysis of concordance lines concerning a minister's family showed that overall, the male ministers family life was portrayed in a more positive light than female ministers'. Conflict emerging from the demands of political life and that of their family was a focus of coverage of both male and female ministers, seeming to contradict research associating discussions of conflict between the demands of politics and family obligations and media's coverage of female politicians (Van Zoonen, 2006). However, the family life of the male ministers was also portrayed as a happy one despite the fact that they would have liked more time to devote to them. A similar positive portrayal of a female politician's family life is lacking resulting in an overall focus on the negative aspects of combining politics with family life.

Theme	Concordance Extracts
	Female Minister's Family
Negative effects	...her intelligence questioned, her competence scorned. Her family and friends are getting upset, she says...
	...makes it difficult to have a normal family life with her husband, David Charlton...
	...of foreign travel was not conducive to family life. However, her promotion...
Prioritising family	...myself to anything else, apart from my family and my friends," she says. This weekend...
Needing support	If you didn't have a partner or family who were there for you and thought...
	I could not do the job. The family have a photograph of me and...
	Male Minister's Family
Importance of family	...in life are said to be his family, politics and sport. He isn't seen hanging...
Limited time	...occasions to be at home with my family, rather than doing 24 functions in 48...
	...prefer to be at home with his family on St Patrick s Day rather than...
	...difficulty of being away from home and family for a lot of the time.
	...rushing to respond: the time with the family is too limited and too precious...
Enjoyment of family	his life, is his family. When it's family time at the weekend or on holidays...
	...the priority in his life, is his family. When it's family time at the weekend...
Negative effects	...1989 his wife, Mary, wondered if the family could afford the risk, what with a...
	...caused considerable upset to myself and my family," he said.
	...side of his chosen career is that family life must suffer. He discovered quickly...
Wife's sacrifice	...set aside her political career to rear her family. She is said, however, to eat, drink, politics politics...
	...that she's the "real" politician in the family which seems a mite unfair to Micheal.
Profile	...club of which he is a member. Family:married to former Fianna Fail activist...
	...There is nothing flash about the Martin family. They holiday every year in Courtmacsherry...

Table 4.23: Concordance Examples of Minister's Immediate Family

Work-life-balance and childcare policy:

Concordance extracts of mentions of the word 'family' showed that in all of the articles featuring politicians in the headlines, there were five instances where policy was mentioned in terms of work-life balance and childcare issues, four of which were in articles featuring female politicians. References to family in the context of work-life balance and childcare issues were identified by manually extracting them from the 820 concordance extracts of the term 'family'. Each of these are presented in Table 4.24. This supports research which showed a tendency in the media to associate female ministers with policy issues regarded as women's issues regardless of their cabinet portfolio (Brikse, 2004; Carroll and Schreiber, 1997; Devitt, 1999; Kahn, 1994, 1996).

1

Gender	Concordance Extracts
M	...facilitating a better balance of work and family life. Maternity benefit, for example, had been...
F	...return to work, such as childcare and family-friendly policies. "This is about recognising that these...
F	...and alternative work practices". "The bulk of family responsibilities still falls on women in...
F	...a relationship and it is crucial that family-friendly initiatives, such as adequate childcare and flexible...
F	...career ambitions and what I would term family ambition. This polar choice is unnecessary, destructive...

Table 4.24: Concordance Examples of 'Family' in the Context of Work-Life Balance Policy

It is not clear from these extracts whether female ministers themselves highlighted these issues or whether they are a result of the media's association them with female ministers. Previous research found more positive media coverage of female politicians if the stories relate to issues traditionally associated with women (Scharrer, 2002). Therefore, even if women are themselves focusing more on childcare and work-life balance issues, it is possible that the cause of this is media bias which affords more positive coverage to ministers who do.

The word 'children' was identified as an important discriminative feature for articles featuring male politicians (Appendix A.5). Examination of concordance extracts of all 1,687 instances of the word children in the corpus showed that these primarily pertained to discussions of policy issues regarding health and social welfare. As these issues most commonly concerned the Minister for Health, it was important to view the frequency of mentions factoring in the gender of the ministers. Factoring into account the time that male and female ministers spent as Ministers for Health and Children, the rates of mentions of the term are almost equivalent for male and female ministers.

Gender	**Mentions of Children**	**Years as Minister** for Health and Children	**Mentions/Year** Per Year
F	909	7.2	125.2
M	778	6.3	123.4

Table 4.25: Mentions of Children Per Year as Minister for Health and Children

The word 'childcare' was identified in the text classification experiments as being associated with female ministers. A qualitative analysis of concordance lines showed that all references to 'childcare' related to the need for an improved childcare system. Of the 83 mentions, only 12 of those were from male politicians. Hence for each year in office, female ministers were associated with childcare 11 times more than their male counterparts. However, the variance in the frequency of mentions of 'childcare' in articles featuring male and female politicians could be considered even if female ministers held the cabinet portfolio associated with this policy issues for a greater proportion of the time. However, male politicians in fact held the cabinet post related to childcare issues, namely Social Community and Family Affairs and Health and Children for longer than female ministers.

Whether the association of female ministers with childcare is a result of their own prioritisation of the issue or the media linking childcare issues with female ministers is not clear. However, the pressure that female ministers may feel to address 'women's issues' is evident in the extract from the Irish Times below. This also aligns with research by Scharrer (2002) which found that more positive coverage is afforded to female politicians when they address issues considered stereotypically concerning women.

> *"New-look caring Harney must try harder - She didn't play weepy, she simply got on with the job. Those were the positive attributes. In getting on with it, however, the kind of solidarity that, rightly or wrongly, a woman in politics can demonstrate was not only left behind, but often trodden underfoot. Signs of womanliness, or gender-related challenges, were sometimes treated as signs of weakness. Harney could be extra tough on the traditional problems of her own sex. Remember her harsh, ill-judged words about young, single mothers? Remember her refusal to join her sister deputies across parties working (in solidarity) to upgrade the so-called female issues such as childcare, health and creating a better gender-balance throughout political life? Harney's political priorities mirrored*

her personal decisions. She did not work to change the family-hostile scheduling with which the Dáil conducts its business. Her interest in expanding the number of women in the workforce coincided with pressure on the labour market rather than happening because she wanted everyone to get a fair crack of the whip. Harney's voice counted loudly in Government because she is a woman, and women stereotypically know about such things." The Irish Times March 2, 2001

Focus on gender

Miller et al. (2010) and Norris (1997) cited the unnecessary focus on the gender of women in politics as evidence of gender bias in the media. The word 'female' was identified in the classification algorithms as being a discriminative factor identifying articles featuring female politicians (Appendix A.8). However, the word male is not associated with articles featuring male politicians. Analysis of the frequency of the words in the corpus showed that for each year in office, the word 'female' was mentioned over 5 times more in association with female ministers than the word 'male' was associated with male ministers (Table 4.26).

Gender	Mentions of Female	Years in Office	Mentions/Year
F	45	38.6	1.2
M	18	84.1	0.2

Table 4.26: Mentions of the Term Female in Articles Featuring Male and Female Politicians Per Year in Office

Concordance analysis of the term ‘female’ found that references to the minister’s own gender accounted for 23 percent of these instances (presented in Table 4.27). In contrast, analysis of concordance extracts of the word ‘male’ showed that of all of the mentions of the word, not one referred to a male politician’s gender. This suggests gender bias in the form of an excessive focus on gender as outlined by Norris (1997) and Miller et al. (2010).

An example of such unnecessary focus on gender is presented by the description of Mary Harney as the ”most successful female politician in the history of the State” (Table 4.27). While this may be true, given her political achievements, she could be described as one of the most successful politicians in Ireland’s history. Such comparison of her political achievements only with other female politicians implies that she is not comparable to male politicians and portrays her as other than the male norm.

Concordance Extracts
...Her objective was to make it to the end of the session without putting her foot in it. Alas, for female politicians, what they say isn't all they have to worry about. A stellar performance can be ruined by wardrobe...
...Her exceptional survival instincts and can-do beliefs single her out among male and female politicians.
...They were delighted with the news of getting a female boss with the reputation of having a good sense of humour, who looks well and can compete on any level with the boys.
...If female prelates were allowed in the Catholic Church, she would have made an excellent bishop.
...There are two things I have in common with our Education Minister Mary Coughlan. The first, we are both female, and the second, we are both from Donegal.
...news of the appointment of the first female Minister for Agriculture and Food swept through the grounds.
...ready to lead the male-dominated Fianna Fail party and become its first female leader.
...Asked if the FF 'boys' club', which has only eight female TDs, is ready for a female leader...
...Ms Hanafin said it was a disappointment to her to lose her seat and it had been a rare privilege to serve as a female minister in government for 11 years.
...would also be visiting the grounds to add to the female side of the equation.
...Ireland has never had a female commissioner. Ms Coughlan was mooted in European circles as a possible candidate for the job. Government officials were told there would be a lot of support for Ms Coughlan if she was put forward.
...under pressure from the head of the European Commission to put forward a female candidate for the post. under pressure from the head of the European Commission to put forward a female candidate for the post.
...The most successful female politician in the history of the State, Ms Harney has been PD leader for 13 years and Tanaiste for nine.
...Remember her refusal to join her sister deputies across parties working (in solidarity) to upgrade the so-called female issues such as childcare, health and creating a better gender-balance throughout political life?
...Calling it "Women's Day" here at the National Ploughing Championships yesterday when the news of the appointment of the first female Minister for Agriculture and Food swept through the grounds.

Table 4.27: All Concordance Extracts of 'Female' Referring to Minister's Gender

Gender equality policy was the issue under discussion in 20 percent of the occurrences of the word 'female'. Men and women were also associated differently with these issues. Female ministers were associated with the issues more frequently and described as supporting them. However male minsters were only described as being under pressure to conform to them. Examples of such extracts are shown in Table 4.28

Gender difference in how ministers are associated with gender equality issues could be a reflection of the political viewpoints and actions of the minsters. However, it could also be a reflection of a bias in the media which sees equality policy as primarily concerning women and therefore excludes male ministers from such debate. The fact that there are no quotes from male ministers in relation to gender equality presents the possibility that journalists do not ask men about such issues.

Gender	**Concordance Extracts**
F	...inequality and the small number of women in upper management were caused by key obstacles like the "relative absence of female role models at senior levels in the workplace" and the "failure to recognise the need to accommodate more flexible and...
M	...union ballot. Many of the female teachers who spoke during the question-and-answer session branded the proposals as "sexist" because a female teacher who takes a number of years out of the profession to have children will lose pension entitlements in respect...
M	...equality and law reform spokeswoman, said Mr Fahey's appointments to the task force "wins the prize as regards lack of female representation". The Minister made 10 appointments to the task force, all of whom are men. Mr Fahey said he found...

Table 4.28: Concordance Examples of 'Female' Referring to Gender Equality

Use of stereotypes

Trimble et al. (2013) showed how stereotypes are often used to portray female politicians, particularly in relation to physical appearance and sexuality. In a study examining adjectives applied to women in British newspapers, Caldas-Coulthard and Moon (2010) found that more descriptions of physical appearance were associated with women. It was therefore expected that by extracting adjectives and using those as features in text classification, an association would be identified between adjectives describing physical appearance and female ministers. However, no such association was uncovered by the classifiers.

Using verbs as features, an association was uncovered between a workplace culture linked to the consumption of alcohol and male ministers. This kind of political culture has been identified in Ireland as alienating women from politics (Joint Committee on Justice, Equality, Defence and Women's Rights, 2009). The word 'drink' was identified as the most discriminative verb associated with articles featuring male ministers. Analysis of concordance lines of the term showed that of all of the instances where a minister is portrayed as drinking alcohol, only one of those pertains to a female minister. All other references relate to a male minister's personal consumption and purchase of alcohol (Table 4.29).

In one example of a reference to a drinking culture, Mary Hanafin, is described as being 'pushed aside' and a political appointment going instead to the Prime Minister's 'drinking buddy'. Such portrayals of an Irish political drinking culture may be a reflection of the political environment rather than a bias on the part of the media. However, systematic association with male politicians with an activity seen as a core part of the workplace culture of politics may serve to portray women's involvement with politics as counter to gender stereotypes (Adcock, 2010).

Gender	Concordance Extracts
M	He said it was the custom at the time to buy drink for canvassers and for "the house" when you went into the pub.
M	Yesterday, in the new regime of his close drinking buddy Brian Cowen, Mr O'Keeffe was handed the plum appointment of Minister for Education, effectively elbowing Mary Hanafin aside.
M	He is extremely sociable, a deadly mimic, with a sharp cynicism, who enjoys a drink.
M	After discussions on policy, the meeting broke up and Government Ministers stood around in groups, sipping drinks and contemplating the next day's vote.
M	The Minister for Transport, Mr Brennan, has refused to confirm or deny that he is the senior politician who allegedly received E5,000 worth of cigars and drink from Aer Rianta in the early 1990s.
F	Following her arrival in the Big Apple on Wednesday night, she joined journalists for drinks in the hotel bar.

Table 4.29: All Concordance Lines of 'Drink' Referring to Personal Consumption/Purchase

The verb 'beat' was identified as the second most important verb associated with male ministers. Concordance analysis showed that these references were primarily mentioned within the context of sporting events and all except one of these were in association with male ministers. The exception featured a female minister leaving a sporting event early because she 'looked bored'. This association of sporting events with male ministers while it does suggest a stereotypical association of men with sport and women as being 'bored' by it, could also be explained by the fact that there was a female minister for sport for only one of the years studied in this research.

Adcock (2010) outlined how sexualised identities are created in media portrayals of female politicians. The term 'woo' was highlighted as a discriminative feature associated with articles featuring female ministers (Appendix A.3). This term is commonly used in the context of an initiation of romance. Analysis of concordance lines of 'woo' show that all except one of the instances are used in articles featuring female politicians.

The term was used metaphorically to describe how women encouraged people to adopt certain political stances or take certain actions. Describing female ministers metaphorically as 'wooing' other parties in negotiations while not using a word with similar connotations in connection with male politicians does suggest the creation of a sexualised identity for female ministers. The identity constructed is of active sexuality, contrary to findings from previous studies which found male sexuality to be portrayed in terms of predatory behaviour (Mills, 2002).

Gender	Concordance Extracts
F	...parties in trying to woo the five independents elected...
F	...the then Enterprise Minister woo her business guests.
F	...Monday Coughlan aims to woo US business...
F	...it is trying to woo the unions back to...
F	...in strong bid to woo Labour...
F	...are also trying to woo politically homeless PDs.
F	...operators, each seeking to woo customers.
M	...Co Mayo to woo community representatives into participating.

Table 4.30: Concordance Examples of the word 'Woo'

Political Style

Differences in the portrayal of the political style and attitude of politicians in the media were identified by the classifier using the General Inquirer Lexicon of action words as features. Of these, the most important discriminative feature for ministers of both genders were words pertaining to attitudes towards government policy or legislation. The term 'revoke' was associated with male ministers while the term 'embrace' was associated with female ministers.

Concordance analysis showed that the term, 'embrace', associated with female politicians, was used primarily as a metaphorical description of their attitude to public policy or political ideology. The word most associated with male politicians, 'revoke' was used in the context of opposing approaches to government policy of revoking legislation (Table 4.31). These findings suggest what Adcock (2010) described as the portrayal of female politicians as "condensing symbols" for the political ideologies they are affiliated with, referring to women being portrayed as being more unconditionally supportive of government policy. Men however, with the increased association with the term 'revoke', are portrayed as more confrontational.

The use of 'embrace', a word that can have romantic connotations, as a metaphor to describe women's attitudes to policy, along with the use of the term 'woo' to describe their negotiating style as described in section 4.2.3 aligns with findings of previous literature that women are portrayed as more emotional then men Pantti (2005). However it also contradictions findings suggesting women are portrayed as passive concerning romantic issues Mills (2002).

Gender	Concordance Extracts
F	every time we really embrace change and are ambitious
F	be open and confidently embrace the challenges of a
F	rewarded, and when we embrace new solutions.
F	We have got to embrace change," she said
F	Mary Harney's embrace of New Right philosophy
M	the report did not embrace proposals for the development
F	Nora Owen, could have revoked the Mahfouz passports
M	Confirming his decision to revoke 24 appointments to the
M	with his proposal to revoke or change the order
M	Martin publishes Bill to revoke Groceries Order
M	Competition Act that will revoke the Order and strengthen

Table 4.31: Concordance Examples of Mentions of 'Embrace' and 'Revoke'

Other words from the General Inquirer lexicon of action words identified as discriminating between male and female politicians pertained to a politician's communication style. The following are these words pertaining to communication ordered according to the weight of each feature or the importance identified in each feature's association to the text categories:

Male Ministers: unite, cooperation, cultivate

Female Ministers: misrepresent, smear, discredit, ridicule, shout

Concordance analysis showed that there were very few occurrences of each word in the corpus. There were also differences in the sense in which the words were used. For example, uses of the word 'smear' showed that the all the female ministers were carrying out the action while the male ministers were the object of them. 'Smear' also referred to a specific Health policy. Another instance of the word described a female politician's lipstick:

> *"A rebellious smear of crimson lipstick held out some promise that the Tanaiste might cut loose if the Opposition's needling became too much."*

The words were grouped according to whether they conveyed a positive or negative portrayal of communication style. Positivity and negativity of the words was judged by the researcher. Of the words listed, the terms 'unite', 'cooperation' and 'cultivate' were judged as positive reflections of diplomatic communication style. The positive words directly correlated with those associated with male politicians (unite, cooperation, cultivate) and the negative words were those most associated with female ministers.

The concordance lines were filtered to extract only those that referred to political communication of the ministers where the ministers were subjects in the communication being described. Aligning with findings in Gidengil and Everitt (2003), this shows that most negative portrayals of political style were of female ministers (Figure 4.4).

Furthermore, while the words associated with female ministers were direct descriptions of speech acts (Caldas-Coulthard, 1995), those associated with male ministers were descriptions of the outcomes or aims of political discourse. Caldas-Coulthard (1995) found that the use of verbs to describe speech acts in British Newspapers reinforced gender stereotypes. Men were found to "shout" and "groan" while women "scream" and "yell" (Caldas-Coulthard, 1995, p.235). In this research however, female ministers were found to shout.

The purpose of the speech acts associated with female ministers also differs from the political outcomes associated with male ministers. The aims of those associated with female ministers are to undermine an opponent while male ministers are associated with political diplomacy. This aligns with Gidengil and Everitt (2003), who found that female politicians are more likely to be described as having an aggressive political communication style.

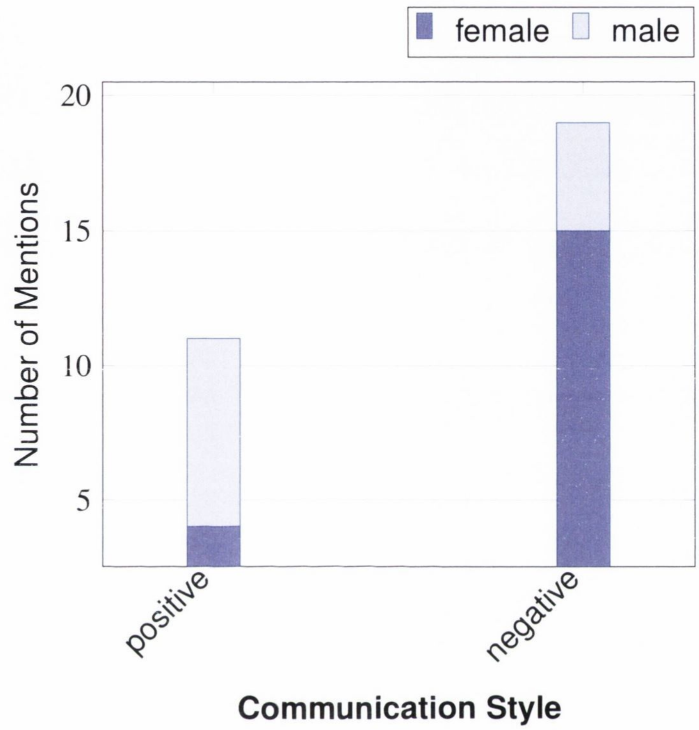

Figure 4.4: Number of Instances Portraying Communication Style According to Gender

Personalised coverage

Female ministers were referred to using their first names more often than male ministers. This aligns with research which found that female politicians are more likely to be referred to in the media using informal versions of their names (Fox, 1997; Uscinski and Goren, 2011). At the text preprocessing stage of this research, names of politicians were replaced with gender-neutral terms which also documented the form of the name used. Results of text classification experiments showed that naming using a minister's first name followed by their surname was associated with female ministers. Using surnames alone were more associated with male ministers. Given that higher status is associated with use of the more formal method of naming, using a surname only (Page, 2003), this finding suggests evidence of media bias.

A pattern of difference in how male and female politicians were associated with terms relating to power emerged in the use of the General Inquirer lexicon of power words as features. The top 5 power words associated with female politicians were 'executive', 'distinguished', 'formidable', 'leader' and 'director'. The top features associated with male ministers were 'jail', 'prohibition', 'squad' and 'chief', suggesting power of institutions. However, concordance analysis showed that only two of these terms were used in a sense which pertained to the ministers themselves and issues of power. For example, the term 'distinguished' was used primarily as a verb which had no relation to portrayals of powerful people. Squad was used metaphorically to refer to football teams. The fact that concordance analysis disproved what had seemed a significant finding from the quantitative analysis of word use, demonstrates the importance of verification of these patterns through qualitative analysis of the context of word use.

Baker (2010) noted that only women were described as 'formidable' in analysis of a corpus of modern British English. This term was also associated with women in this research. Concordance analysis showed that the term 'formidable' was used to portray both male and female politicians. There was a total of 9 mentions of the word in the text. Of these, 5 referred to female ministers while 3 referred to male ministers. While over a 15 year period, this level of difference does not seem high. However when you include the time-period that the newspaper corresponds to, female candidates were described as 'formidable' almost four times as often as male ministers (Details presented in Table 4.32). These findings suggest that in the same way that research has examined descriptions of female politicians as strident (Childs, 2004), further research would be useful to assess the implications of describing them as 'formidable'.

Gender	Formidable	Years in Office	Mentions/Year
F	5	38.6	.12
M	3	84.1	.03

Table 4.32: Descriptions of Politicians as 'Formidable' Per Year in Office

Masculine narrative/metaphor

Media coverage of politics has been critiqued for using metaphors of war and sport traditionally associated with men in political discourse (Gidengil and Everitt, 1999; Trimble et al., 2013). While terms pertaining to activities of war were not identified by the classifiers as terms associated with ministers of either gender, terms pertaining to power inherent in institutions of the state, including the military, and positions of power within it are associated with male ministers. Terms associated with competitive sport are also linked with male ministers. There were also more terms associated with male ministers then female ministers. This corroborates the finding from Koller's (2004) that there are more gender schemas evident in business magazines for businessmen than for businesswomen reflecting the fact that "discursive and cognitive structures determined by hegemonic masculinity show more cultural models of masculinity than of femininity". The following are the discriminative features which pertain to governmental power and sport:

Male ministers: jail, prohibition, squad, chief, politician, chairmen, council, mayor, president, championship, aristocracy, champion, guard, commander, presidency

Female ministers: government, officer

Analysis of the concordance lines of these terms showed that the words not frequently used in direct reference to the ministers themselves. However, the systematic association of male politicians with newspaper articles mentioning high-status roles within the state, military and competitive sport does suggest a bias that excludes female ministers from such discourse. Female ministers are associated with two terms which are linked to institutions of the state, 'government' and 'officer'. The relative absence of association of female politicians with the police force, army or high-status institutional roles suggests that female ministers may be disassociated from those positions and institutions of government. Further research is needed to attribute to bias on the part of the newspaper to this.

The link in newspaper coverage between male ministers and terms pertaining to the police force and the army suggest that male politicians in Ireland are associated with "military masculinity" (Higgins, 2010, p.151). The pattern identified by the classification experiments associating male ministers with power linked to the military and institutions of the state, can be interpreted as perpetuating gendered stereotypes regarding women and their relationship with power (Okimoto and Brescoll, 2010).

The terms 'riot' and 'ban', which are reminiscent of military actions, were also associated with male politicians when the General Inquirer Lexicon of political words were used as features (Appendix A.1). However, in contrast to these findings, female politicians are associated with the term 'invasion', which can be used in the context of an act of war while male politicians are associated with words more associated with peace negotiations, 'conciliation', 'mediation' and 'neutrality'. Concordance analysis of these terms however showed a wide variety of contexts in which the words were used.

Newspaper section

Analysis of the results of text classification when the newspaper sections that the articles appeared in were used as features, yielded contradictory results. While there is evidence of stereotypical associations of male and female ministers with certain topics aligning with previous research on the nature of gender bias in media coverage of politicians, female ministers also seem to dominate in the higher profile sections of the newspapers which would suggest bias in favour of female politicians.

Male politicians feature more in sections stereotypically associated with male pursuits such as motoring, farming and sport. They are also more prevalent in the sections relating to science and international politics. Female ministers appear most in the letters, opinion, entertainment and lifestyle sections. This aligns with research showing more personalised coverage of female politicians in the media then their male counterparts (Devitt, 1999; Garcia-Blanco and Wahl-Jorgensen, 2011; Ross and Sreberny, 2000). Female politicians also featured more in the health section although the ministry of health was held by male ministers for longer than female ministers (Table 3.4). This suggests a gender bias identified by Brikse (2004) and Carroll and Schreiber (1997) where female politicians are associated with health and social issues regardless of their actual role. However, female ministers also appear in the politics section, in feature articles, editorial pages and and on the front pages of newspapers more than male ministers, which suggests there is evidence of gender bias in their favour.

These findings show that there are stereotypical associations of ministers of both genders with topics stereotypically associated with their gender. There is also more of a focus on personal coverage of female ministers. However, coverage of female ministers tends to feature more in the prominent sections of the newspapers than their male counterparts, indicating a gender bias in favour of female politicians.

4.2.4 Changes in Levels of Gender Bias

In order to access whether there were changes in the coverage of male and female politicians between 1997 and 2011, the corpus was divided according to the Dáil term the article appeared in. The period between 1997 and 2011 included three Dáil terms including the 28th, 29th and 30th Dáil Eireann (Parliament of Ireland). Detailed descriptions of the corpus are presented in section 3.7.

The findings of this research showed that there was no change in disparity between the volume of coverage afforded to male and female ministers. However the difference in the content of the articles decreased over time. Table 4.33 presents an overview of the findings of this research.

Gender Bias	**Summary of Findings**
Quantity	No change evident in difference of quantity of coverage afforded to male and female ministers
Content	Level of difference in the content of articles featuring male and female politicians decreasing over time

Table 4.33: Overview of Findings of Changes in Gender Bias

Gender Bias in Quantity of Coverage

No evidence was found of gender bias in terms of the quantity of coverage afforded to ministers in any of the Dáil terms studied. In each of the Dáil terms, the coverage of female politicians outnumbered that of male politicians. The magnitude of this difference also remained constant.

Figure 4.5 shows the average number of articles for male and female politicians per year in office. Female politicians featured in headlines approximately twice as much as their male counterparts. This analysis shows that gender bias or lack thereof, as measured by the volume of coverage afforded to ministers has remained constant in Ireland between 1997 and 2011.

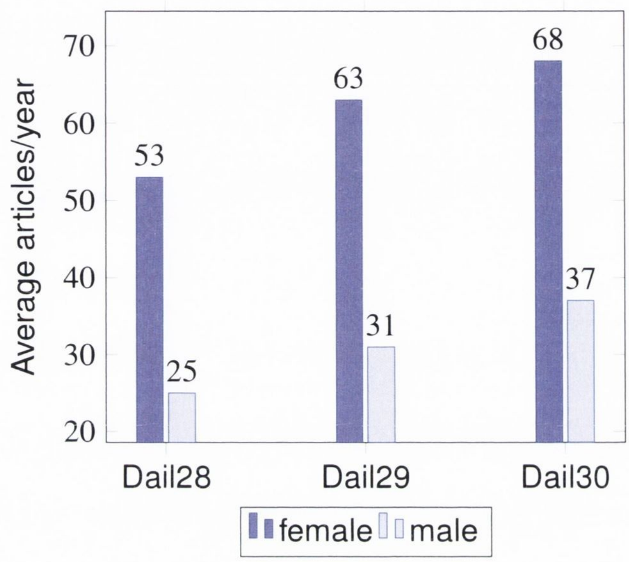

Figure 4.5: Articles Featuring Politicians in Headlines per Year in Office for Each Dáil term

Gender Bias in Content of Coverage

To examine the patterns of change in levels of gender bias between 1997 and 2011, sub-corpora were created categorising newspaper articles according to each Dáil term. Overall trends in gender bias were assessed by comparing the levels of accuracy achieved by text classification experiments in distinguishing between articles featuring male and female ministers. This indicated the levels of difference between the content of articles featuring male and female ministers. Findings showed that if gender bias exists in the substance of newspaper articles, may be decreasing over time.

The approach to text classification used in these experiments was based on findings from the first round of text classification experiments outlined in

Section 4.2. The support vector machine learning algorithm was therefore used along with a boolean representation of features.

The classification accuracies in all but three of the experiments reduced between the 29th and 30th Dáil (Table 4.34). This suggests that differences between the content of articles featuring male and female ministers decreased over time. This aligns with findings of Bystrom and Hennings (2013) and Norris (1997) that show differences in how male and female politicians are featured in the media, while still in existence, is decreasing over time.

Analysis of the discriminative features for these experiments showed that the reduced size of the corpora led to words relating to specific details about political events being used to predict the gender of the minister featured in the article rather than identifying overall patterns in how language is used. This contributed, for example, to the particularly high accuracy (92 percent) gained in using unigrams as features in classifying articles from the 28th Dail. For this reason, a deep analysis of the concordance lines of the discriminative features was not suitable. Analysis of levels of gender bias instead focused on abstracting overall trends in the levels of difference between articles featuring male and female ministers, as outlined in Table 4.34.

Feature Sets	28th	29th	30th	Difference(30th-28th)
Vocabulary measures	51.5	55.3	54.3	2.8
Adjectives from articles	69.6	66.5	71.3	1.7
POS tags	52.2	53.9	52.5	0.4
POS bigrams	53.6	54.6	52.9	−0.7
GI political words	55.1	51.8	53.7	−1.4
POS trigrams	59.2	57.7	57.4	−1.7
Stop words	55.5	53.6	53.5	−1.9
GI power words	64.8	58.5	62.2	−2.6
Adjectives from sentences featuring minister	66.1	59.0	63.3	−2.8
Sentiment	53.3	51.4	48.5	−4.8
Verbs from articles	63.8	59.5	57.3	−6.5
Verbs from sentences featuring minister	63.8	55.5	56.8	−6.9
Newspaper section	64.8	57.8	55.3	−9.4
GI action word	74.5	62.7	64.2	−10.3
Bigrams	95.2	83.8	84.4	−10.9
Unigrams	92.5	76.4	79.8	−12.6
Baseline	51.5	51.9	52.1	

Table 4.34: Differences in Classifier Prediction Accuracy For Each Dáil Term

4.2.5 Newspaper Sections and Gender Bias

Articles in different sections of the newspapers focus on particular topics and are often subject to different editorial guidelines. For example, subjective opinions are more evident in the opinion section rather than the front pages of newspapers. No previous research has addressed whether there are differences in levels of gender bias among sections of newspapers.

In order to explore the existence of gender bias in different sections of newspaper, sub-corpora were developed according to newspaper sections. Text classification experiments were run and the levels of accuracy compared. Differences in overall volume of coverage afforded to male and female politicians in each section were also examined. An overview of the findings is presented in Table 4.35 with full details provided in the subsequent section.

Section	Theme	Summary of Findings
All	Quantity	More coverage of female ministers in all except sport and motor sections
		Female ministers for transport featured in motors section of newspaper less than male ministers for transport
		More coverage of female ministers in business and hfinance section
Letters	Different Policy Focus	Female ministers more associated with social policy related to care of people
	Political Style	More active and assertive actions associated with male ministers while female ministers are associated with more passive and compliant actions.
Front Page		Female ministers associated with activities pertaining to communication while male ministers are associated with direct action.
Opinion and Analysis	Focus on Gender	Female ministers associated with the term 'woman' but male ministers not associated with the term 'man'.
Ministerial Roles	Policy Focus	Female ministers for transport featured in motors section of newspaper less than male ministers for transport

Table 4.35: Overview of Findings of Gender Bias in Different Newspaper Sections - Details Presented in Subsequent Sections

Quantity of Coverage

A breakdown of the number of articles featuring male and female politicians according to each newspaper section shows that female politicians receive greater coverage than male politicians in most sections (Table 4.36). In those sections where female politicians are represented more than male politicians, the scale of this difference is greater than those sections where male politicians are represented more than female politicians. The two sections that feature male ministers substantially more than female ministers are sections that are most stereotypically associated with male interests, *Motors* and *Sports*.

The greatest disparity in the volume of coverage afforded to male and female ministers is in the *Ireland* and *Business and Finance* sections. The *Ireland* section focuses on news events in Ireland. The strong association of female ministers with articles in both of these sections contradicts findings in previous studies which showed that male politicians are more strongly associated with economic and financial issues than female politicians (Brikse, 2004; Bystrom et al., 2004; Carroll and Schreiber, 1997).

Section	Female	Male	Diff.(F-M)
Arts	0.16	0.06	0.10
Business and Finance	8.32	2.08	6.24
Editorial	0.98	0.29	0.70
Education and Living	0.93	0.67	0.27
Entertainment	0.05	0.01	0.04
Farming	0.23	0.26	−0.03
Features	0.28	0.15	0.13
Front page	3.76	1.37	2.39
Health	1.74	0.12	1.62
Ireland	35.88	21.33	14.55
Letters	2.20	0.61	1.60
Lifestyle	0.03	0.00	0.03
Magazine	0.00	0.07	−0.07
Motors	0.03	0.13	−0.10
News features	0.96	0.40	0.55
Opinion and Analysis	1.76	0.59	1.17
Politics	1.71	0.59	1.12
Property	0.13	0.02	0.11
Science	0.03	0.01	0.01
Sport	0.28	0.88	−0.59
Travel	0.00	0.02	−0.02
Weekend	0.21	0.23	−0.02
World	0.23	0.27	−0.04
Total	60.44	30.26	30.18

Table 4.36: Number of Articles in Newspaper Sections

Gender Bias in the Content of Coverage

In order to explore whether there was variance in levels of gender bias in different newspaper sections, content from three sections of the newspaper was examined. These three sections were selected to capture differing styles in the papers. The sections included the *letters* section, *opinion and analysis* and articles from the *front page*.

The letters section was selected since the language in it is not subject to the main newspaper editorial guidelines. There is more likelihood therefore that gender bias may be expressed in this section. The opinion and analysis section also includes subjective opinions about events and individuals and is authored by journalists from the newspaper. The front page was selected as this section typically is the most news focused and editorially controlled section of the paper.

The text classification experiments carried out used feature extraction methods that were found to be most useful in earlier experiments. Support vector machine learning algorithms were also used along with a boolean approach to representing features.

Feature sets
Unigrams (single words)
General Inquirer Lexicons
Verbs
Adjectives

Table 4.37: Feature Sets Used for Newspaper Section Sub-Corpora

As expected, articles from the letters section showed the greatest difference in the content of articles featuring male and female ministers (Figure 4.6). Using unigrams as features for example yielded 86.7 percent accuracy in predicting the gender of the minister featured. This would suggest that if gender bias explained differences in the coverage of male and female politicians, the levels of bias are greatest in the letters to the editor. There was no consistent pattern in the classification accuracy achieved from front page articles compared with those from the opinion and analysis sections.

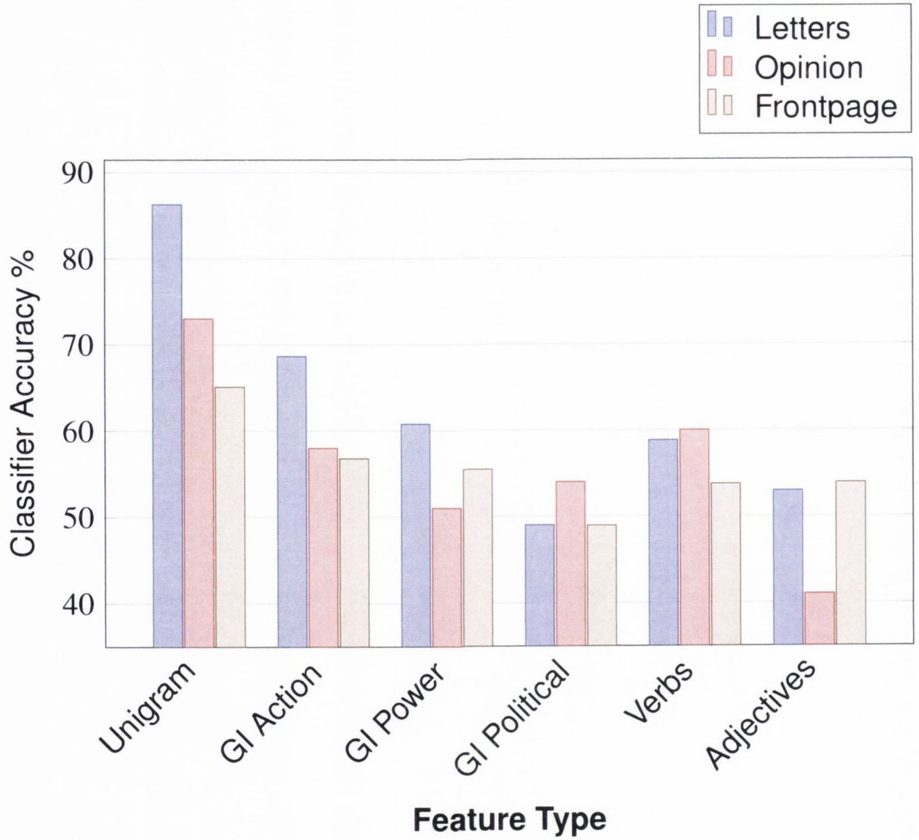

Figure 4.6: Classification Accuracy for Newspaper Sections

Letters to the Editor - Gender Differences in Content

There was a notably high level of difference between articles featuring male and female ministers in letters to the editor when single words were used as features (Table 4.7). Other feature sets which yielded high accuracy were the General Inquirer power and action lexicons. Analysis of the discriminative features showed that some did suggest evidence of gender bias while others were a result of a smaller corpus size allowing the machine learning algorithm to learn information about political events at a time and associate these with politicians of either gender. Accuracy results using single words as features yielded very high accuracy (86.3 percent). Analysis of the discriminative features and concordance lines showed that this was influenced by the fact that many of the letters featuring female ministers concerned Mary Harney while she was Minister for Health and Children. The letters concerning male politicians are distributed more evenly among politicians with different roles in cabinet. The machine learning algorithm therefore used single words that referred specifically to Mary Harney and issues of the time concerning health policy to infer the gender of the subject. This was less informative regarding gender bias than larger corpora which included a broader spectrum of politicians, roles and covered more political events and issues.

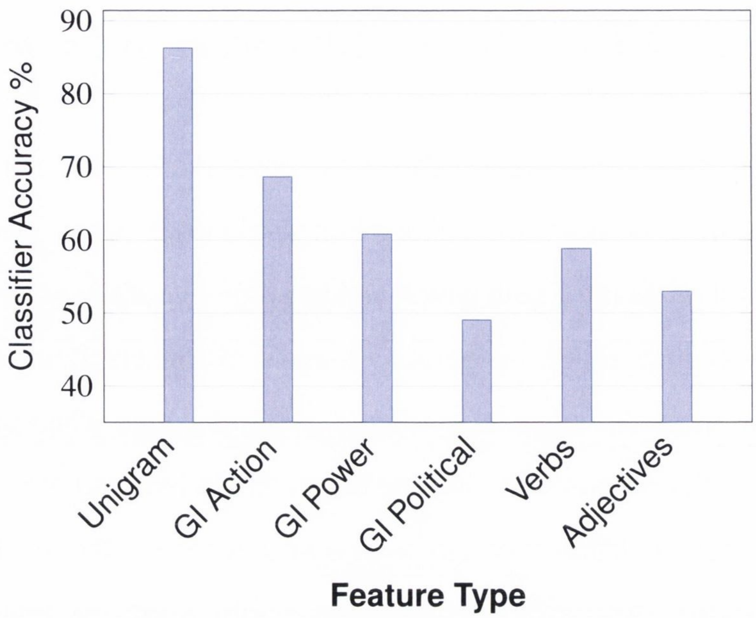

Figure 4.7: Classification Accuracy for Letters to The Editor

An association of female ministers with health policy discussions related to care of people, suggests a tendency to associate social policies with female ministers in media coverage (Bystrom et al., 2012). The term 'care' is a discriminative feature most associated with female politicians. Concordance analysis showed that mentions of care occurred primarily in conjunction with health policy. Accounting for differences in the duration male and female politicians were ministers for health, 'care' was mentioned in letters to the editor featuring female ministers over 2.5 times as often as those featuring male ministers.

Gender	Mentions of care	Years as Health Minister	Mentions/Year
F	31	6.3	4.9
M	13	7.3	1.8

Table 4.38: Mentions of 'Care' in Letters to the Editor

Words from the General Inquirer action lexicon associated with male politicians include 'drive', 'direct', 'process', 'prepare', 'progress' and 'argue'. To the researcher, these words suggested a more assertive tone than the terms associated with female politicians. These words included ‘continue’, ‘sort’, ‘support’, ‘seek’ and ‘contribution’ (Appendix A.9). Concordance analysis confirmed that the primary use of these words in the articles were in relation to policy. These passive action words associated with female ministers is reminiscent of the portrayal of female politicians as "condensing symbols" for existing policies and ideologies rather than the originators and proponents of them (Adcock, 2010).

Front Page - Gender Differences in Content

Compared with letters to the editor and articles from the opinion and analysis section, it was expected that the least amount of gender bias would be evident in this section as it was subject to the greatest level of editorial control. Classification results if they were to be taken as an indication of gender bias, supported this expectation (Figure 4.8). There was little difference in the articles featuring male and female politicians using General Inquirer political and power lexicons or verbs. Two approaches to feature selection showed significant difference in text classification results. These features were unigrams and general inquirer action words (Appendix A.11 and A.10).

Feature analysis of the unigram discriminative features showed that the classifiers identified specific political events and associated those with ministers of either gender. In the discriminative features identified using the General Inquirer action lexicon words associated with female ministers are more focused on communication while more words associated with male ministers concern action (Appendix A.10). These words were manually interpreted by the researcher and presented in Table 4.39.

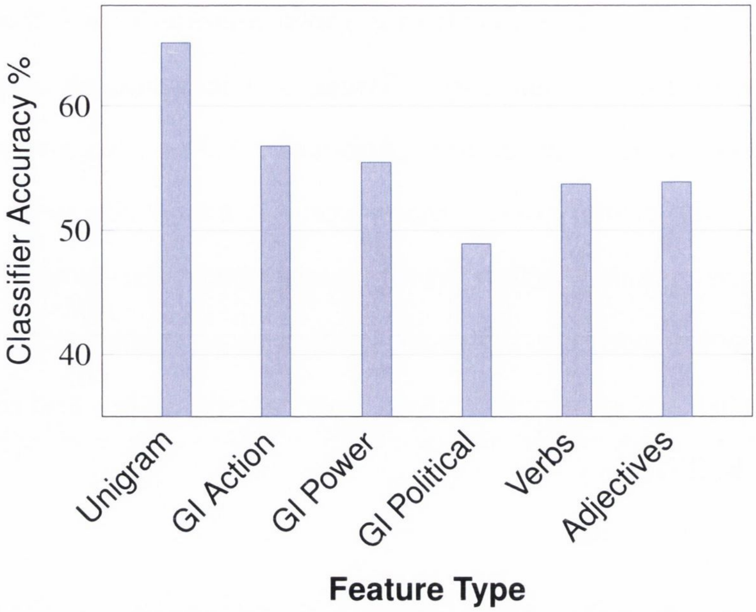

Figure 4.8: Classification Accuracy for Front Page Articles

Female Ministers	Male Ministers
offer	prevent
seek	whip
signal	carry
encourage	action
participate	task
proceed	plan
accommodate	effect
	extend

Table 4.39: Action Words Associated With Male and Female Ministers in Front-Page Articles

Opinion and Analysis - Gender Differences in Content

Three of the classification experiments conducted on newspaper articles from the Opinion and Analysis section yielded classification accuracy which suggested a substantial difference in the articles featuring male and female ministers (Figure 4.9. Analysis of the discriminative features showed some patterns which align with findings in earlier experiments suggesting gender bias.

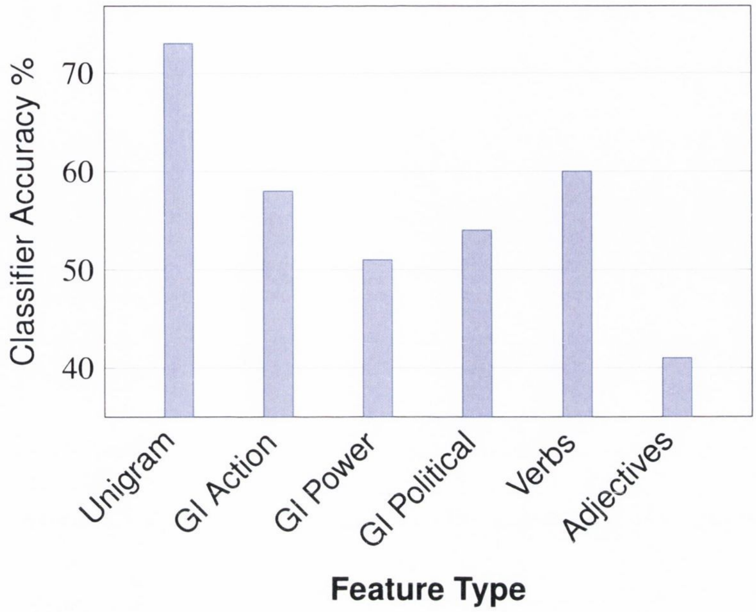

Figure 4.9: Classification Accuracy for Articles in Opinion and Analysis Section

Female politicians were associated with the term 'woman' while men were not associated with the term 'man'. This points towards a gender bias linking issues pertaining to women with female politicians or an increased focus on the gender of female politicians.

4.2.6 Ministerial Positions and Gender Bias

Sub-corpora were developed which categorised newspaper articles featuring ministers holding the same portfolios. This was done in order to identify whether levels of gender bias were linked with certain cabinet roles. The focus of this analysis was to highlight overall trends and patterns in comparison with those from the original corpus including all of the ministerial roles.

Theme	Summary of Findings
Quantity	Female ministers receive more coverage in all ministerial posts except that of Agriculture
Content of Articles	Ministers of Transport and Natural Resources show highest overall levels of difference in the content of the articles
Transport and Sport Ministers	Greatest disparity in the content of newspaper coverage using GI lexicons and adjectives as features
Social and Family Affairs	Greatest disparity in content of newspapers using verbs as features

Table 4.40: Summary of Findings of Gender Bias in Coverage of Ministerial Positions

Gender Bias in Quantity of Coverage

The quantity of coverage as measured by appearances in newspaper headlines for each role in cabinet posts is presented in Figure 4.10. This shows that female ministers were featured in the headlines of articles more frequently than their male colleagues in all except the role of Minister for Agriculture. In four of the departments, Health, Education and Science, Enterprise and Arts, Sports and Tourism they received substantially more coverage. These results indicate that there is no gender bias in coverage of male and female ministers in terms of quantity of coverage in all but one of the departments. If gender bias were to be measured by the quantity of articles, then these results would indicate gender bias in favour of male politicians.

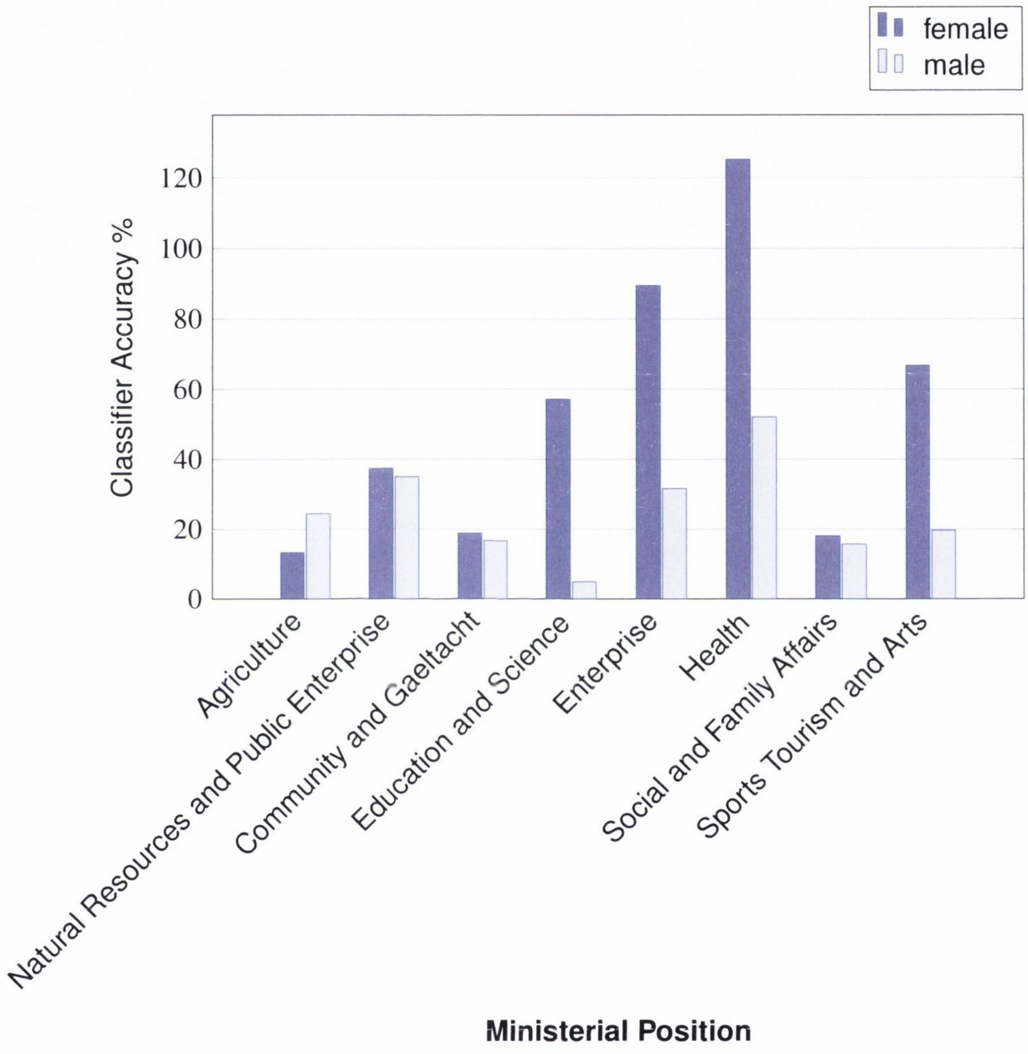

Figure 4.10: Average Number of Articles Featuring Ministers in Headlines for Each Year in Cabinet

Gender Bias in the Content of Coverage

There was variance in the level of accuracy gained by the classification experiments run on articles pertaining to ministers of different departments. Articles featuring ministers in the department of transport and natural resources for example showed the highest overall differences (Figure 4.11). However, feature analysis showed that many of the particularly high accuracy rates were a result of the size of the sub-corpora leading the algorithms to reply on terms associated with particular political events to predict the gender of the politician. An example of this is the high accuracy gained using verbs as features. Analysis of the discriminative features showed that this was a reflection of differing economic climates while the male and female ministers were in power. The female minister who held the position of social and family affairs did so after the economic collapse which began in 2008. Her male counterpart held this position while the economy was doing well. This lead to different subjects of concern in the content of the articles and features indicating this were used to discriminate between articles featuring male and female ministers. Larger corpora that incorporated articles from a broader range of political events between 1997 and 2011 relied less on words pertaining to particular political events to discriminate between articles featuring politicians from either gender.

High classification accuracy was obtained for articles featuring ministers for transport using sections of the newspaper as features. This showed a difference between how often male ministers for transport featured in the *motors* section compared with female ministers. This could suggest gender bias on the part of the editor of that section of the paper or it could be a reflection of differing political priorities on the part of the ministers. Further qualitative analysis of this coverage would be necessary to address this.

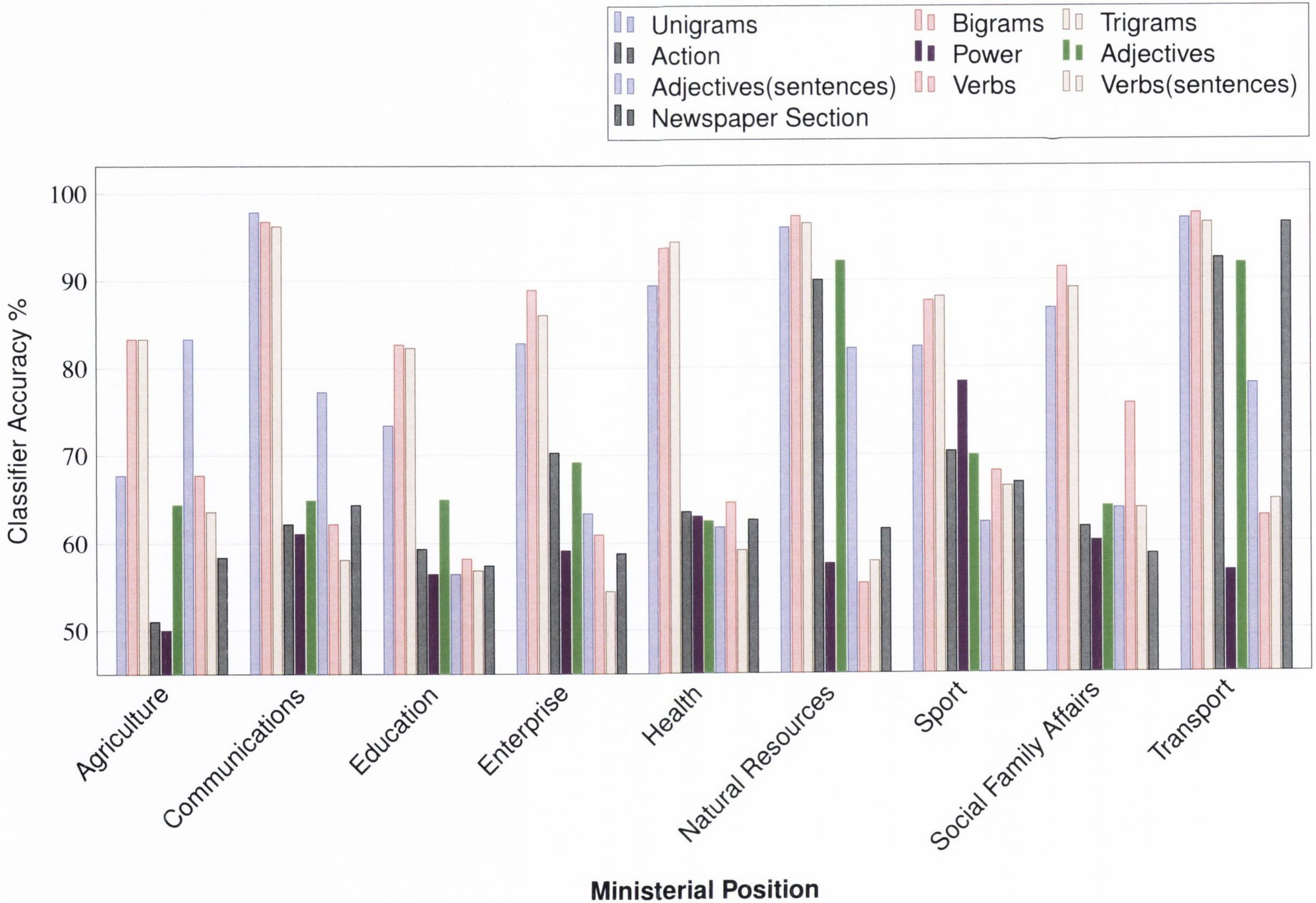

Figure 4.11: Classifier Accuracy for Each Ministerial Position

4.3 Corpus Analysis of Sentences Naming Ministers

This section presents the findings pertaining to gender bias in all newspaper articles which named ministers between 1997 and 2011. Evidence of gender bias is investigated by calculating the quantity of coverage of politicians and also examining the substance of that coverage using text classification methods. An overview of the findings of analysis of this corpus is presented in Table 4.41. These findings are broadly consistent with the findings from experiments described earlier in this chapter.

Theme	Summary of Finding
Quantity	Female ministers mentioned more in newspaper articles than male ministers
Political Style	Female minister's political style compared with Margaret Thatcher
	Portrayed as using sexual identity in political negotiation
Focus on Family	Marital status of female ministers more important than that of male ministers
Focus on Gender	Female ministers grouped with and evaluated against other female ministers
	Gender of female politicians unnecessarily mentioned
Policy Focus	Gender equality in politics associated with female ministers
	Policy in relation to mothers associated with female ministers
Personalised Coverage	Female ministers' dress was mentioned 18 times more than male ministers per year in office

Table 4.41: Overview of Findings of Gender Bias in Sentences Naming Ministers

The corpus was composed of 47,981 articles. The content of the articles varied widely, from articles in which politicians were central to the stories covered to those where politicians were just named incidentally. Including the entire articles in machine learning experiments resulted in a corpus where the dimensionality of the data was extremely high and therefore both computationally inefficient. Therefore, in order to isolate the text particularly relevant to the ministers, the corpus was further refined to extract text most likely to pertain to the ministers. This was conducted by extracting the sentences that named the ministers. This resulted in a corpus comprised of 90,674 documents, each containing a sentence that mentioned a minister. While this approach has the advantage of isolating terms most closely associated with politicians in an article, the sentences are removed from the overall context of the articles they appear in. The disadvantage of loosing some of the overall context of the articles was addressed by conducting experiments using the entire content of the articles in the corpus that was developed using the search criteria that a politician's name must appear in the headline.

4.3.1 Gender Based Difference in Quantity of Coverage

Female ministers were mentioned in articles over twice as often in the Irish Times as their male counterparts and almost three times as often in the Irish Independent (Figure 4.12). However, this analysis did not take into account the fact that some ministers feature more prominently in some articles. The more central the minister is to the story, the more often they are mentioned in the article. To address this issue, the number of times ministers were mentioned in sentences in articles was also counted. This showed that both the Irish Times and the Irish Independent mentioned female ministers in sentences over 2.5 times as often as male ministers (Figure 4.13).

These results were consistent with the disparity in volume of coverage between male and female politicians identified in the corpus based on articles naming ministers in headlines. This shows that female politicians are not only mentioned in news headlines more often but this disparity also translates to the number of times they are mentioned in the content of newspaper articles. These results suggest, not only that there is no gender bias in favour of male politicians but that there may be gender bias against them (Bystrom et al., 2001; Fuertes-Olivera, 2007; Miller et al., 2010; Pearce, 2008).

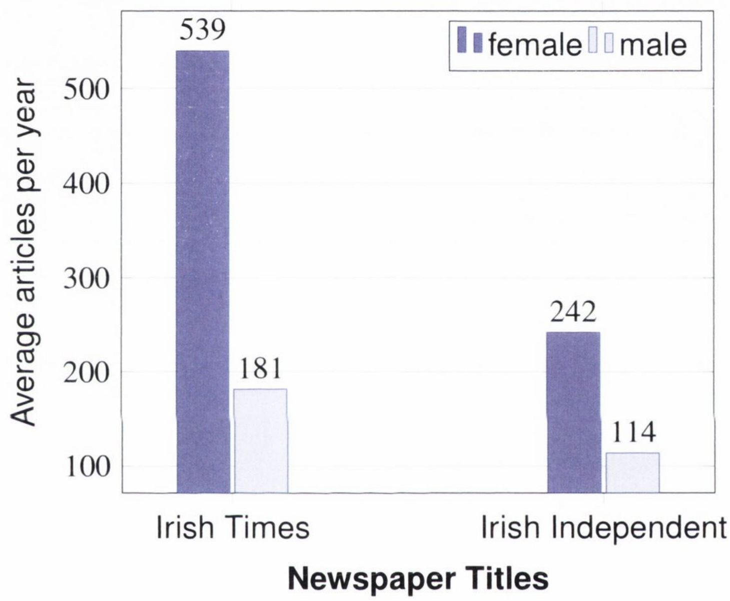

Figure 4.12: Average Number of Articles Mentioning Ministers

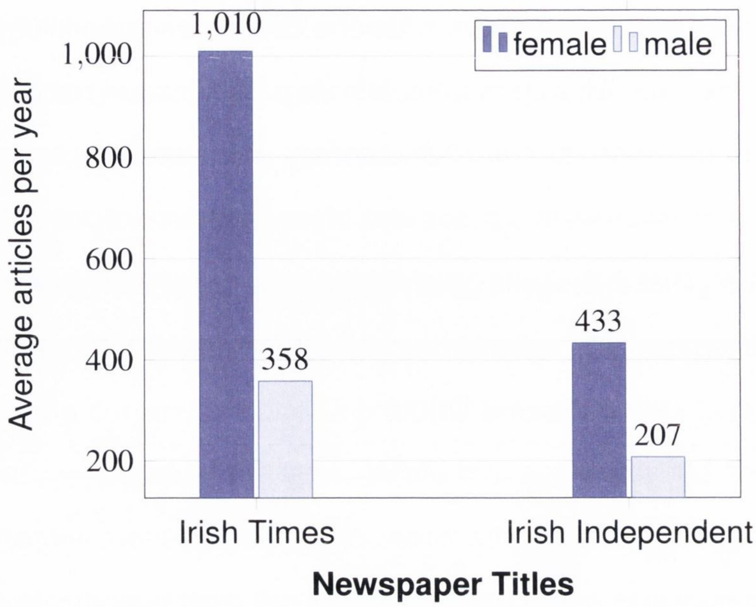

Figure 4.13: Average Number of Sentences Mentioning Ministers

Analysis of the volume of coverage afforded to each minister showed that coverage of Mary Harney exceeded that of any other ministers. Other female ministers, while not exceeding their male colleagues to the same degree, were all among the top ministers in terms of the level of coverage afforded to them. Figure 4.14 shows the average number of times per year that each minister was named in in an article in the Irish Times. This paper is used to compare the quantity of coverage among all of the ministers as it spans the entire duration of the research.

Mary Harney's coverage was almost twice that of the nearest male minister. This may be explained by the fact that she was Deputy Prime Minster for part of this time and was also leader of a smaller coalition partner. However, Eamon Ryan was also the leader of a coalition party and Brian Cowen, who was also a minister in this study, was Prime Minister from 2008. The trend of overall higher quantities of coverage being afforded to female politicians is also supported by the other four female ministers. The second highest in terms of volume of coverage was afforded to Mary O'Rourke. Overall, the five female politicians are in the top 50th percentile of all ministers in terms of the quantity of coverage they receive. Therefore, if as previous studies have found, gender bias against female politicians is manifested in lower volume of coverage being afforded to them (Bystrom et al., 2001; Fuertes-Olivera, 2007; Miller et al., 2010; Pearce, 2008), then these results would suggest that there is a gender bias against male politicians.

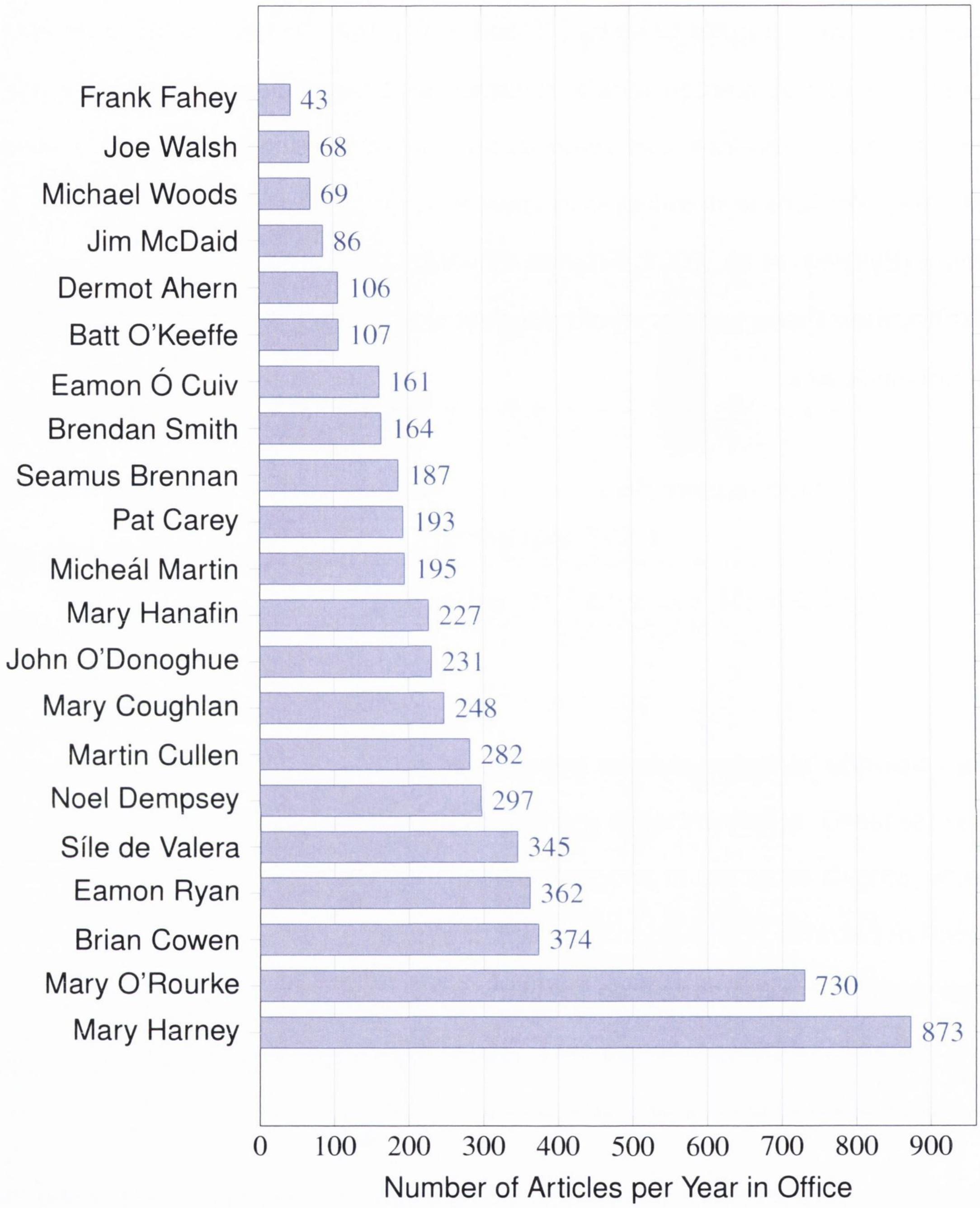

Figure 4.14: Average Number of Times Minister is Mentioned in an Article, Per Year in Office

4.3.2 Gender Bias in Content of Sentences

While overall classifier accuracies were below 60 percent, analysis of the discriminative features highlighted patterns that were relevant to gender bias.

As the accuracies were low, analysing the discriminative features alone to determine whether gender bias existed, was insufficient. Therefore, concordance analysis was an important part of verifying whether gender bias existed or not.

Feature Sets	Classifier Accuracy
Adjectives	56.6
GI action words	59.8
GI power words	53.8
Newspaper section	54.3
Unigrams	75.8
Verbs	57.7

Table 4.42: Classifier Accuracy for Sentences Naming Ministers

Feature Sets	Classifier Accuracy
Adjectives	65.5
GI action words	64.5
GI power words	59.2
Newspaper section	58.7
Unigrams	75.2
Verbs	57.4

Table 4.43: Classifier Accuracy for Articles Naming Ministers in Headlines

In text classification experiments on this corpus, the feature sets used were those found most useful in analysis of the corpus of newspaper articles featuring politicians in the headlines. Text classification accuracies for this set of experiments were lower overall than those gained from the corpus containing articles featuring ministers in headlines (Table 4.42). Classification accuracy for articles featuring ministers in headlines are presented for comparison purposes in Table 4.43. Two of the feature sets used, newspaper section and the General Inquirer power lexicon yielded classification accuracy of less than 5 percent over the baseline of 50 percent.

Given the results outlined in Table 4.42, four of the feature sets warranted feature analysis as these were the only experiments which resulted in accuracies that were 5 percent over the baseline. These feature sets included adjectives, General Inquirer Lexicon of action words, unigrams and verbs. Analysis of the discriminative features identified by the classifiers using these feature sets is described in the next section.

Political Style

The name 'Thatcher' is most associated with female politicians (Appendix A.12). Concordance analysis showed that this referred to 16 mentions of Margaret Thatcher, the former Prime Minister of Great Britain. Each of these mentions occurred in sentences featuring Irish female ministers. Frequently these involved comparisons between Mary Harney and Margaret Thatcher. Given that no male politician in the period between 1997 and 2011 was compared with Margaret Thatcher, this suggests a tendency to compare politicians of the same gender, as exemplified by the following excerpt:

> *The women I admire and find most attractive are those who have that tiny touch of testosterone which provides the perfect finish to their femininity; I am thinking of courageous politicians like Margaret Thatcher and Hillary Clinton and, closer to home, Mary Hanafin, Beverly Flynn and Mary Harney. (Irish Independent)*

This comparison of female politicians with other female politicians aligns with findings mentioned previously in this chapter where in evaluating Mary Harney's political career, her gender was emphasised unnecessarily in the description of her as the "most successful female politician in the history of the State" (Section 4.2.3).

The word 'female' is also associated with female politicians while the word 'male' is not associated with male politicians. Concordance analysis of the word 'female' identified a few themes in how the word was used. One of these is in relation to an evaluation of female politicians (Table 4.44). Female ministers were compared with other female politicians but not with male politicians. This supports the findings set out in Section 4.3.2. This tendency towards same-gender comparisons of female politicians, particularly their political style, points to gender bias in the highlighting of their gender and portrays them as outside the male political norm (Miller et al., 2010; Norris, 1997).

Gender-based evaluation
"I can't believe Mary Harney's a female - she seems to lack a lot of humanity.
...Compared to the other female senior Irish politicians, she is eminently likeable: less patrician than Mary Hanafin; less abrasive than Mary O'Rourke; less grumpy than Mary Harney...
...good sense of humour, she is fondly regarded as 'one of the lads' around Leinster house, in contrast to other female TDs.
Mary Harney, Mary Robinson and Mary McAleese have bottom - you can fill in the shifty female names who do not have.
The most successful female politician in the history of the State, Ms Harney has been PD leader for 13 years and Tanaiste for nine.
One problem with the theory of a wish list of women leaders is that our most prominent female representatives in Cabinet, the Marys Coughlan and Harney, haven't exactly covered themselves in glory, so far...
...all this, despite never playing the gender card (and even playing against it), Coughlan couldn't avoid the fact she was female.

Table 4.44: Concordance Examples of 'Female' in Context of Evaluation

The word 'kiss' is most associated with female ministers (Appendix A.13). This aligns with the finding set out previously in this chapter which suggested a sexual identity was created relation to female ministers (Section 4.2.3). There were only 8 references to ministers kissing or being kissed and all but 2 of those pertained to female ministers. Two of the uses of the term were metaphorical. One of these, 'cold kiss', reflects the tendency identified by Mills (2002) to use metaphors to describe female sexuality in terms of heat or lack thereof. In this instance the term 'cold' is used metaphorically to describe the nature of the kiss. However that kiss being referenced is also metaphorical. This concordance analysis included different forms of the verb including kiss, kissed and kissing. These findings suggest that a female minister's political style incorporates use of their sexuality while male political style involves political negotiation and diplomacy.

Gender	Concordance line
F	First we had the lips-on-lips kiss between Michael McDowell and Mary Harney.
F	... must have felt his political lifeblood chill as Harney's cold kiss of political support sealed his fate.
F	... Calling Harney's endorsement the kiss of death for Ryan, he reads on with obvious enjoyment...
F	... when elected leader of the Progressive Democrats, by embracing and kissing former party leader Mary Harney.
F	No heartbreak hotel as Ahern and Coughlan kiss and make up...
F	... Nor even when Mary O'Rourke kissed Donie Cassidy after he took her seat.
M	...a huddle of media, his mother Mildred planted a big kiss on his cheek.
M	I saw Bertie kissing John O'Donoghue...

Table 4.45: Concordance Examples of the Word 'kiss'

Focus on Family

The word 'married' is associated with female ministers (Appendix A.14). Analysis of the sentences containing the word 'married' showed that many of them refer to a female politician's marital status and others refer to policy issues concerning married couples. While the term 'single' is also associated with female ministers, concordance analysis showed that none of those instances referred to their marital status but were used within the context of policy issues. The following extracts demonstrate the tone of discussions in relation to one female politician's marital status and whether she had children:

- *... Mary Harney admitted that she might not have made a life in politics had she married and had babies in her youth.*
- *Harney, who only married in recent years and has no children of her own, has made significant personal sacrifices to pursue her long, but varied, career in politics.*

In these extracts Mary Harney is described as 'admitting' the incompatibility of her life in politics with family. She 'only married' recently and has 'no children of her own'. Not having children is portrayed as a significant personal sacrifice and her life is portrayed as being less fulfilled without children.

Focus on Gender

As mentioned in a previous section, the word 'female' is associated with female politicians but 'male' is not associated with male politicians. Table 4.46 presents concordance lines where the word 'female' was used to point out the gender of a female politician. Many of these references to gender seem unnecessary, suggesting a bias which focuses on a female politician's gender (Miller et al., 2010; Norris, 1997). While some uses may seem balanced and objective such as pointing out the best male and female politicians in parliament, (eg. 'Jim Higgins the most impressive male TD in danger of losing his seat Jan O'Sullivan, the most impressive female TD in danger of losing her seat...'), the division of politicians into male and female categories in evaluating their performance implies differences in their performance which prevents consideration of politicians as one group regardless of gender.

The word 'Marys' was identified as a discriminative feature identifying female politicians. This was used in the media as a term to reference the female politicians who shared first names. This grouping of the female politicians emphasised their gender and grouped them based on this. While pointing out their shared first names may not be gender biased in itself, the outcome of it is to reinforce the grouping of female ministers based on their gender.

Focus on Gender

...despite all his great claims to be the political mastermind of the age, it was Harney's female intuition which mostly came to the fore.
The job has been very much a female preserve, as Niamh Bhreathnach and Mary O'Rourke held it for long periods in recent years.
She was joined by fellow female Cabinet colleagues, Mary Hanafin and Mary Coughlan.
Previous winners, selected by the female members of the Oireachtas under the chairmanship of Minister Mary O'Rourke...
Jim Higgins the most impressive male TD in danger of losing his seat Jan O'Sullivan, the most impressive female TD in danger of losing her seat...
New Taoiseach Brian Cowen yesterday made Mary Coughlan the most powerful female politician in the country as he opted to reshuffle Bertie Ahern's Cabinet rather than wield the axe.
Mr Cowen gambled on personal friendship and fresh female appeal when he selected former agriculture minister Mary Coughlan for the deputy leadership of Government.
Mary Coughlan is not the first of his female targets.
... And the presence of two female Irish ministers, Mary Coughlan and Mary Hanafin, on the delegation led by Taoiseach Bertie Ahern is attracting attention.
Harney is currently The mother of the Dáil the female member with the longest unbroken service.
Mary Harney is the longest-serving female TD - she was first elected for Fianna Fail in 1981, a year before Mrs O'Rourke.
... a half year ago a cabinet reshuffle led to somewhat of a surprise when Mary Coughlan was appointed the first female Minister for Agriculture.
Mary Harney is one of the most successful female politicians this country has ever known.
Harney's candidature, should it occur, has one potential advantage for the Government as a female nomination may help as it seeks to hold on to an Irish place...

Table 4.46: All Concordance Examples of 'Female' Focusing on Minister's Gender

Different Policy Focus

Discussions of women in politics is associated with female politicians as can be seen in the concordance extracts in Table 4.47. The apparent lack of association between male politicians and the topic of women in politics portrays the issue as one only concerning and of benefit to women.

Discourse on Women in Politics

The low rate of female participation raises uncomfortable questions for State agencies, for the Government and for Irish business culture in general, Ms Harney said.
Tanaiste Mary Coughlan laid on the charm and a drinks reception for all female parliamentarians on Wednesday evening.
Of 150 winners in 50 years, only 11 have been women, though there have been some notable female finalists, such as the Tanaiste, Mary Harney, former president Mary Robinson and the Minister for Social and Family Affairs...
Ms Mary Coughlan, his colleague in the adjoining Donegal South West constituency, was the only new female to be appointed to Cabinet.
Famously, the Kildare Street Club only started accepting female members in the mid-Nineties, and once refused membership to Health Minister Mary Harney because of her gender.
Harney, in common with many women TDs, is instinctively opposed to quotas but not happy with the current level of female representation.
Hanafin, of course, like Hillary, would also offer primary voters the attraction of electing Ireland's first female taoiseach.
..."women in politics" was no longer the issue it once was, given that both our President and our Tanaiste are female, another book on the subject will be launched by Mary O'Rourke next week.
Conor Lenihan, for instance, regardless of their sex, but sometimes I have to agree with Lucinda Creighton's point that many female TDs are judged far more harshly than their male counterparts.
Way back when Mary Harney was a girl, she sought out other female deputies in the hope they'd meet regularly and quietly advance some women's issues in Dáil Eireann.

Table 4.47: Concordance Examples of 'Female' in Discourse on Women in Politics

Gender issues are seen as a concern for female politicians but not male politicians. The term 'sexist' is associated with female politicians. On analysing these sentences, most occurrances in the context of a description of the sexism of the political environment or their response to it (Table 4.48). Sexism is then portrayed as as an issue concerning only female politicians, aligning with the finding that discussions of women in politics are associated only with female politicians.

Political Discourse
Harney retorted that he had used abusive and sexist language.
Yet this rebuke from Mary Harney about sexist language came in discussion of a serious matter
Tanaiste Mary Harney condemned Michael McDowell for his use of sexist language in his statement.
Using sexist language about a lady like Mary Harney?
an article opposing the referendum, were "absurd", "personally abusive and sexist".
O'Rourke, who has perhaps best negotiated the shark-infested waters of sexist politics, but who is often, frequently disparagingly, referred to as
is hard to escape the feeling that there is something sexist about the way the newspapers keep picking on Ms Coughlan.
Coughlan said: If the deputy wishes to throw a condescending, sexist remark across the House, that is fine.
Fine Gael's Leo Varadkar slipped up when he rather sexistly compared Coughlan to Sarah Palin.

Table 4.48: All Concordance Extracts of 'Sexism'

Such gender based differences in policy discussions is seen in the association of the word 'mothers' with female politicians. Concordance analysis show that this word is used primarily in the context of discussions of policy concerning mothers. This supports research which showed that female ministers are more likely to be associated with issues that are considered to be women's issues (Brikse, 2004; Carroll and Schreiber, 1997; Kahn, 1996; Scharrer, 2002).

Along with being associated with different policy, evidence suggests that male politicians are associated with policy more generally. Of the top 100 single words as discriminative features, more terms associated with male politicians refer to policy issues while terms associated with female politicians are mostly general descriptive terms. This aligns with previous research which found that coverage of male politicians is likely to be more focused on policy issues (Kahn, 1996; Kahn and Goldenberg, 1991).

Personalised Coverage

Analysis of this corpus showed that female minister's physical appearance was commented on more than male ministers'. The word 'dress' was identified as a feature identifying female ministers. While the overall number of occurrences of the word are low, the vast majority of those occurrences are in connection with female ministers. Female Ministers dress was commented on 18 times more than male ministers (Table 4.49).

The concordance examples presented in Table 4.50 demonstrate how the female ministers' dress was commented on. There were two instances where male ministers' dress was described. One was in neutral terms. However, the other described Minister Eamon Ryan as dressed like a 'French intellectual' in a derisive tone.

No. Occurrences	Years in Office	Mentions/Year
17	38.6	0.44
2	84.1	0.02

Table 4.49: Mentions of 'Dress' Per Year in Office

Personalised Coverage
Ms Harney wore a long, pale lilac dress and coat.
Minister for Health Mary Harney, dressed in an exotic green blue Shanghai Lil style of trouser...
Agriculture Minister Mary Coughlan dropped in, dressed quite sedately which was a disappointment after her blingtastic ensemble...
Agriculture Minister Mary Coughlan, dressed in a seriously bling silver dress and jacket, had a big winner with Incline...
Instead, Mary Coughlan, dressed in a chic grey and black top and smart skirt...
Mary Coughlan, seemingly clad in a grey dressing gown, gazed disinterestedly from the government benches.
...mocking Mary Coughlan in the Dáil for her "colour co-ordinated" dress sense.
Ms Hanafin and Ms Coughlan always dress suitably and soberly.
O'Leary's favourite politician, Tanaiste Mary Coughlan, was in a grey dress.
...or, more recently, Mary Coughlan's floral dress that distracted attention from Cowen's statement to the nation...
Tanaiste Mary Coughlan was there, dressed in a hot pink number and accompanied by her son...
The party's pin-up wunderkind, Eamon Ryan, arrived to open proceedings dressed like a French intellectual...
Other Dáil contenders having a punt included casually dressed Communications Minister Noel Dempsey, Junior Minister Sean Power and former...

Table 4.50: All Concordance Lines of 'Dress'

4.4 Corpus Analysis of 2011 Irish Presidential Election Coverage

This section presents the findings of experiments conducted on the corpus of articles covering the 7 candidates of the Irish Presidential Election, two of which were female. An overview of these findings is presented in Table 4.51. The criteria used were to identify articles where the election candidate was named in the headline. The number of articles in the corpus totalled 705 including 195 articles featuring female candidates and 610 articles featuring male candidates. Each article was manually checked to ensure that it was relevant to the election candidate in question. A detailed description of the corpus was presented in Chapter 3, section 3.6.2.

Theme	Summary of Finding
Quantity	No gender bias evident in volume of coverage
Personalised Coverage	Use of more informal version of female candidates' names
Family focus	Female candidates' spouses featured more than male candidates' spouses
Use of Stereotypes	Stereotypical portrayals of candidates' spouses
	Female candidates portrayed as less in control of emotion

Table 4.51: Overview of Findings of Corpus Analysis of Coverage of Presidential Election

4.4.1 Evaluation of Approaches to Text Classification in Detecting Gender Bias

A full range of feature sets were used in this corpus. The smaller size of the corpus enabled feature types that significantly increase the size of the dataset, for example tri-grams, to be used. This section presents an evaluation of the feature types which were found to be informative regarding gender bias. Gender bias in the text is then explored through analysis of the discriminative features.

Analysis of the feature sets showed that those that were valuable in terms of highlighting potential evidence of gender bias were the same as those found valuable in experiments on the corpus pertaining to cabinet ministers. These are shown in Table 4.52 along with the classification accuracy gained using each feature set. This analysis addresses the first research question concerning which approach to feature extraction was found to be informative in identifying gender bias in text.

Feature sets which yielded a classification accuracy of below 5 percent over the baseline of 53.7 percent were not considered valuable in discriminating between texts featuring male or female candidates. Experiments using other features resulted in high classification accuracy. However, analysis of the discriminative features showed that the accuracy gained was either a result of reliance on particular events in the campaign to indicate gender bias rather than patterns in language use or the feature sets themselves were not informative in relation to the research question. For example, although parts of speech n-grams were useful in distinguishing between articles featuring male and female ministers the were not interpretable in relation to the topic of gender bias. This indicates that there may be differences in style in articles concerning male and female election candidates, but the discriminative features do not facilitate interpretation of the discriminative features in relation to gender bias.

Informative Features	Precision	Non-informative Features	Precision
Single Words	85.6	Character n-rams	87.1
GI Action Words	80.7	Bigrams	85.9
Verbs	76.1	Trigrams	83.8
Adjectives	74.6	Part of Speech Trigrams	74.0
GI Power Words	74.6	Part of Speech Bigrams	64.3
Stop words	66.9	Sentiment	56.6
		Part of Speech	61.2
		Author Gender	54.5
		Section	57.1
		Word Length	53.7
		Headline Length	53.7
		Headline Hapax	53.7
		GI Political Words	53.7
		Word Count	53.7
		Content Hapax	53.5

Table 4.52: Classification Accuracy for Features Used to Analyse Presidential Election Corpus

Feature Sets	Precision	Analysis of feature sets
Informative Features		
Unigrams	85.6	High accuracy and interpretable discriminative features
GI action words	80.7	
Verbs	76.1	
Adjectives	74.6	
GI Power Words	74.6	
Non-informative Features		
POS Trigrams	74.0	High accuracy signaling possibility of differences in style. However, discriminative features not interpretable in relation to gender bias.
Stop words	66.9	
POS Bigrams	64.3	
POS	61.2	
Character n-rams	87.1	High accuracy. However, features highlighted same patterns as using unigrams as features but results less interpretable.
Bigrams	85.9	
Trigrams	83.8	
Section	57.1	Low classification accuracy
Sentiment	56.6	
Author Gender	54.5	
Word Length	53.7	
Headline Length	53.7	
Headline Hapax	53.7	
GI Political Words	53.7	
Word Count	53.7	
Content Hapax	53.5	

Table 4.53: Evaluation of Feature Extraction for Presidential Election Corpus

4.4.2 Gender-Based Difference in Volume of Coverage

There was no clear evidence of gender bias regarding the quantity of coverage afforded to male and female candidates (Figure 4.15). While on average, women received 80 percent of the coverage male candidates did (Table 4.54), there was a large disparity between the amount of coverage afforded to the two female candidates diminishing the apparent equality the average calculations suggest.

Candidate Gender	Total Articles	Number of Candidates	Average Per Candidate
Male	610.0	5.0	122.0
Female	195.0	2.0	97.5

Table 4.54: Overall Volume of Coverage in Presidential Election

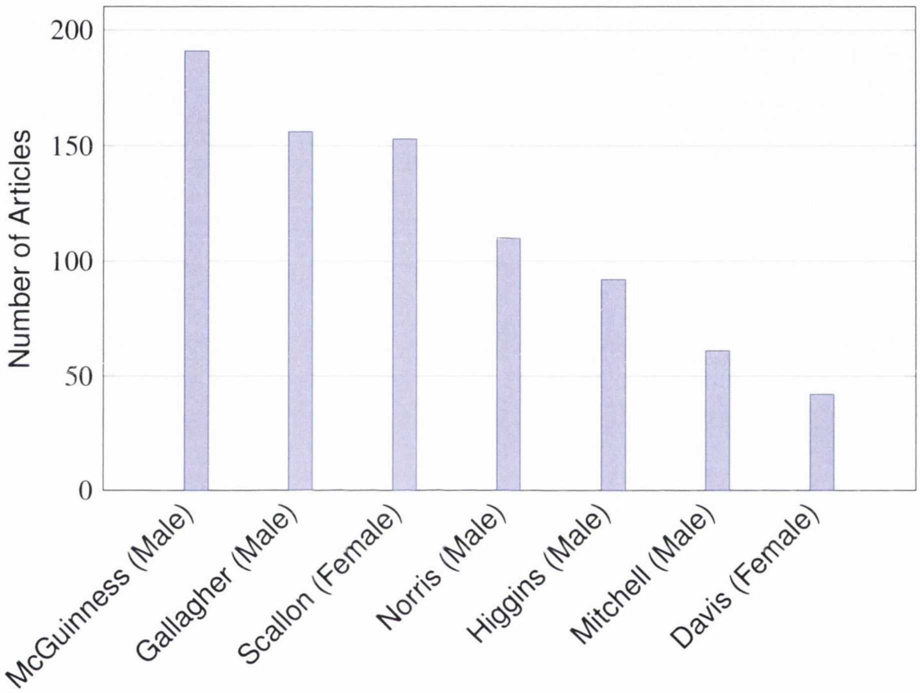

Figure 4.15: Number of Articles Featuring Presidential Election Candidates in Headlines

The profiles of the candidates were not directly comparable. One of the female candidates, Dana Rosemary Scallon, was third in terms of the quantity of coverage afforded to her. This was almost equal to the coverage of Sean Gallagher, who was the frontrunner for much of the campaign. The other female candidate, Mary Davis, received the least amount of coverage in terms of how frequently she appeared in newspaper headlines. This may have been explained by the fact that she was the only candidate who came to the campaign without already having a public profile. The other candidates were prominent politicians or had celebrity status in Ireland. In contrast to previous studies which concluded gender bias because of the fact that a female candidate's popularity was disproportional to the amount of coverage they received (Heldman et al., 2005), Mary Davis was never a front runner in the election.

4.4.3 Gender Bias in Content of Articles

This section presents the analysis of the discriminative features identified as important in distinguishing between articles featuring male and female presidential election candidates. The main theme which emerged from the results was that the coverage of female candidates was more personalised than that of male candidates. There was also evidence of stereotypical portrayal of candidates in relation to emotion.

The feature sets that were highlighted as potentially informative are as follows:

- Unigram
- General Inquirer Action Lexicon
- General Inquirer Power Lexicon
- Adjectives
- Verbs

Analysis of the discriminative features showed that the higher classification accuracy gained from experiments on this corpus relative to the accuracy gained from the corpus pertaining to cabinet ministers was a result of a smaller dataset allowing information about specific political events to be used to infer the gender of the politician in an article. Specific events and issues were used to discriminate between articles featuring male and female candidates rather than overall trends in the language being focused on. However, eliminating event

Naming Presidential Candidates
Dana is Dana. I made tea during her interview on Prime Time...
All the candidates, except Mr Norris and Dana Rosemary Scallon...
Mary Davis may be the only threat to Michael D....
...the unofficial Senator David Norris for president Facebook page...
...seven people on the ballot paper: Mr Norris, Dana, Michael D Higgins (Labour), Gay Mitchell (FG), Martin McGuinness (SF) and Independents Sean Gallagher and Mary Davis...

Table 4.55: Examples of the Naming of Presidential Candidates

related features highlighted some patterns in the use of language which could indicate gender bias. These patterns are presented in this section along with concordance analysis to evaluate the context of the use of certain terms.

Personalised Coverage

Mentions of candidate surnames are more closely associated with male candidates rather than female candidates. While this would seem to align with previous research citing gender bias in the tendency to use more informal versions of women's names in the media (Page, 2003), one of the female candidates had a public profile previous to the election and her commonly used name was her first name, Dana. Analysis of concordance lines verified that this name was used in the media to refer to her. Micheal D Higgins, a male election candidate, was also often referred to as 'Michael D', a name he was commonly known by. However, it was not as widely used as 'Dana'. The concordance extracts shown in Table 4.55 present examples of the differences in the formality used in naming election candidates.

Family Focus

The terms 'family' and 'husband' are associated with female candidates. The term 'wife', however, is not listed as an important discriminative feature identifying male candidates. Analysis of concordance lines showed that female candidates' spouses were mentioned then the male candidates' spouses. Individually however, there is variance between the candidates of both genders (Figure 4.16). Sean Gallagher's spouse is mentioned more often than any other male candidate is, reflecting the fact that a public identity was created for his wife as a part of his campaign strategy. Dana Rosemary Scallon's husband being mentioned more than any other candidate's spouse may have been because of an event where he was involved in a car accident with her. However, this does not seem to justify the vast difference in the number of mentions. Mary Davis's husband however, received among the lowest number of mentions. The variance shown here demonstrates the limitations of studies which focus on small samples of election candidates or politicians.

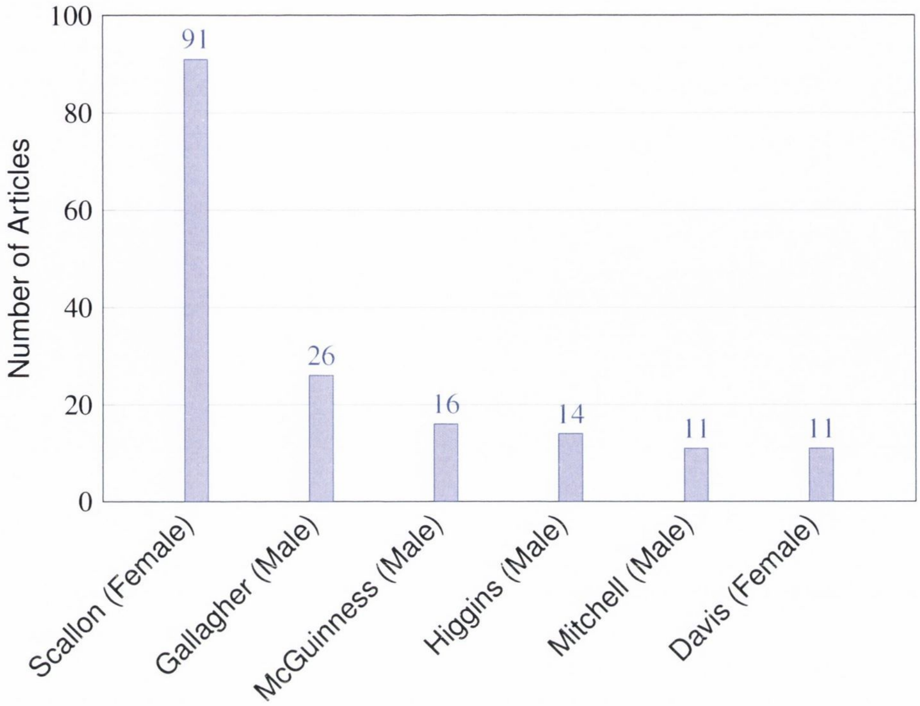

Figure 4.16: Mentions of Candidate's Spouse

Qualitative differences are evident in how the spouses were portrayed. This is evident in concordance lines presented in Table 4.56. These examples were purposively selected to illustrate ways in which spouses are portrayed. The examples also include the spouses of ministers of both genders being spoken of in a neutral and balanced ways. Some instances illustrate how the male candidates' spouses were described as being in a supportive role. For example Sean Gallagher's 'beautiful' wife, Trish, is described as being 'at his side' during a canvass, while Dana Rosemary Scallon's husband 'kept very close'. While these actions are the same, there is a tone of being the stereotypical supportive wife implied by using the commonly used term 'at his side'. The term used to describe husbands who accompany their wives is 'keeping close' which implies, by contrast, a protective intent. Other male candidates' spouses were portrayed as following their husband 'eagerly' and 'grasp[ing]' their hands.

Describing Presidential Candidate's Spouses
The big secret weapon of Sean Gallagher is his beautiful wife.
...his wife Trish who has been at his side throughout the campaign.
He instinctively put out his hand and his wife Sabina grasped it. The two walked hand in hand...
Gay Mitchell fearlessly led the way, followed eagerly by his wife, Norma.
...released a bank statement for Mr McGuinness and his wife, which details their expenditure...
The candidate was accompanied by her husband Damien Scallon, who kept very close during the walkabout.
Last week the election candidate and her husband had claimed that a tyre which blew out.
Ms Davis confirmed both she and her husband Julian were on the board of Social Entrepreneurs Ireland.
Davis denies role in charity contract going to husband's PR firm...

Table 4.56: All Concordance Examples of 'Husband' and 'Wife'

The word 'family' is the most important feature identifying female candidates. Analysis of concordance lines of the word 'family' showed that the vast majority of instances of the word 'family' pertained to Dana Rosemary Scallon. These mentions pertained to specific events during the campaign that involved her family. In contrast with this, there was only one reference to Mary Davis's family. This variance between the kind of coverage of the two female candidates prevents generalised conclusions regarding gender bias.

Use of Stereotypes

The most important discriminative feature used to distinguish between articles featuring female candidates was the word 'upset'. This was uncovered using the General Inquirer lexicon of action words as features (Appendix A.16). There was not one word describing an emotional state in the top 50 discriminative features associated with men. Concordance analysis showed that while there was a wide variety of contexts of the use of the term 'upset' it was only used directly in relation to female candidates. This aligns with Pantti (2005)'s finding that female politicians are more likely to be portrayed as less in control of emotion.

Subjective Language

Concordance analysis shows that each adjective identified as a discriminative feature by the classifier relates to specific events in the campaign. There is however a difference in the levels of subjectivity of the adjectives associated with male and female politicians. There is a tone of dramatisation associated with female candidates but not with the male candidates. Table 4.57 shows these adjectives.

Adjectives Associated with Female Candidates
malicious
huge
defamatory
terrible
disastrous
outrageous

Table 4.57: Subjective Adjectives Associated with Female Candidates

4.5 Conclusion

This research presented a new approach to analysing text for gender bias. A broad range of techniques for analysing text were explored and the most useful of these identified. The support vector machine learning algorithm was found to yield the best results along with a binary approach to text classification. The feature types which uncovered the most relevant information regarding gender bias were single words, the General Inquirer Lexicons of power and action words, adjectives, verbs and newspaper sections. These findings address the first set of research questions of this research.

The findings showed that there were differences in how Irish newspapers covered male and female politicians and some of these could be attributed to gender bias. Many of the themes which emerged aligned with the main themes identified in the literature review of feminist research on gender bias and of bias in the representation of politicians in the media. These findings are presented below according to each theme identified.

Quantity of Coverage:

- More coverage afforded to female ministers
- Differences in volume of coverage of ministers remained consistent between 1997 and 2011
- Only male Ministers for Agriculture received more coverage than female ministers per year in office (Figure 4.10)
- Male ministers dominate only in *sport* and *motors* sections of newspapers

Family Relationships and Roles:

- Minister's husband mentioned 4 times as often per year in office as male minister's spouse
- Presidential election candidate's husband mentioned 3 times more than male candidate's spouse
- Stereotypical portrayals of political spouses
- Marital status of female ministers more important
- Conflict between demands of job and family life emphasised in relation to female ministers
- Male ministers associated with discussions of families in economic terms while female ministers are associated with discussions of family in relation to social policy or in personal terms.

Different Policy Focus:

- Stereotypical association of male and female ministers with policy issues
- Discussions of work-life-balance policies associated with female ministers
- Childcare policy primarily associated with female ministers
- Female ministers portrayed as supporting gender equality policies while male ministers portrayed as under pressure to conform to them
- Policy in relation to mothers associated with female ministers
- Gender equality in politics associated with female ministers
- In Letters section, female ministers more associated with social policy related to care of people
- More coverage of male ministers in relation to policy
- Female ministers featured more predominantly in highest profile newspaper sections
- Female ministers for transport featured in *motors* section of newspaper less than male ministers

Focus on gender:

- Female ministers grouped with and evaluated against other female ministers
- Gender of female politicians mentioned unnecessarily

Use of Stereotypes:

- Association of political drinking culture with male ministers
- Sexualised identity created for female ministers
- Female candidates portrayed as less in control of emotion

Political Style:

- Female ministers portrayed as more accepting of government policies
- Association of female ministers with negative political communication style
- Female ministers compared with other female ministers
- Sexual identity portrayed as part of ministers' political negotiation strategy
- In the *Letters* section, more active and assertive actions are associated with male ministers while female ministers are associated with more passive and compliant actions
- In *Front page* articles, female ministers associated with activities pertaining to communication while male ministers are associated with direct action.
- In *Opinion and Analysis* section, male ministers more associated with actions related to policy than female ministers.

Personalised Coverage:

- Female ministers' dress was mentioned 18 times more than male ministers' per year in office
- There was more personalised coverage of female ministers
- Female ministers more likely to be described as 'formidable'
- More informal version of female ministers and candidates' names used

Masculine Narrative/Methaphors:

- Male ministers associated with terms pertaining to the military, institutions of state and sport

Gender Bias over Time:

- Level of difference in the content of articles featuring male and female politicians decreasing over time, suggesting a decrease in gender bias

The findings of this thesis showed that coverage of politicians in Irish newspapers contains evidence of gender bias. It also demonstrated how text classification techniques can be used to identify patterns in large volumes of media content which can subsequently be qualitatively analysed to identify gender bias.

The kinds of gender bias identified in the coverage of politicians in Irish newspapers align with previous studies of the representation of women in the media. This research also uncovered new ways in which such bias can be manifested in the content of newspaper coverage of politicians. Identification of these were facilitated by the corpus-driven approach adopted in this thesis.

Chapter 5

Conclusions

5.1 Summary of the Thesis

This research provides an original insight into gender bias in the coverage of politicians in Irish newspapers. Newspaper articles featuring Irish cabinet ministers over the 15 year period between 1997 and 2011 were analysed along with coverage of candidates in the 2011 Irish Presidential Election. The content of articles featuring male and female ministers was examined to identify evidence of gender bias against female politicians.

There were two key objectives of this research. The first objective was to create and test a new methodological approach to analysing gender bias in text. The purpose of this was to contribute to the existing range of methods available to researchers using Corpus Linguistics in analysing gender and language (Newbold et al., 2002; Scharkow, 2013; Wiedemann, 2013). The approach also addresses some criticisms of research in gender and language which predominantly relies on manual content analysis (Baker, 2014; Neuendorf, 2011).

The second objective was to examine whether there is bias in the coverage of politicians in Ireland and what the nature of it is. Ireland currently holds the 92nd position globally for the level of women's participation in politics (IPU, 2013), a figure which has not improved in recent years. Given that gender bias in the media has been shown to discourage women from entering politics (Fox and Lawless, 2004; Heldman et al., 2005), it is important to examine media coverage of Irish politics for evidence of gender bias. To date there has been little analysis of bias in the Irish media (Ahmad et al., 2011; Brandenburg, 2005) and this research has addressed this research gap.

The central focus of the study was on newspaper coverage of cabinet ministers in Ireland. To achieve this, a sample of ministers with comparable roles was selected. Two corpora of articles were created using different search criteria. One corpus contained articles where the cabinet ministers were a central focus of the article content. The second corpus was a broader-ranging corpus including all mentions of the ministers. A third corpus contained coverage of the candidates in the 2011 Irish Presidential Election. These corpora provided the opportunity to use different methods to explore the textual content and examine different aspects of political coverage in Ireland.

The corpora of newspaper articles were divided according to whether they featured male or female politicians. The content of the articles was then processed and key information was extracted, based on findings from previous research on the representation of women in the media (Mills, 2002; Trimble et al., 2013; Van Zoonen, 2006) and related studies on text classification in computational linguistics (Argamon et al., 2009; Diermeier et al., 2012; Koppel et al., 2009). A machine learning algorithm was used to generate models to predict the gender of the politician featured in the articles based on learned patterns differentiating between articles featuring male and female politicians. Since this research was the first study to utilise machine learning in order to detect gender bias, a broad range of methods to identify patterns in text were implemented. Throughout the research, these techniques were evaluated and the most useful ones identified.

The patterns identified by the machine learning algorithm which differentiated between articles featuring male or female politicians were analysed to identify whether they constituted gender bias or not. The context in which some differences occurred was explored by extracting occurrences from the texts using concordance analysis. The following sections summarise the findings of this research.

5.1.1 Detecting Gender Bias using Text Classification

These research findings indicate that machine learning can be effectively used to uncover patterns in how male and female politicians are represented in the media. A key focus in text classification is on identifying the best features to extract from texts for analysis related to the research goal (Whitelaw et al., 2005). To address this, a broad range of approaches to extracting information or features from the newspaper articles was explored. Multiple machine learning algorithms and methods of representing texts were also tested.

The advantages of both the corpus-driven and corpus-based approaches to identifying patterns in text were demonstrated by the findings gained from using different approaches to feature extraction. Using all of the words occurring in newspaper content as features for analysis by the machine learning algorithm identified unexpected patterns of difference in the texts featuring male and female politicians. Extracting features based on predefined lexicons, thereby using a more corpus-based approach, was valuable in its capacity to focus the analysis of the text on specific themes.

The following are the findings of the research pertaining to each research question:

Research Question 1: How can automatic text classification be used to explore differences in the coverage of male and female politicians in order to identify gender bias?

> **1(a):**What machine learning algorithm is suitable for identifying gender bias?
> The best results in this study were produced using a Support Vector Machine learning algorithm.
>
> **1(b):**Which is the optimal approach to representing features in text classification experiments to identify gender bias?
> The binary representation of features produced classification accuracies comparable to those gained using the tf-idf representation. However, the binary representation ranked the features suggesting gender bias as more important discriminative factors. Given this, it was concluded that the binary representation was the most useful in terms of the research goal.
>
> **1(c):** What approach to feature extraction is most informative in identifying gender bias in text?
> The following are the main approaches to feature extraction that were the must useful in terms of identifying gender bias in newspaper articles:
>
> - Single words
> - General Inquirer Lexicon of action words

- General Inquirer Lexicon of power words
- Adjectives
- Verbs
- Newspaper section

5.1.2 Gender Bias in Irish Newspapers

The findings showed that there was evidence of gender bias in the coverage of female politicians in Ireland. In previous studies of gender bias in the media, bias had been analysed in terms of the quantity and the content of the coverage (Gidengil and Everitt, 1999). The present research showed a distinct positive bias towards female politicians in terms of the volume of coverage they received. This was evident both in their appearances in headlines and the content of the articles. However, differences were evident in how male and female politicians were featured in articles which may be attributable to gender bias. The following sections set out the main kinds of gender bias that were identified, thereby addressing the second research question.

Research Question 2: What differences in Irish newspaper coverage of male and female politicians indicate gender bias?

Volume of Coverage

This research showed that female ministers received far greater volume of coverage in Irish newspapers than their male counterparts. There was also no evidence of gender bias in the quantity of coverage afforded to the candidates of the 2011 Irish Presidential election.

On average, female ministers received almost twice as much coverage as male ministers. Each of the female ministers studied were in the top 50th percentile in terms of the volume of coverage per year in office. The only cabinet post where female ministers received less coverage than male ministers was the Ministry for Agriculture. The sections in the newspapers where male ministers featured more often than female ministers were in the Sport and Motor sections. These sections are stereotypically male domains. However, female ministers also featured more strongly in the higher profile sections of the newspaper such as the Front Page and in the Business and Finance sections.

These findings contradict those from previous studies which have shown that female politicians receive less than, or equal coverage as their male counterparts (Bystrom et al., 2001; Fuertes-Olivera, 2007; Miller et al., 2010; Pearce, 2008). This may be due to the fact that female politicians are gender schema inconsistent (Bem, 1981) and therefore may attract more media attention. Verification of this requires further research however. Given that female ministers were more prominently featured in the Business and Finance sections, these results also question previous studies that showed male politicians being more closely associated with issues pertaining to business and finance (Brikse, 2004; Bystrom et al., 2004; Carroll and Schreiber, 1997).

Focus on Family Relationships and Roles

Marital status was more important in coverage of female politicians than male politicians as evidenced by the number of mentions of their marital status and comments on their spouse. For each year in office, female politicians' husbands were mentioned 4 times more often than men's wives. Similar trends were evident in the coverage of the Presidential candidates. The spouses of male politicians were also portrayed in stereotypical ways. Wives were frequently portrayed as being 'at home', supporting and helping their husbands. In contrast, the relationships of female ministers were portrayed as more equal partnerships.

In references to a female minister's children, where relevant, the conflict between the demands of political and family life were focused on, thereby emphasising the negative effects of politics on family life for women. Not not having children was also portrayed as an incompleteness in the life of female politicians and a necessary sacrifice for success in politics. However, while male ministers' desire to have more time with their family was mentioned, their family life was portrayed as as a happy one which they enjoyed.

These findings show that among women in political life, family was focused on more than their male counterparts. This supports the findings of Brikse (2004), Spears et al. (2000) and Van Zoonen (2006). While some newspaper coverage suggested that female politicians tend to emphasise their own role in their family, the media portrayal of their family life as conflicting with their political career suggests that this view did not originate with the politicians (Miller and Peake, 2013; Ross and Sreberny, 2000).

Policy Focus

Differences were evident in the kind of policy areas associated with male and female politicians in Irish newspapers. Female politicians tended to be associated with issues that might stereotypically be considered women's issues. This aligns with previous research that identified similar differences in the issues politicians were associated with, regardless of their formal role in politics (Carroll and Schreiber, 1997; Drew, 2000; Kahn, 1996; Kahn and Goldenberg, 1991;

Scharrer, 2002).

The issues female politicians were associated with in this corpus related to work-life-balance, childcare, gender equality and social policies relating to mothers. When male ministers were mentioned in conjunction with gender equality policy in the workforce or in politics, they were portrayed as under pressure to conform to the polices rather than as advocates for them. Even the language used to refer to families in a policy context differed for men and women. Male ministers were associated with discussions of families as an economic unit by using the term 'household' while female ministers were associated with families in terms of social policy.

Focus on Gender

The gender of female politicians was mentioned significantly more often than that of male politicians. Many of these references were found to be irrelevant to the issue being discussed. For each year in office a female politician's gender was mentioned 5 times more than for male politicians. Female politicians were also grouped and evaluated against other female politicians. This identification of female politicians as a distinct group that are not comparable to male politicians, reinforces perceptions of their participation in politics as outside the norm (Adcock, 2010).

Use of Stereotypes

Supporting the findings by White (2012), stereotypical portrayals of women were found in the coverage of both ministers and election candidates in this research. Instead of being portrayed as diplomatic or convincing, female candidates were portrayed as using sexualised identities as part of their negotiating strategies. Also, while some male politicians were portrayed as being part of a political culture centred on the consumption of alcohol, female ministers were not connected with this culture. Aligning with research by Pantti (2005), terms used to describe the emotional state of female politicians also suggested they were less in control of their emotions than male candidates.

Political Style

Female politicians were portrayed as more accepting of government policies than male politicians and were associated with more passive and compliant actions. While the cause of this was not clear, it did reflect Adcock's (2010) description of the portrayal of female politicians in the media as 'condensing symbols' for government policy.

The political style of female politicians in Ireland was compared with that of Margaret Thatcher. There were no comparisons, of male politicians with Margaret Thatcher. Female politicains were often described as 'formidable'. This, along with other instances of same-gender comparisons shows how female politicians are portrayed as having a distinct political style that is outside the male political norm. Their political style is also portrayed in more negative terms, compared with that of men.

Personalised Coverage

The trend towards greater personalisation of media coverage of politicians disproportionably affects female politicians (Van Zoonen, 2006). In this study, male politicians featured in articles focused on policy issues while articles featuring female politicians focused more on personal coverage. The personal attire of female ministers was also mentioned 18 times more often than for male ministers per year in office. Even the way male and female politicians were named differed, with more formal naming conventions being used to refer to male politicians.

Masculine Narrative/Metaphor

The use of metaphor in political discourse has been cited as alienating women from politics (Gidengil and Everitt, 1999; Trimble et al., 2013). While the use of war in political discourse in Irish politics was not prevalent, positions of power within the military and institutes of state were associated with male politicians. Discourse connected with sport was also primarily associated with male politicians. While the context of these discourses varied, the cumulative effect of systematic association of terms associated with power with male politicians suggests that there is an overall effect portraying female politicians as disassociated from these manifestations of power.

5.2 Contributions of the Research

5.2.1 Methodological Contribution

This research sets out how automatic text classification can be used to identify gender bias in the newspaper coverage of male and female politicians. This approach allowed a large volume of text to be identified and patterns of difference in the coverage of male and female politicians to emerge.

This study used text classification techniques within a feminist theoretical framework. In doing so, it contributes to research promoting the use of computational methods in the discipline of gender and language in particular (Baker, 2014) and in the social sciences generally (Wiedemann, 2013). It achieves this by testing a broad range of approaches to text classification and identifying the most useful in terms of the research goal. The main methodological decisions that need to be made in utilising a text classification approach to text analysis are highlighted, numerous options are explored and the most useful ones identified.

This study also presents how both a corpus-driven and corpus-based approach to analysing large samples of text can be implemented by altering the features extracted from the research. This allows the same overarching methodological approach to be used to explore hypothesis and allow patterns to emerge from the data which may not have been expected by the researcher.

5.2.2 Contribution to Theory

To date, there has been little research on bias in the Irish media, due in part to a belief that the Irish media is objective (Brandenburg, 2005). This research questions that assumption by uncovering systematic gender bias in how two of the best selling newspapers in the country portray female politicians.

A factor contributing to Ireland's poor record in women's political representation could be the level of gender bias in the media coverage of male and female politicians. This research outlines how this bias is manifested in the content of newspaper articles. In achieving this, the research explored ways to monitor media coverage for evidence of gender bias.

Many of the forms of gender bias identified in this research support findings from previous studies on how female politicians are represented in the media. New ways in which gender bias is manifested in newspaper articles are also uncovered, thus contributing to the body of research on how female politicians are portrayed in the media.

5.3 Limitations of the Research

A limitation of this research concerns the generalisability of the results. While gathering coverage of Irish ministers over a 15 year period and coverage of candidates of a Presidential election resulted in a large corpus of articles, the results cannot be reliably generalised to conclude that such gender bias exists in all coverage of female politicians. Similar to other research on the representation of women in politics, the scope of the study is bounded by the number of women in political life. This study was limited by the fact that there were only 5 female cabinet ministers in Ireland between 1997 and 2011. In the Presidential election, 2 of the 7 candidates were women. The distinctiveness of each of the candidates in this election made it difficult to identify patterns in the texts that could only be explained by the candidate's gender.

Creating datasets for studies in Corpus Linguistics can be an error-prone process. While much effort was devoted to ensuring the accuracy of every article in this research, the larger corpora were more difficult to verify reliably. The most significant source of errors was due to the fact that some of the searches returned articles featuring different people who shared the same name as the politicians in the sample. Strategies were put in place to identify these errors. However, the larger a corpus is, the more such errors may have remained.

The writing style of newspaper content can make applying natural language processing techniques error prone. This was seen in the number of words in the corpus that were tagged with incorrect parts-of-speech illustrating Stamatatos's (2009) point that the more pre-processing applied to text the more errors are accumulated. Due to this, an approach to minimising the pre-processing of newspaper content was adopted. However, using more complex natural language processing techniques in extracting features from the text to analyse in this research may have yielded more patterns in the data that this research did not identify.

The use of quantitative approaches to analysing text, such as text classification, seeks to reduce the levels of subjective judgement on part of the researcher, thereby increasing the scientific reliability of the results. However, as demonstrated by this thesis, there are several ways in which researcher bias can influence the findings of this research. Ascertaining whether discriminative features are attributable to gender bias relies on interpretation of those features. In research like this in particular, which is being carried out from a feminist standpoint, this interpretation could be open to researcher bias as other researchers with a different viewpoint may explain the same features differently. To address this every effort has been made to ensure the rationale and methodology behind the interpretation of data is as transparent as possible.

5.4 Future Research

This thesis presents numerous possibilities for further research. Some of these relate to the findings of this research and others emerge from its limitations.

In relation to studies of the coverage of female politicians in the media, the systematic approach to identifying bias undertaken in this research presents an opportunity to explore implementing real-time monitoring of newspaper coverage. Media coverage could be regularly analysed for evidence of similar patterns of language that were identified by this research. This work would integrate well with the research monitoring attributes of media coverage (Ahmad et al., 2011; Flaounas et al., 2011).

The kinds of gender bias identified in this study are wide-ranging. More qualitative studies could be used to explore some of the patterns in more depth to understand the origin of them fully. This Irish based study found that there was consistently more coverage of female politicians than male politicians. Since this has not been the case in any other study of the quantity of coverage of female politicians, further research is required to understand why this is the case. It is possible that it is a reflection of the work that was being done by female politicians in this sample or it may be a media bias in favour of female politicians specific to Ireland.

The findings of this thesis are specific to Ireland. It would be beneficial to replicate the study in other countries to examine whether the patterns relating to gender bias identified might be found outside Ireland.

At present, there are no language analysis tools specifically for research in the area of Gender and Language. This research highlights the potential benefits to be gained from the creation of corpus-based computational linguistics tools. The findings show that the use of lexicons to extract information about the newspaper articles yields targeted results. However the lexicons that were used were not created specifically to identify gender bias. The findings from studies such as this could be used to create lexicons which could be used as a resource for researchers in gender and language. Further research is also needed towards the creation of corpora designed for corpus-based studies in gender and language. This would accelerate the refinement of techniques for the analysis of topics in gender and language.

Bibliography

Adcock, C. (2010), 'The politician, the wife, the citizen, and her newspaper', *Feminist Media Studies* **10**(2), 135–159.

Ahmad, K., Daly, N. and Liston, V. (2011), What is new?: News media, general elections, sentiment, and named entities, *in* 'Proceedings of the Workshop on Sentiment Analysis: Where AI meets Psychology (SAAIP)', Chaing Mai, Thailand, pp. 80–88.

Ahrens, K. (2009), *Politics, Gender, and Conceptual Metaphors*, Palgrave Macmillan.

Akbani, R., Kwek, S. and Japkowicz, N. (2004), Applying support vector machines to imbalanced datasets, *in* J.-F. Boulicaut, F. Esposito, F. Giannotti and D. Pedreschi, eds, 'Machine Learning: ECML 2004', Vol. 3201, Springer Berlin Heidelberg, pp. 39–50.

Ali, O., Flaounas, I. N., De Bie, T., Mosdell, N., Lewis, J. and Cristianini, N. (2010), 'Automating news content analysis: An application to gender bias and readability', *Journal of Machine Learning Research - Proceedings Track* **11**, 36–43.

Argamon, S., Goulain, J., Horton, R. and Olsen, M. (2009), 'Vive la difference!:

Text mining gender difference in French literature', *Digital Humanities Quarterly* **3**(2).

Argamon, S., Koppel, M., Fine, J. and Shimoni, A. R. (2003), 'Gender, genre, and writing style in formal written texts', *Text* **23**(3), 321–346.

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

Argamon, S. and Levitan, S. (2005), Measuring the usefulness of function words for authorship attribution, *in* 'Proceedings of ACH/ALLC', Victoria, BC, Canada, June 2005.

Argamon, S., Whitelaw, C., Chase, P., Hota, S. R., Garg, N. and Levitan, S. (2007), 'Stylistic text classification using functional lexical features', *Journal of the American Society for Information Science and Technology* **58**(6), 802–822.

Baker, P. (2008), *Sexed texts: Language, gender and sexuality*, Equinox, London.

Baker, P. (2010), 'Will Ms ever be as frequent as Mr?: A corpus-based comparison of gendered terms across four diachronic corpora of british english', *Gender and Language* **4**(1), 125–149.

Baker, P. (2014), *Using Corpora to Analyze Gender*, Bloomsbury, London.

Banwart, M. C., Bystrom, D. G. and Robertson, T. (2003), 'From the primary to the general election: A comparative analysis of candidate media coverage in mixed-gender 2000 races for Governor and US Senate', *American Behavioral Scientist* **46**(5), 658–676.

Bem, S. L. (1981), 'Gender schema theory: A cognitive account of sex typing', *Psychological review* **88**(4), 354.

Bem, S. L. and Bem, D. J. (1973), Training the woman to know her place: The social antecedents of women in the world of work, Technical report, Pennsylvania State Department of Education, Bureau of Pupil Personnel Services.

Bird, S. (2006), Nltk: The natural language toolkit, *in* 'Proceedings of the COLING/ACL on Interactive presentation sessions', Association for Computational Linguistics, Stoudsburg, PA, USA, pp. 69–72.

Boulis, C. and Ostendorf, M. (2005), A quantitative analysis of lexical differences between genders in telephone conversations, *in* 'Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics', Vol. 43, Stroudsburg, PA, USA, p. 435.

Brandenburg, H. (2005), 'Political bias in the Irish media: A quantitative study of campaign coverage during the 2002 general election', *Irish Political Studies* **20**(3), 297–322.

Brikse, I. (2004), *The Role of the Mass Media in the (re)distribution of Power: A Comparative Report*, Mass media in the (re)distribution of power, European Union in terms of the Community Framework Strategy on Gender Equality (2001-2005), Riga, Latvia.

Bristor, J. M. and Fischer, E. (1993), 'Feminist thought: Implications for consumer research', *Journal of Consumer Research* **19**(March), 518–536.

Burns, S., Eberhardt, L. and Merolla, J. L. (2013), 'What is the difference between a hockey mom and a pit bull?: Presentations of Palin and gender stereotypes in the 2008 presidential election', *Political Research Quarterly* **66**(3), 687–701.

Burton, D. (1982), *Through glass darkly: through dark glasses*, Language and Literature: An Introductory Reader in Stylistics, Allen and Unwin, London.

Busse, J. A. and Clifton Green, T. (2002), 'Market efficiency in real time', *Journal of Financial Economics* **65**(3), 415–437.

Butler, J. (1990), *Gender trouble, 2nd Edition*, Routledge, London.

Bystrom, D., Brown, N. and Fiddelke, M. (2012), Barriers bent but not broken: Newspaper coverage of local and state elections, *in* R. Melody, ed., 'Women and Executive Office: Pathways and Performance', Lynne Rinner, Boulder, pp. 159–179.

Bystrom, D. G., Robertson, T. A. and Banwart, M. C. (2001), 'Framing the fight - an analysis of media coverage of female and male candidates in primary races for governor and US Senate in 2000', *American Behavioral Scientist* **44**(12), 1999–2013.

Bystrom, D. G., Robertson, T., Banwart, M. C. and Kaid, L. L. (2004), *Gender and candidate communication: Videostyle, webstyle, newstyle*, Routledge, New York.

Bystrom, D. and Hennings, V. M. (2013), Newspaper coverage of women running for the US Senate in 2012: evidence of an increasingly level playing field?, *in* C. L. Armstrong, ed., 'Media Disparity: A Gender Battleground', Lexington Books, Lanham.

Caldas-Coulthard, C. R. (1995), Man in the news: The misrepresentation of women speaking in news-as-narrative-discourse, *in* S. Mills, ed., 'Language and Gender: Interdisciplinary Perspectives', Longman, London, pp. 226–239.

Caldas-Coulthard, C. R. and Moon, R. (2010), 'Curvy, hunky, kinky: Using corpora as tools for critical analysis', *Discourse and Society* **21**(2), 99–133.

Cameron, D. (1992), *Feminism and linguistic theory*, Macmillan, London.

Caracelli, V. J. and Greene, J. C. (1993), 'Data analysis strategies for mixed-method evaluation designs', *Educational evaluation and policy analysis* **15**(2), 195–207.

Carroll, S. and Schreiber, R. (1997), Media coverage of women in the 103rd congress, *in* P. Norris, ed., 'Women, Media and Politics', Oxford University, New York.

Casey, M. and Kate, S. (1980), *The handbook of nonsexist writing*, Harper and Row, New York.

Chang, Y.-W. and Lin, C.-J. (2008), 'Feature ranking using linear svm', *Causation and Prediction Challenge Challenges in Machine Learning, Volume 2* p. 47.

Chiang, C.-F. and Knight, B. (2011), 'Media bias and influence: Evidence from newspaper endorsements', *The Review of Economic Studies* **78**(3), 795–820.

Childs, S. (2004), 'A feminised style of politics?: Women MPs in the House of Commons', *The British Journal of Politics and International Relations* **6**(1), 3–19.

Coates, J. (2003), *Men talk: Stories in the making of masculinities*, Blackwell Pub., Malden, MA.

Coates, J. (2004*a*), *Women, men and language: A sociolinguistic account of gender differences in language*, Longman, London.

Coates, J. (2004*b*), *Women talk: Conversation between women friends*, Blackwell, Oxford.

Corney, M., de Vel, O., Anderson, A. and Mohay, G. (2002), Gender-preferential text mining of e-mail discourse, *in* 'Computer Security Applications Conference, 2002. Proceedings. 18th Annual', IEEE, pp. 282–289.

Creswell, J. W. (2014), *Research design : qualitative, quantitative, and mixed methods approaches*, SAGE Publications, London.

Curran, J. (2002), *Media and power*, Routledge, New York.

D'Alessio, D. and Allen, M. (2000), 'Media bias in presidential elections: A meta-analysis', *Journal of Communication* **50**(4), 133–156.

Dave, K., Lawrence, S. and Pennock, D. M. (2003), Mining the peanut gallery: Opinion extraction and semantic classification of product reviews, *in* 'Proceedings of the 12th international conference on World Wide Web', ACM, pp. 519–528.

de Graaf, R. and van der Vossen, R. (2013), 'Bits versus brains in content analysis: Comparing the advantages and disadvantages of manual and automated methods for content analysis', *Communications* **38**(4), 433–443.

Denzin, N. K. (2009), *The research act: A theoretical introduction to sociological methods*, Transaction Publishers, New York.

Denzin, N. and Lincoln, Y. (1994), *Handbook of qualitative research*, Sage Publications.

Derrida, J. (1976), *Of grammatology*, Johns Hopkins University Press, Baltimore.

Devitt, J. (1999), *Framing gender on the campaign trail: Women's executive leadership and the press*, Women's Leadership Fund, New York.

Diermeier, D., Godbout, J.-F., Yu, B. and Kaufmann, S. (2012), 'Language and ideology in congress', *British Journal of Political Science* **42**(01), 31–55.

Drew, E. (2000), *Career Trajectories: Convergence or Divergence?*, Macmillan Press, London.

Durant, K. T. and Smith, M. D. (2007), Predicting the political sentiment of web log posts using supervised machine learning techniques coupled with feature selection, *in* 'Advances in Web Mining and Web Usage Analysis', Springer-Verlag, Berlin Heidelberg, pp. 187–206.

Eagly, A. H., Makhijani, M. G. and Klonsky, B. G. (1992), 'Gender and the evaluation of leaders: A meta-analysis', *Psychological Bulletin* **111**(1), 3.

Ehrlich, S. and King, R. (1994), 'Feminist meanings and the (de)politicization of the lexicon', *Language in Society* **23**(01), 59–76.

Flaounas, I., Ali, O., Turchi, M., Snowsill, T., Nicart, F., De Bie, T. and Cristianini, N. (2011), Noam: news outlets analysis and monitoring system, *in* 'Proceedings of the 2011 ACM SIGMOD international conference on Management of data', ACM, pp. 1275–1278.

Flaounas, I., Turchi, M., Ali, O., Fyson, N., De Bie, T., Mosdell, N., Lewis, J. and Cristianini, N. (2010), 'The structure of the EU mediasphere', *PLoS One* **5**(12), 14243.

Foucault, M. (1970), *The Archaeology of Knowledge*, Pantheon Books, New York.

Fowler, R., Hodge, R. I. V., Kress, G. and Trew, T. (1979), *Language and control*, Routledge, London.

Fox, R. L. (1997), *Gender dynamics in congressional elections*, Sage Publications, Thousand Oaks, USA.

Fox, R. L. and Lawless, J. L. (2004), 'Entering the arena?: Gender and the decision to run for office', *American Journal of Political Science* **48**(2), 264–280.

Friedan, B. (1963), *The feminine mystique*, Pengiun, London.

Frith, K., Shaw, P. and Cheng, H. (2005), 'The construction of beauty: A cross-cultural analysis of women's magazine advertising', *Journal of Communication* **55**(1), 56–70.

Fuertes-Olivera, P. A. (2007), 'A corpus-based view of lexical gender in written Business English', *English for Specific Purposes* **26**(2), 219–234.

Galligan, Y. (1998), *Women and politics in contemporary Ireland: From the margins to the mainstream*, Pinter, London.

Garcia-Blanco, I. and Wahl-Jorgensen, K. (2011), 'The discursive construction of women politicians in the European press', *Feminist Media Studies* **12**(3), 422–441.

Gauntlett, D. (2008), *Media, gender and identity: An introduction*, Routledge, London.

Gidengil, E. and Everitt, J. (1999), 'Metaphors and misrepresentation: Gendered mediation in news coverage of the 1993 canadian leaders debates.', *Harvard International Journal of Press/Politics* **4**(1), 48–65.

Gidengil, E. and Everitt, J. (2003), 'Talking tough: gender and reported speech in campaign news coverage', *Political Communication* **20**, 209–232.

Goatly, A. (2007), *Washing the Brain Metaphor and Hidden Ideology*, Vol. 23, John Benjamins Publishing, Amsterdam.

Grijzenhout, S., Marx, M. and Jijkoun, V. (2014), Sentiment analysis in parliamentary proceedings, *in* B. Kaal, I. Maks and A. van Elfrinkhof, eds, 'From Text to Political Positions: Text analysis across disciplines', Vol. 55, John Benjamins Publishing Company, Amsterdam, The Netherlands, pp. 117–134.

Grimmer, J. and Stewart, B. M. (2013), 'Text as data: The promise and pitfalls of automatic content analysis methods for political texts', *Political Analysis* **21**(3), 267–297.

Guyon, I., Weston, J., Barnhill, S. and Vapnik, V. (2002), 'Gene selection for cancer classification using support vector machines', *Machine learning* **46**(1-3), 389–422.

Halliday, M. A. K. (1973), *Explorations in the functions of language*, Elsevier North-Holland, New York.

Hansen, A., Cottle, S., Negrine, R., Newbold, C. and Halloran, J. D. (1998), *Mass communication research methods*, Macmillan, London.

Harrington, K., Litosseliti, L., Sauntson, H. and Sunderland, J. (2008), *Gender and language research methodologies*, Palgrave Macmillan, Basingstoke.

Heldman, C., Carroll, S. J. and Olson, S. (2005), 'She brought only a skirt: Print media coverage of Elizabeth Dole's bid for the republican presidential nomination', *Political Communication* **22**(3), 315–335.

Heyer, G., Quasthoff, U. and Wittig, T. (2006), *Text Mining: Wissensrohstoff Text*, W3L-Verlag, Herdecke.

Higgins, P. (2010), *A nation of politicians: gender, patriotism, and political culture in late eighteenth-century Ireland*, University of Wisconsin Press, Madison, Wisconson.

Hines, C. (1999), 'Rebaking the pie: The women as dessert metaphor', *Reinventing identities: The gendered self in discourse* pp. 145–162.

Hirst, G., Riabinin, Y., Graham, J. and Boizot-Roche, M. (2014), Text to ideology or text to party status?, *in* B. Kaal, I. Maks and A. van Elfrinkhof, eds, 'From Text to Political Positions: Text analysis across disciplines', Vol. 55, John Benjamins Publishing Company, Amsterdam, The Netherlands, pp. 93–116.

Holmes, G., Donkin, A. and Witten, I. H. (1994), Weka: A machine learning workbench, *in* 'Intelligent Information Systems, 1994. Proceedings of the 1994 Second Australian and New Zealand Conference on', IEEE, pp. 357–361.

Holmes, J. (1994), 'Inferring language change from computer corpora: some methodological problems', *ICAME journal* **18**, 27–40.

Holmes, J. (2003), *Gender Identity in New Zealand English*, Vol. 3, John Benjamins Publishing, Amsterdam.

Holmes, J. and Sigley, R. (2002), 'What's a word like girl doing in a place like this?: Occupational labels, sexist usages and corpus research', *Language and Computers* **36**, 247–264.

Hota, S., Argamon, S., Koppel, M. and Zigdon, I. (2006), 'Performing gender: Automatic stylistic analysis of shakespeare's characters', *Digital Humanities* pp. 82–88.

Hyde, J. S. (2005), 'The gender similarities hypothesis', *American psychologist* **60**(6), 581.

IPU (2013), *Parline Database*, Geneva, Switzerland. Available: http://www.ipu.org/parline/.

Jespersen, O. (1922), *Language, its nature, development, and origin*, G. Allen and Unwin, London.

Jiang, M. and Argamon, S. (2008), Political leaning categorization by exploring subjectivities in political blogs, *in* 'Proceedings of 4th International Conference on Data Mining (DMIN)', Citeseer, pp. 647–653.

Joint Committee on Justice, Equality, Defence and Women's Rights (2009), *Women's Participation in Politics*, Houses of the Oireachtas, Ireland.

Kahn, K. F. (1994), 'The distorted mirror: Press coverage of women candidates for statewide office.', *Journal of Politics* **56**(154-173).

Kahn, K. F. (1996), *The political consequences of being a woman: How stereotypes influence the conduct and consequences of political campaigns*, Power, conflict, and democracy: American politics into the Twenty-first century, Columbia University Press, Chichester, New York.

Kahn, K. F. and Goldenberg, E. N. (1991), 'Women candidates in the news: An examination of gender differences in us senate campaign coverage', *Public Opinion Quarterly* **55**(2), 180–199.

Kim, S.-M. and Hovy, E. (2006), Automatic identification of pro and con reasons in online reviews, *in* 'Proceedings of the COLING/ACL on Main conference poster sessions', Association for Computational Linguistics, pp. 483–490.

Koller, V. (2004), 'Businesswomen and war metaphors: 'Possessive, jealous and pugnacious'?', *Journal of sociolinguistics* **8**(1), 3–22.

Koppel, M., Akiva, N., Alshech, E. and Bar, K. (2009), Automatically classifying documents by ideological and organizational affiliation, *in* 'Intelligence and Security Informatics, 2009. ISI'09. IEEE International Conference on', IEEE, pp. 176–178.

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

Kramarae, C. (1981), *Women and men speaking: Frameworks for analysis*, Newbury House Publishers, Rowley, Massachusetts.

Kucukyilmaz, T., Cambazoglu, B. B., Aykanat, C. and Can, F. (2006), Chat mining for gender prediction, *in* 'Advances in Information Systems', Springer, Berlin Heidelberg, pp. 274–283.

Lakoff, G. and Johnson, M. (1980), 'The metaphorical structure of the human conceptual system', *Cognitive Science* **4**(2), 195–208.

Lakoff, R. T. (1975), *Language and woman's place*, Harper and Row, New York.

Lavery, L. (2013), 'Gender bias in the media?: An examination of local television news coverage of male and female house candidates', *Politics and Policy* **41**(6), 877–910.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabsi, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D. and Van Alstyne, M. (2009), 'Computational social science', *Science* **323**(5915), 721–723.

Lei, X. (2006), 'Sexism in language', *Journal of Language and Linguistics* **5**(1), 87–94.

Litosseliti, L. (2006), *Gender and language theory and practice*, Hodder Arnold, New York.

Lombard, M., SnyderDuch, J. and Bracken, C. C. (2002), 'Content analysis in mass communication: Assessment and reporting of intercoder reliability', *Human Communication Research* **28**(4), 587–604.

Luhn, H. P. (1958), 'The automatic creation of literature abstracts', *IBM Journal of research and development* **2**(2), 159–165.

Luyckx, K. and Daelemans, W. (2008), Authorship attribution and verification with many authors and limited data, *in* 'Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1', Association for Computational Linguistics, pp. 513–520.

Macnamara, J. (2003), 'Media content analysis: Its uses, benefits and best practice methodology', *Asia Pacific Public Relations* **6**(1), 1–34.

Manning, C., Prabhakar, R. and Hinrich, S. (2008), *Introduction to information retrieval*, Cambridge University Press, Cambridge.

Martin, J. R. and White, P. R. (2005), *The language of evaluation*, Palgrave Macmillan, Basingstoke.

Martyna, W. (1978), 'What does 'he' mean?', *Journal of Communication* **28**(1), 131–138.

Masser, B., Lee, K. and McKimmie, B. M. (2010), 'Bad woman, bad victim?: Disentangling the effects of victim stereotypicality, gender stereotypicality and benevolent sexism on acquaintance rape victim blame', *Sex Roles* **62**(7-8), 494–504.

McQuail, D. (1992), *Media Performance: Mass Communication in the Public Interest*, Sage, London.

Miller, M. K. and Peake, J. S. (2013), 'Press effects, public opinion, and gender coverage of Sarah Palin's vice-presidential campaign', *The International Journal of Press/Politics* **18**(4), 482–507.

Miller, M. K., Peake, J. S. and Boulton, B. A. (2010), 'Testing the Saturday night live hypothesis: Fairness and bias in newspaper coverage of Hillary Clinton's presidential campaign', *Politics and Gender* **6**(02), 169–198.

Millett, K. (1970), *Sexual politics*, Doubleday, New York.

Mills, S. (2002), *Feminist stylistics*, Routledge, London, New York.

Mollin, S. (2012), 'Revisiting binomial order in english: Ordering constraints and reversibility', *English Language and Linguistics* **16**(01), 81–103.

Motschenbacher, H. (2013), 'Gentlemen before ladies?: A corpus-based study of conjunct order in personal binomials', *Journal of English Linguistics* **41**(3), 212–242.

Moulton, J., Robinson, G. M. and Elias, C. (1978), 'Sex bias in language use:" neutral" pronouns that aren't', *American Psychologist* **33**(11), 1032.

Neuendorf, K. A. (2002), *The content analysis guidebook*, Sage Publications, Thousand Oaks, California.

Neuendorf, K. A. (2011), 'Content analysisa methodological primer for gender research', *Sex Roles* **64**(3-4), 276–289.

Neuman, L. W. (2007), *Social Research Methods, 6th Edition*, Pearson, Boston.

Newbold, C., Boydt-Barrett, O. and Bulck, V. d. (2002), *The media book*, Hodder Arnold Publication, London.

Niven, D. and Zilber, J. (2001), 'How does she have time for kids and congress?', *Women and Politics* **23**(1-2), 147–165.

Norris, P. (1997), *Women, media, and politics*, Oxford University Press, Oxford.

Nowson, S. and Oberlander, J. (2006), The identity of bloggers: Openness and gender in personal weblogs, *in* 'AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs', pp. 163–167.

Okimoto, T. G. and Brescoll, V. L. (2010), 'The price of power: Power seeking and backlash against female politicians', *Personality and Social Psychology Bulletin* **36**(7), 923–936.

Opsomer, R., Knoth, P., van Polen, F., Trapman, J. and Wiering, M. (2008), Categorizing children automated text classification of CHILDES files, *in* 'Proceedings 20th Belgian-Netherlands Conference on Artificial Intelligence', University of Twente, Enschede, the Netherlands, pp. 209–15.

Otterbacher, J. (2010), Inferring gender of movie reviewers: exploiting writing style, content and metadata, *in* 'Proceedings of the 19th ACM international conference on Information and knowledge management', ACM, pp. 369–378.

Page, R. (2010), 'New challenges for feminist stylistics: The case of girl with a one track mind', *Journal of Literary Theory* **4**(1), 81–97.

Page, R. E. (2003), "Cherie: lawyer, wife, mum': Contradictory patterns of representation in media reports of Cherie Booth/Blair', *Discourse and Society* **14**(5), 559–579.

Pang, B. and Lee, L. (2008), 'Opinion mining and sentiment analysis', *Foundations and trends in information retrieval* **2**(1-2), 1–135.

Pantti, M. (2005), 'Masculine tears, feminine tears and crocodile tears: Mourning Olof Palme and Anna Lindh in Finnish newspapers.', *Journalism Studies* **6**(3), 357–377.

Pauwels, A. (2000), Language in transition: A study of the title 'ms' in contemporary australian society, *in* A. Pauwels, ed., 'Women and Language in Australian and New Zealand Society', Australian Professional Publications, Sydney, pp. 129–54.

Pearce, M. (2008), 'Investigating the collocational behaviour of man and woman in the BNC using Sketch Engine', *Corpora* **3**(1), 1–29.

Pennebaker, J. W., Francis, M. E. and Booth, R. J. (2001), *Linguistic inquiry and word count: LIWC 2001*, Lawrence Erlbaum Associates, Mahwah, NJ.

Pichler, P. E. E. (2009), *Gender and spoken interaction*, Palgrave Macmillan, New York.

Pilaszy, I. (2005), Text categorization and support vector machines, *in* 'The Proceedings of the 6th International Symposium of Hungarian Researchers on Computational Intelligence'.

Pollach, I. (2012), 'Taming textual data: The contribution of corpus linguistics to computer-aided text analysis', *Organizational Research Methods* **15**(2), 263–287.

Pollak, S., Coesemans, R., Daelemans, W. and Lavra, N. (2013), 'Detecting contrast patterns in newspaper articles by combining discourse analysis and text mining', *Pragmatics* **21**(4).

Porter, M. F. (2001), 'Snowball: A language for stemming algorithms, available at: http://snowball.tartarus.org/texts/introduction'.

Prasad, P. (2005), *Crafting qualitative research: Working in the postpositivist traditions*, ME Sharpe, New York.

Price, N. D., Papin, J. A., Schilling, C. H. and Palsson, B. O. (2003), 'Genome-scale microbial in silico models: The constraints-based approach', *Trends in biotechnology* **21**(4), 162–169.

Radic-Bojanc, B. and Silaski, N. (2012), 'Metaphors of hegemonic masculinity: Women's understanding of sport metaphors in political discourse', *Gender Studies* **11**(1), 245–257.

Riffe, D. and Freitag, A. (1997), 'A content analysis of content analyses: Twenty-five years of journalism quarterly', *Journalism and Mass Communication Quarterly* **74**(3), 515–524.

Riloff, E., Patwardhan, S. and Wiebe, J. (2006), Feature subsumption for opinion analysis, *in* 'Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing', Association for Computational Linguistics, pp. 440–448.

Robinson, G. and Saint-Jean, A. (1991), Women politicians and their media coverage, *in* K. Megyery, ed., 'Women in Canadian politics: Toward equity in representation', Vol. 6, Dundurn Press, Toronto, pp. 127–145.

Romaine, S. (1999), *Communicating gender*, Lawrence Erlbaum Associates, New Jersey.

Romaine, S. (2000), *Language in Society: An Introduction to Sociolinguistics*, 2 edn, Oxford University Press, Oxford.

Romaine, S. (2001), *A corpus-based view of gender in British and American English*, Vol. 1, Benjamins, Amsterdam, pp. 153–175.

Ross, K. (2000), Women behaving badly? Framing gender in political communication, *in* 'Proceedings of Political Studies Association UK 50th Annual conference', Political Studies Association, London, pp. 10–13.

Ross, K., Evans, E., Harrison, L., Shears, M. and Wadia, K. (2013), 'The gender of news and news of gender a study of sex, politics, and press coverage of the 2010 British General Election', *The International Journal of Press/Politics* **18**(1), 3–20.

Ross, K. and Sreberny, A. (2000), *Women in the house: Media representation of British politicians*, Hampton Press, Cresskill, New Jersey, pp. 79–99.

Ross, K. and Sreberny-Mohammadi, A. (1997), 'Playing house: Gender, politics and the news media in Britain', *Media, Culture and Society* **19**(1), 101–109.

Ruping, S. (2005), Interpreting classifiers by multiple views, *in* '22nd ICML Workshop on Learning with Multiple Views', International Conference on Machine Learning.

Sabin, R. E., Goodwin, K. A., Goldstein-Stewart, J. and Pereira, J. A. (2008), Gender differences across correlated corpora: Preliminary results, *in* 'Proceedings of the Twenty-First International FLAIRS Conference', Miami, Florida, pp. 207–212.

Sanbonmatsu, K. (2006), 'Do parties know that women win?: Party leader beliefs about women's electoral chances', *Politics and Gender* **2**(04), 431–450.

Scharkow, M. (2013), 'Thematic content analysis using supervised machine learning: An empirical evaluation using German online news', *Quality and Quantity* **47**(2), 761–773.

Scharrer, E. (2002), 'An 'improbable leap': A content analysis of newspaper coverage of Hillary Clinton's transition from first lady to Senate candidate', *Journalism Studies* **3**(3).

Schiebinger, L., Klinge, I., Snchez de Madariaga, I., Schraudner, M. and Stefanick, M. (2011-2013), *Gendered Innovations in Science, Health and Medicine, Engineering and Environment*, Standford University, http://ec.europa.eu/research/gendered-innovations/.

Sebastiani, F. (2002), 'Machine learning in automated text categorization', *ACM computing surveys (CSUR)* **34**(1), 1–47.

Semino, E. (2008), *Metaphor in discourse*, Cambridge University Press Cambridge.

Sigley, R. and Holmes, J. (2002), 'Looking at girls in corpora of English', *Journal of English Linguistics* **30**(2), 138–157.

Somasundaran, S. and Wiebe, J. (2010), Recognizing stances in ideological on-line debates, *in* 'Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text', Association for Computational Linguistics, pp. 116–124.

Spears, G., Seydegart, K. and Gallagher, M. (2000), *Who makes the news*, Global Media Monitoring Project. Available at www.whomakesthenews.org.

Spender, D. (1980), *Man made language*, Pandora, London.

Sreberny-Mohammadi, A. and Ross, K. (1996), 'Women MPs and the media: Representing the body politic', *Parliamentary Affairs* **49**(1), 103–115.

Stamatatos, E. (2009), 'A survey of modern authorship attribution methods', *Journal of the American Society for Information Science and Technology* **60**(3), 538–556.

Stamatatos, E., Fakotakis, N. and Kokkinakis, G. (2001), 'Computer-based authorship attribution without lexical measures', *Computers and the Humanities* **35**(2), 193–214.

Stanton, E. C. (1990), *Declaration of sentiments*, M. Cole.

Steen, G. (1999), 'From linguistic to conceptual metaphor in five steps', *Amseterdam Studies in the Theory and History of Linguistic Science* (4), 57–78.

Stempel, G. H. and Stewart, R. K. (2000), 'The internet provides both opportunities and challenges for mass communication researchers', *Journalism and Mass Communication Quarterly* **77**(3), 541–548.

Stone, P. J., Dunphy, D. C. and Smith, M. S. (1966), *The General Inquirer: A Computer Approach to Content Analysis*, MIT Press, Mass.

Sunderland, J. and Litosseliti, L. (2002), *Gender identity and discourse analysis*, John Benjamins Publishing, Philadelphia, PA.

Swift, K. and Miller, C. (1981), *The Handbook of Non-Sexist Writing*, The Women's Press, London.

Taylor, C. (2013), 'Searching for similarity using corpus-assisted discourse studies', *Corpora* **8**(1), 81–113.

Tetlock, P. C. (2007), 'Giving content to investor sentiment: The role of media in the stock market', *The Journal of Finance* **62**(3), 1139–1168.

Tremblay, M. and Belanger, N. (1997), 'Femmes chefs de partis politiques et caricatures editoriales: Canadian federal elections 1993', *Recherches feministes* **10**(1), 35–75.

Trimble, L., Wagner, A., Sampert, S., Raphael, D. and Gerrits, B. (2013), 'Is it personal?: Gendered mediation in newspaper coverage of Canadian national party leadership contests 1975-2012', *The International Journal of Press/Politics* **18**(4), 462–481.

Trudgill, P. (1971), *The social differentiation of English in Norwich*, Cambridge University Press, Cambridge.

Uscinski, J. E. and Goren, L. J. (2011), 'What's in a name?: Coverage of Senator Hillary Clinton during the 2008 democratic primary', *Political Research Quarterly* **64**(4), 884–896.

Van Zoonen, L. (2006), 'The personal, the political and the popular: A woman's guide to celebrity politics', *European Journal of Cultural Studies* **9**(3), 287–301.

Vefali, G. M. and Erdentug, F. (2010), 'The coordinate structures in a corpus of new age talks: 'man and woman'/'woman and man'', *Text and Talk-An Interdisciplinary Journal of Language, Discourse and Communication Studies* **30**(4), 465–484.

Wardhaugh, R. (1998), *An Introduction to Sociolinguistics (Blackwell Textbooks in Linguistics)*, Blackwell, New York.

Weedon, C. (1997), *Feminist practice and poststructuralist theory*, 2nd edn, Blackwell, Oxford.

WEF (2013), *The Global Gender Gap Report 2013*, World Economic Forum.

White, S. (2012), 'Mothers and whores: The relationship between popular culture and women in politics', *Mapping Politics* **4**(september).

Whitelaw, C., Garg, N. and Argamon, S. (2005), Using appraisal groups for sentiment analysis, *in* 'Proceedings of the 14th ACM international conference on Information and knowledge management', ACM, pp. 625–631.

Wiedemann, G. (2013), 'Opening up to big data: Computer-assisted analysis of textual data in social sciences', *Forum Qualitative Social Research* **14**(2).

Wilson, T., Wiebe, J. and Hwa, R. (2004), Just how mad are you?: Finding strong and weak opinion clauses, *in* 'AAAI', Vol. 4, pp. 761–769.

Witten, I. H., Frank, E. and Hall, M. A. (2011), *Data Mining: Practical Machine Learning Tools and Techniques: Practical Machine Learning Tools and Techniques*, Elsevier Science, Burlington.

Woolf, V. (2008), *A Room of One's Own 1929*, Bridgewater, NJ.

Wu, H. C., Luk, R. W. P., Wong, K. F. and Kwok, K. L. (2008), 'Interpreting tf-idf term weights as making relevance decisions', *ACM Trans. Inf. Syst.* **26**(3), 1–37.

Yu, B., Kaufmann, S. and Diermeier, D. (2008), 'Classifying party affiliation from political speech', *Journal of Information Technology and Politics* **5**(1), 33–48.

Appendix A

Examples of Discriminative Features

Female	**Male**
invasion	conciliation
dependency	riot
non-proliferation	mediation
ratification	decree
negotiation	neutrality
dáil	defence
protocol	ban
parliament	adoption
	yoke
	subordination
	patriotism
	submission

Table A.1: Corpus Naming Ministers in Headlines - Top Features of the SVM Classifier (idft): General Inquirer Lexicon of Political Words as Features Extracted from Sentences Naming Politicians

Female	Male
...	...
credibly	june
anti-inflation	forth
duboin	bord
double	drive
pacific	farmers
...	...

Table A.2: Extract of Top Features of the SVM-BOW Classifier: Unigram Features, Minimum Frequency 1

Female	**Male**
embrace	revoke
misrepresent	participant
wash	inhibit
seal	shell
smear	recreation
shake	circulation
enterprise	tempt
trail	reactive
murder	rejection
aspire	catch
equip	sing
ate	scare
bet	gesture
curb	club
bore	convene
wrestle	evaluation
woo	twist
resurrect	loan
rebuff	escape
writer	burst
shout	unite
cheer	suppress
snap	dominate
bury	rope
expel	celebration
slight	transport
prevail	sail
hang	erect
evade	prosecute
uplift	fled
undid	civility
ridicule	refund
researcher	assignment
ration	chase
quicken	cleanse
quest	cooperation
punish	cultivate
obstruct	deceive
multiply	deception
misbehave	defy
isolate	disarmament
intercede	discard
infect	disperse
fetch	drill
feign	expend
excel	experiment
err	feast
engulf	frolic
discredit	hunter
decipher	hustle

Table A.3: Corpus Naming Ministers in Headlines - SVM Classifier (binary): General Inquirer Lexicon of Action Words as Features Extracted from from Sentences naming Politician

Female	Male
made	many
was	may
been	get
when	did
you	only
time	at
over	never
more	last
us	way
because	for
any	support
all	too
be	this
much	would
the	as
how	but
against	very
there	from
not	asked
after	said
do	so
before	has
should	will
which	like
by	during
me	on
their	also
no	some
am	and
if	into
in	per
up	with
what	they
were	where
now	is
my	have
can	are

Table A.4: Corpus Naming Ministers in Headlines - SVM Classifier (binary): Stop Words as Features

Female	**Male**
hse *(health service executive)*	transport
enterprise	rural
islands	marine
heritage	communications
employment	david
mcdowell	sport
pharmaceutical	farmers
managers	asti
trade	tourism
define	natural
first-name-last-name	for
leader	document
commissioners	passed
left	eastern
told	cork
stores	cutbacks
double	agricultural
padraig	energy
application	fuel
moral	firm
small	dr
amalley	lists
donnell	green
joyce	navan
assurance	strategies
family	willie
political party	m
stagg	smoking
analysis	institute
jobs	boards
sir	understand
patients	doesn't
signed	system
nightmare	own
childcare	community
holiday	notion
enright	threat
argued	children
hand	teachers
patient	backbenchers
subsidy	education
regan	planned
genuinely	meat
amounts	elections
fas *(Training and Employment Authority)*	breach
accident-and-emergency	we
industries	progress
features	complaint
finance	retain
deposits	stakeholders

Table A.5: Corpus Naming Ministers in Headlines - SVM Classifier (tf-idf min. freq. 2): Unigram Features

Female	**Male**
hse *(health service executive)*	transport
enterprise	marine
islands	rural
heritage	for
managers	asti
mcdowell	communications
sir	farmers
pharmaceutical	david
analysis	passed
define	sport
double	firm
associated	avenue
holiday	natural
stores	retain
arising	tourism
left	document
leader	attendance
telecom	lists
violence	smoking
genuinely	complaint
amounts	eastern
childcare	issued
statistics	institute
temporary	threat
application	cutbacks
stance	pr
features	agricultural
someone	charging
padraig	understand
find	cork
stagg	tullamore
disastrous	boards
enright	energy
size	backbenchers
family	approved
became	green
immediately	doesn
donnell	strategies
mater	elections
first-name-last-name	progress
bank	rights
timetable	breach
seriously	households
patients	meat
shop	own
commissioners	education
mergers	fuel
brussels	ned
amalley	pursue
regan	distress

Table A.6: Corpus Naming Ministers in Headlines - SVM Classifier (binary min. freq. 2): Unigram Features

Female	**Male**
10-point	Shannon-Heathrow
half-rate	ODonoghue
beet	non-judicial
Longford/Westmeath	horrendous
four-seat	renewable
income-tax	offensive
differential	defensive
happen	infected
privately-run	developmental
O'Donnell	packaged
programme	stark
bird	under-age
merger	waiting-list
told	metric
alarming	parental
uneasy	North-Central
10-day	idealistic
golden	safest
cut-throat	forth
unemployed	horrific
hidden	curious
cutting-edge	45-year-old
vice-chairman	pensionable
on-line	preventive
17-year-old	Russian
hesitant	pan-European
romantic	infectious
Later	right
salary	hard-hitting
tax	manageable
AT&T	consequential
Four-thousand	radioactive
Labour-Sinn	well-equipped
O'Malley	travelling
Re-appointed	three-man
attached	thosesurveyed
autumn	sweep
com-parison	strictest
cost-neutral	strange
cross-examining	smaller-scale
dribble	silver-medal
furnish	second-highest
half-thought-out	participant
level-crossing	paedophile
motor-tax	over-45s
non-authorised	one-bed
place	missing
press-ganged	lump-sum
ratified	litigation-based

Table A.7: Corpus Naming Ministers in Headlines - SVM Classifier (binary): Adjectives as Features

Female	Male
hse *(health service executive)*	transport
enterprise	rural
islands	marine
heritage	communications
employment	david
mcdowell	farmers
managers	sport
pharmaceutical	asti
define	tourism
trade	natural
commissioners	document
leader	for
left	eastern
first-name-last-name	passed
padraig	cork
told	cutbacks
stores	dr
amalley	fuel
double	firm
polparty	agricultural
moral	institute
application	willie
donnell	energy
small	navan
amounts	boards
childcare	strategies
joyce	lists
jobs	understand
stagg	m
signed	notion
family	avenue
assurance	smoking
nightmare	backbenchers
patients	green
sir	doesn't
regan	own
finance	route
enright	education
holiday	threat
industries	children
analysis	stakeholders
sake	system
patient	community
requirements	households
fas	teachers
genuinely	meat
argued	denial
female	breach
subsidy	complaint
accident-and-emergency	renewable

Table A.8: Corpus Naming Ministers in Headlines - SVM-Classifier (binary min. freq. 3): Unigram Features

Female	Male
get -0.1972	transport 0.1261
enterprise -0.1687	own 0.1179
reform -0.1591	action 0.1159
come -0.1531	hard 0.1143
look -0.1388	pass 0.1129
progress -0.1357	play 0.107
ask -0.1261	market 0.1064
fine -0.1243	development 0.1016
answer -0.1218	order 0.0962
back -0.1192	battle 0.0957
free -0.1125	risk 0.0929
election -0.1114	act 0.0897
leader -0.1093	circle 0.0873
issue -0.1016	change 0.0847
take -0.1016	consumption 0.0815
wait -0.1	attempt 0.081
give -0.0965	pleased 0.0797
stop -0.0938	total 0.0781
vote -0.0932	pay 0.0775
watch -0.0891	drive 0.0743

Table A.9: Corpus Naming Ministers in Headlines (Letters to the Editor) - Top Features of the SVM-Classifier (binary): General Inquirer Action Lexicon Features with Feature Weights

Female	Male
enterprise -0.5081	prevent 0.5834
motion -0.4467	transport 0.5631
offer -0.4462	initiative 0.4912
seek -0.4229	whip 0.4634
spokesman -0.4034	administration 0.4311
school -0.3858	widen 0.4186
attempt -0.3733	better 0.3945
begin -0.3688	carry 0.3929
signal -0.3632	post 0.3887
lead -0.3616	recreation 0.3801
administrative -0.356	action 0.3688
encourage -0.3503	task 0.3671
participate -0.3454	distribution 0.3562
proceed -0.341	plan 0.3544
contribution -0.3394	loan 0.3459
accommodate -0.3383	competitive 0.3416
project -0.3349	charge 0.3396
branch -0.3316	abuse 0.3376
erect -0.3241	effect 0.3234
alter -0.3111	extend 0.3157

Table A.10: Corpus Naming Ministers in Headlines (Front Page Articles) - Top Features of the SVM Classifier (binary): General Inquirer Action Lexicon Features with Feature Weights

Female	Male
new -0.1041	five 0.0863
spokesman -0.0998	tourism 0.0763
service -0.0923	e 0.0748
across -0.09	allow 0.07
enterprise -0.0898	campaign 0.0664
over -0.0811	month 0.0662
about -0.0704	while 0.0654
hse -0.0699	review 0.0654
irish -0.0658	people 0.0641
went -0.0627	wider 0.0635
number -0.0624	farmers 0.0613
tuesday -0.0618	asked 0.0597
nothing -0.0603	known 0.0571
either -0.0597	i 0.0567
changed -0.0592	dr 0.0562
situation -0.0592	plan 0.055
full -0.0575	aer 0.0534
which -0.0569	bring 0.0529
employment -0.0562	we 0.0527

Table A.11: Corpus Naming Ministers in Headlines (Front Page Articles) - Top Features of the SVM Classifier (binary): Unigram Features with Feature Weights

Female	Male
brazilian -2.5076	communications 2.5449
pg -2.5067	metro 2.405
thatcher -2.1864	mccole 2.3791
singer -2.1552	pork 2.376
hse -2.1512	fionnuala 2.3465
cunningham -2.1382	transport 2.3281
off-licence -2.1356	maire 2.2595
seagate -2.1173	bone 2.2201
disclosures -2.0415	amgen 2.1849
values -2.0359	shareholding 2.1704
suspicions -2.0321	grain 2.0913
haugh -2.0272	irfu 2.0777
awash -2.0	mayors 2.0624
bleats -2.0	handicap 2.042
captive -2.0	biofuels 2.039
cereals -2.0	sky 2.0262
fur -2.0	harvest 2.017
graham -2.0	cartel 2.0125
post-importation -2.0	ringsend 2.0
residence -2.0	residue 2.0
tips -2.0	nchds 2.0
wrapped -2.0	chequebook 2.0
subs -1.9935	doses 2.0
female -1.9879	korean 1.9894
florida -1.9803	smoke-free 1.9874
madden -1.9569	surcharges 1.9806
wrixon -1.9384	hauliers 1.9744
grehan -1.9381	macken 1.9601
calamity -1.9295	nuala 1.955
ida-backed -1.9246	reeves 1.9226
quibbles -1.9202	anti-smoking 1.9158
sexist -1.9161	tape 1.9117
pharmacists -1.9032	dilger 1.9115
cox -1.8875	river 1.9071
beware -1.8872	tullamore 1.8963
marys -1.8858	claymon 1.8909
fas -1.8787	inadvertently 1.8835
hangar -1.8678	aspersions 1.8811
twomey -1.8676	famine 1.8808
newcastle -1.8635	presentations 1.8713
rating -1.8553	hepatitis 1.8657
god -1.8548	tags 1.8615
demotion -1.8539	worsen 1.8601
feeney -1.8379	energy 1.8596
auschwitz -1.8303	esb 1.8564
swore -1.8231	lca 1.8361
grealish -1.8164	anaesthetic 1.8353
accountancy -1.816	cavan/monaghan 1.8334
philosophy -1.8076	new-born 1.8324
mid -1.7987	ocean 1.8321
...	...
mothers -1.7316	incorporate 1.7398
...	...
healthcare -1.4847	last-mentioned 1.4821

Table A.12: Corpus Naming Ministers in Sentences - Top Features of the Svm-Bool Classifier: Unigram Features with feature weights

Female	Male
tail -1.3985	shell 2.9701
paint -1.3357	transport 2.4071
modify -1.2192	harvest 2.183
singer -1.1614	recreation 2.1587
bury -1.161	install 2.0339
experiment -1.161	deceive 2.0
rate -1.161	overrun 2.0
attach -1.1609	revoke 1.9983
compare -1.1608	squeeze 1.9142
deliberate -1.1608	output 1.8414
flew -1.1608	wrestle 1.7444
recruit -1.1608	tennis 1.728
differ -1.1607	shatter 1.7151
wage -1.1607	contend 1.706
mislead -1.1606	installation 1.704
struggle -1.1606	rejection 1.6621
peel -1.1605	concentration 1.6474
sung -1.1605	devastate 1.6313
care -1.1604	penetration 1.6226
dress -1.1604	swim 1.6062
punish -1.1603	sail 1.6007
haste -1.1602	loan 1.5968
acquire -1.0702	retreat 1.5888
expression -1.0217	wind 1.5585
depart -1.0214	spear 1.5447
salute -1.0037	exploration 1.5308
alarm -1.0	shoulder 1.5048
attendant -1.0	designate 1.4936
bestow -1.0	ridicule 1.4921
compliment -1.0	mediate 1.467
conspire -1.0	roof 1.4468
continuous -1.0	coach 1.4379
crept -1.0	dip 1.4074
defy -1.0	heat 1.3904
descry -1.0	amend 1.3897
dominate -1.0	festival 1.3782
drank -1.0	construct 1.349
embarrass -1.0	collusion 1.3446
embrace -1.0	detection 1.3384
engulf -1.0	strike 1.3365
expel -1.0	ascertain 1.3164
fathom -1.0	lure 1.3126
flash -1.0	rip 1.3125
grapple -1.0	tease 1.3071
herd -1.0	limit 1.3055
hop -1.0	applaud 1.2998
imitation -1.0	driver 1.2662
induce -1.0	frame 1.254
intercede -1.0	creativity 1.2495
kiss -1.0	assassinate 1.2488

Table A.13: Corpus Naming Ministers in Sentences - Top Features of the SVM Classifier (binary): General Inquirer Lexicon of Action Words as Features with Feature Weights

Female	Male
scored -2.1648	control 2.6831
ag -2.0003	aggravated 2.4906
angry -2.0	designate 2.0619
ball -2.0	congested 2.0371
count -2.0	swimming 2.0256
fas -2.0	ticketing 2.0
holidayed -2.0	speeding 2.0
launches -2.0	routing 2.0
muttered -2.0	pack 2.0
names -2.0	monopolise 2.0
plays -2.0	justdo 2.0
posting -2.0	intimated 2.0
racked -2.0	culture 2.0
removes -2.0	boarding 2.0
suspends -2.0	angling 2.0
travels -2.0	evolve 2.0
visualise -2.0	whereby 1.988
capt -1.9999	halted 1.9658
defends -1.9999	acceptance 1.9535
matron -1.9999	groomed 1.9511
single -1.9999	breakdown 1.9446
distract -1.9998	uninsured 1.938
newest -1.9998	plucked 1.9308
salute -1.9998	unused 1.9084
unravel -1.9998	decline 1.9028
remarks -1.9997	dislikes 1.8751
survives -1.9997	attained 1.8694
warns -1.9996	undertaking 1.8459
smile -1.9982	specialise 1.8387
plus -1.9906	driving 1.8181
sunbed -1.9806	insulting 1.8129
trim -1.9611	imported 1.8026
authorised -1.9609	shepherded 1.7934
purported -1.9548	forcing 1.783
merit -1.9308	shareholding 1.7755
observe -1.929	owe 1.7492
declares -1.9215	welded 1.7438
deregulate -1.9093	reconsidered 1.7435
suicide -1.8979	dredging 1.7351
sterling -1.8945	authorising 1.7333
opaque -1.8922	cycling 1.7311
shock -1.8855	checked 1.7264
biased -1.8846	exported 1.7214
weekdays -1.8813	native 1.7074
admired -1.8597	tracked 1.6934
rid -1.859	undoubted 1.6907
withhold -1.8359	sporting 1.6848
introduces -1.8294	interview 1.6575
retains -1.8281	matches 1.6533
...	...
married -1.6308	elevated 1.3126

Table A.14: Corpus Naming Ministers in Sentences - Top Features of the SVM Classifier (binary): Verbs as Features with Feature Weights

Female	Male
abused -1.1356	promoted 1.0
abstained -1.0	discuss 1.0
grant -1.0	criticised 0.8641
mckeown -1.0	killed 0.6453
performed -1.0	predicted 0.6443
molested -0.9964	become 0.6443
see -0.9727	blake 0.6287
informed -0.8382	invited 0.5209
riven -0.8262	warned 0.5134
siding -0.8262	represent 0.4878
tweed -0.8086	convicted 0.4878
clean -0.7191	attacking 0.4878
tells -0.7191	joined 0.4806
served -0.7163	says 0.4692
epitomised -0.6853	maybe 0.4513
ferry -0.6853	win 0.4162
born -0.6046	increase 0.4162
getting -0.5539	held 0.3926
losing -0.5539	are 0.392
appointed -0.5042	am 0.3846
sided -0.4892	appeared 0.3792
supported -0.4892	don't 0.3753
finish -0.4785	went 0.3744
kill -0.4785	gone 0.3579
preceded -0.4785	d 0.3579
print -0.4785	seeking 0.3541
someone -0.4785	entitled 0.3481
saved -0.473	hailing 0.3413
shouldn't't -0.473	flock 0.3413
nothing -0.4699	charlie 0.3413
secure -0.4699	noto 0.3335
told -0.4642	answer 0.3335
denying -0.4605	take 0.3299
sued -0.4605	having 0.3171
forced -0.445	title 0.3153
manhattan -0.4417	tried 0.3144
traded -0.4417	organising 0.3144
proposed -0.4342	being 0.3115
scrapped -0.4342	became 0.3106
burgh -0.4322	mairead 0.31
encountered -0.4322	standing 0.3085
working -0.4322	published 0.2953
got -0.4299	instructed 0.2953
had -0.4298	accepted 0.2897
came -0.4193	s 0.2894
said -0.3975	notwithstanding 0.2837
creating -0.3948	approached 0.2837
did -0.3596	is 0.2816

Table A.15: Corpus Naming Presidential Candidates in Headlines - Top Features of the SVM Classifier (binary): Verbs as Features with Feature Weights

Female	Male
upset -0.706	rally 1.1234
conflict -0.7028	account 0.9149
evidence -0.6718	minister 0.8662
control -0.649	performance 0.7763
court -0.6407	loan 0.7331
launch -0.6326	shot 0.7118
seat -0.5851	box 0.705
decision -0.551	research 0.6137
abuse -0.5491	invite 0.6097
bring -0.5396	peace 0.5525
interview -0.5292	face 0.5272
renounce -0.5284	transfer 0.5158
mean -0.5234	return 0.5105
still -0.4662	delivery 0.506
commercial -0.4523	exact 0.504
pull -0.4418	answer 0.504
state -0.4307	highlight 0.4937
singer -0.4214	speak 0.4873
row -0.3995	help 0.4805
achieve -0.3898	raise 0.4693
motion -0.3764	fine 0.4327
mistaken -0.3734	block 0.4272
busy -0.3667	spokesman 0.4183
campaigner -0.3587	ambush 0.4052
won -0.3548	association 0.3978
sign -0.3544	active 0.3925
trail -0.3483	expand 0.3915
inspire -0.3446	meet 0.3892
wage -0.3384	change 0.3864
race -0.3338	write 0.3845
destroy -0.3336	jump 0.3765
rush -0.3295	tell 0.3712
open -0.3293	institute 0.3682
broke -0.3288	own 0.3671
plan -0.3251	represent 0.3669
learn -0.3185	election 0.3634
leave -0.3128	come 0.3608
hurt -0.3116	leader 0.3513
festival -0.3058	pressure 0.35
construct -0.3032	clean 0.3464
storm -0.2989	draw 0.3406
join -0.2977	ground 0.3378
publish -0.2974	struggle 0.3336
make -0.2915	close 0.3287
interpretation -0.2863	win 0.3187
give -0.2849	detective 0.3186
exercise -0.2741	enterprise 0.316
break -0.2732	selection 0.3031
occasion -0.2636	meant 0.3019
schedule -0.2588	land 0.3008

Table A.16: Corpus Naming Presidential Candidates in Headlines - Top Features of the SVM Classifier (binary): General Inquirer Lexicon of Action Words as Features with Feature Weights

Female	**Male**
enterprise	transport
hse	communications
employment	recreation
eilish	david
leader	marine
islands	dr
double	for
government	education
heritage	green
violence	asti
sir	farmers
drawing	fees
unfortunately	agricultural
companies	approached
warning	avenue
trade	using
u-turn	system
imagine	assured
opens	bord
greedy	cutbacks
frontbench	tourism
rebecca	forth
sake	wondering
foster	natural
locomotive	complaint
athlone	resumption
begins	road
developments	drive
levels	uses
first-name-last-name	deteriorating
mortality	rural
facilities	agriculture
insurance	limit
mcdowell	consensus
aba	attendance
cooper	protection
husband	energy
argued	ned
half-rate	cover
things	decline
mater	west
pressing	drastic
freedom	theatres
be	pre-empt
telecom	beef
amalley	tight
dismay	cairo
small	june
polparty	via
modelled	passed

Table A.17: Corpus Naming Ministers in Headlines - Top Features of the SVM Classifier (binary): Unigram Features min frequency 10

Appendix B

Sentiment Lexicon

Word	Polarity	Intensity						
no aberrant	positive	median	not aberrant	positive	median	ultimately	maximize	
no abhorrent	positive	median	not abhorrent	positive	median	purely	positive	median
no able	negative	median	not able	negative	median	purer	positive	high
no abnormal	positive	median	not abnormal	positive	median	pure	positive	median
no abnormally	positive	median	not abnormally	positive	median	TRUE	positive	median
no abominable	positive	median	not abominable	positive	median	genuine	positive	high
no absolute	decrease	not absolute	decrease	genuinely	positive	high		
no absolutely	minimize	not absolutely	minimize	really	decrease			
no absorbed	negative	high	not absorbed	negative	high	absolutely	maximize	
no absorbing	negative	median	not absorbing	negative	median	completely	maximize	
no absurd	positive	high	not absurd	positive	high	immensely	maximize	
no absurdly	positive	high	not absurdly	positive	high	incredibly	maximize	
no acceptable	negative	low	not acceptable	negative	low	hugely	maximize	
no accommodating	negative	median	not accommodating	negative	median	incomparably	maximize	
no accomplished	negative	median	not accomplished	negative	median	perfectly	maximize	
no accurate	negative	median	not accurate	negative	median	prodigiously	maximize	
no acutely	decrease	not acutely	decrease	profoundly	maximize			
no adaptable	negative	median	not adaptable	negative	median	supremely	maximize	
no adept	negative	very high	not adept	negative	very high	surpassingly	maximize	
no adequate	negative	low	not adequate	negative	low	terrifically	maximize	
no admirable	negative	median	not admirable	negative	median	tremendously	maximize	
no adoring	negative	high	not adoring	negative	high	utterly	maximize	
no adroit	negative	high	not adroit	negative	high	wholly	maximize	
no adult	negative	median	not adult	negative	median	terribly	maximize	
no adventurous	negative	median	not adventurous	negative	median	unmistakably	maximize	
no affectionate	negative	median	not affectionate	negative	median	unbelievably	maximize	
no affectionately	negative	median	not affectionately	negative	median	staggeringly	maximize	
no afraid	positive	median	not afraid	positive	median	such	increase	
no aggressive	positive	median	not aggressive	positive	median	rather	increase	
no aghast	positive	very high	not aghast	positive	very high	very	increase	
no agile	negative	median	not agile	negative	median	greatly	increase	
no agitated	positive	median	not agitated	positive	median	extremely	increase	
no alien	positive	median	not alien	positive	median	amazingly	increase	
no allright	negative	median	not allright	negative	median	deliriously	increase	
no alluring	negative	high	not alluring	negative	high	more	increase	
no also-ran	positive	median	not also-ran	positive	median	quite	increase	
no altogether	decrease	not altogether	decrease	far	increase			
no altruistic	negative	median	not altruistic	negative	median	mostly	increase	
no amazing	negative	median	not amazing	negative	median	definitely	increase	
no amazingly	decrease	not amazingly	decrease	surprisingly	increase			
no ambitious	negative	median	not ambitious	negative	median	entirely	increase	
no amoral	positive	median	not amoral	positive	median	highly	increase	
no amorphous	positive	median	not amorphous	positive	median	fully	increase	
no amusing	negative	median	not amusing	negative	median	largely	increase	
no amusingly	negative	median	not amusingly	negative	median	entertained	positive	median
no angry	positive	median	not angry	positive	median	thoroughly	increase	
no animated	negative	median	not animated	negative	median	remarkably	increase	
no annoying	positive	high	not annoying	positive	high	strongly	increase	
no annoyingly	positive	high	not annoyingly	positive	high	notably	increase	
no anxious	positive	median	not anxious	positive	median	altogether	increase	
no anxiously	positive	median	not anxiously	positive	median	exceedingly	increase	
no appalling	positive	median	not appalling	positive	median	totally	increase	

no appealing	negative	median	not appealing	negative	median	strikingly	positive	median
no appreciable	negative	low	not appreciable	negative	low	exceptionally	increase	
no apprehensive	positive	median	not apprehensive	positive	median	specially	positive	median
no appropriate	negative	median	not appropriate	negative	median	positively	increase	
no appropriately	negative	median	not appropriately	negative	median	considerably	increase	
no arbitrary	positive	median	not arbitrary	positive	median	vastly	increase	
no arcane	positive	median	not arcane	positive	median	unusually	increase	
no arousing	negative	median	not arousing	negative	median	peculiarly	increase	
no arresting	negative	median	not arresting	negative	median	peculiarly	negative	median
no arrestive	negative	median	not arrestive	negative	median	intensely	increase	
no arrogant	positive	median	not arrogant	positive	median	extraordinarily	increase	
no arrogantly	positive	median	not arrogantly	positive	median	excessively	increase	
no articulate	negative	median	not articulate	negative	median	emphatically	increase	
no artificial	positive	median	not artificial	positive	median	comprehensively	increase	
no artistic	negative	median	not artistic	negative	median	significantly	increase	
no artistically	negative	median	not artistically	negative	median	pretty	increase	
no ascetic	positive	median	not ascetic	positive	median	pretty	positive	median
no assiduous	negative	median	not assiduous	negative	median	mightily	increase	
no assured	negative	median	not assured	negative	median	inordinately	increase	
no assuredly	negative	median	not assuredly	negative	median	astonishingly	increase	
no astonished	positive	median	not astonished	positive	median	substantially	increase	
no astonishing	negative	median	not astonishing	negative	median	reasonably	increase	
no astonishingly	decrease	not astonishingly	decrease	moderately	decrease			
no astounding	negative	high	not astounding	negative	high	enormously	increase	
no astute	negative	median	not astute	negative	median	acutely	increase	
no asymmetric	positive	median	not asymmetric	positive	median	somewhat	decrease	
no asymmetrical	positive	median	not asymmetrical	positive	median	relatively	decrease	
no atrocious	positive	high	not atrocious	positive	high	fairly	decrease	
no atrociously	positive	high	not atrociously	positive	high	slightly	decrease	
no attentive	negative	median	not attentive	negative	median	hardly	decrease	
no attentively	negative	median	not attentively	negative	median	barely	decrease	
no attractive	negative	median	not attractive	negative	median	mildly	decrease	
no attractively	negative	median	not attractively	negative	median	partly	decrease	
no auspicious	negative	median	not auspicious	negative	median	passably	decrease	
no authentic	negative	median	not authentic	negative	median	scarcely	decrease	
no avaricious	positive	median	not avaricious	positive	median	tolerably	decrease	
no average	positive	median	not average	positive	median	partially	decrease	
no awesome	negative	high	not awesome	negative	high	imperceptibly	decrease	
no awful	positive	high	not awful	positive	high	accommodating	positive	median
no awfully	positive	high	not awfully	positive	high	adaptable	positive	median
no awkward	positive	median	not awkward	positive	median	brave	positive	median
no awkwardly	positive	median	not awkwardly	positive	median	braver	positive	high
no awry	positive	median	not awry	positive	median	careful	positive	median
no backward	positive	median	not backward	positive	median	cautious	positive	median
no backwardly	positive	median	not backwardly	positive	median	circumspect	positive	median
no bad	positive	median	not bad	positive	median	conscientious	positive	median
no badly	positive	median	not badly	positive	median	considerate	positive	median
no balanced	negative	median	not balanced	negative	median	constant	positive	median
no banal	positive	median	not banal	positive	median	courageous	positive	high
no barbarous	positive	median	not barbarous	positive	median	dauntless	positive	high
no barely	increase	not barely	increase	dependable	positive	median		
no barren	positive	median	not barren	positive	median	devoted	positive	median
no base	positive	median	not base	positive	median	exact	positive	median
no baser	positive	high	not baser	positive	high	faithful	positive	median
no basic	negative	median	not basic	negative	median	fearless	positive	high
no beastly	positive	median	not beastly	positive	median	firm	positive	median
no beauteous	negative	median	not beauteous	negative	median	flexible	positive	median
no beautiful	negative	median	not beautiful	negative	median	flexible	positive	median
no beautifully	decrease	not beautifully	decrease	gallant	positive	high		
no beggarly	positive	median	not beggarly	positive	median	gutsy	positive	high
no beguiling	positive	median	not beguiling	positive	median	heroic	positive	high
no believable	negative	median	not believable	negative	median	honest	positive	median
no beloved	negative	median	not beloved	negative	median	honorable	positive	median
no beneficial	negative	median	not beneficial	negative	median	indefatigable	positive	median
no benevolent	negative	median	not benevolent	negative	median	loyal	positive	median
no best	negative	max	not best	negative	max	meticulous	positive	median
no better	negative	high	not better	negative	high	patient	positive	median
no bewitching	negative	high	not bewitching	negative	high	persevering	positive	median
no big	negative	median	not big	negative	median	plucky	positive	median
no bitter	positive	median	not bitter	positive	median	reliable	positive	median
no bizarre	positive	high	not bizarre	positive	high	reliable	positive	median
no bizarrely	positive	high	not bizarrely	positive	high	resolute	positive	high
no blabbermouth	positive	median	not blabbermouth	positive	median	safe	positive	median
no blah	positive	median	not blah	positive	median	scrupulous	positive	median
no blameless	negative	max	not blameless	negative	max	scrupulous	positive	median
no bland	positive	median	not bland	positive	median	stalwart	positive	median
no blandly	positive	median	not blandly	positive	median	staunch	positive	high
no blatant	positive	median	not blatant	positive	median	steadfast	positive	high
no bleak	positive	median	not bleak	positive	median	steady	positive	median
no blithe	negative	median	not blithe	negative	median	thorough	positive	median
no blue	positive	high	not blue	positive	high	tireless	positive	median
no blunt	positive	median	not blunt	positive	median	trustworthy	positive	median
no bluntly	positive	median	not bluntly	positive	median	trusty	positive	median
no bogus	positive	median	not bogus	positive	median	unafraid	positive	median
no bold	negative	median	not bold	negative	median	undaunted	positive	median
no bonafide	negative	median	not bonafide	negative	median	unfailing	positive	median
no borderline	positive	low	not borderline	positive	low	unflagging	positive	high
no bored	positive	median	not bored	positive	median	unswerving	positive	high
no boring	positive	median	not boring	positive	median	untiring	positive	median
no bouncy	negative	high	not bouncy	negative	high	unvarying	positive	median
no brainless	positive	median	not brainless	positive	median	unwavering	positive	high
no brave	negative	median	not brave	negative	median	upright	positive	median
no braver	negative	high	not braver	negative	high	valiant	positive	median
no bravest	negative	max	not bravest	negative	max	veracious	positive	median
no breathtaking	negative	high	not breathtaking	negative	high	vigilant	positive	median

no breathtakingly	negative	high	not breathtakingly	negative	high	wary	positive	median
no brief	negative	median	not brief	negative	median	watchful	positive	median
no brilliant	negative	high	not brilliant	negative	high	capricious	negative	median
no brilliantly	negative	high	not brilliantly	negative	high	cowardly	negative	high
no brutal	positive	high	not brutal	positive	high	despondent	negative	median
no brutally	positive	high	not brutally	positive	high	disloyal	negative	median
no bulkier	positive	median	not bulkier	positive	median	distracted	negative	low
no bulky	positive	median	not bulky	positive	median	erratic	negative	median
no buoyant	negative	median	not buoyant	negative	median	FALSE	negative	median
no busy	negative	median	not busy	negative	median	fickle	negative	median
no byzantine	positive	high	not byzantine	positive	high	foolhardy	negative	high
no callous	positive	median	not callous	positive	median	gutless	negative	median
no callow	positive	median	not callow	positive	median	hasty	negative	median
no calm	negative	median	not calm	negative	median	headstrong	negative	median
no candid	negative	median	not candid	negative	median	impatient	negative	median
no candidly	negative	median	not candidly	negative	median	impetuous	negative	high
no cantankerous	positive	median	not cantankerous	positive	median	impulsive	negative	median
no capable	negative	median	not capable	negative	median	inconstant	negative	median
no capricious	positive	median	not capricious	positive	median	irresolute	negative	median
no captivated	negative	high	not captivated	negative	high	madcap	negative	median
no captivating	negative	high	not captivating	negative	high	obstinate	negative	median
no careful	negative	median	not careful	negative	median	rash	negative	median
no carefully	negative	median	not carefully	negative	median	reckless	negative	high
no careless	positive	median	not careless	positive	median	shaky	negative	median
no caring	negative	median	not caring	negative	median	soft	negative	median
no catchy	negative	median	not catchy	negative	median	spineless	negative	high
no cautious	negative	median	not cautious	negative	median	stubborn	negative	median
no cautiously	negative	median	not cautiously	negative	median	timid	negative	median
no celebrated	negative	median	not celebrated	negative	median	traitorous	negative	median
no central	negative	median	not central	negative	median	treacherous	negative	high
no challenging	negative	median	not challenging	negative	median	undependable	negative	median
no chaotic	positive	median	not chaotic	positive	median	unfaithful	negative	median
no charitable	negative	median	not charitable	negative	median	unloyal	negative	median
no charmed	negative	median	not charmed	negative	median	unpredictable	negative	median
no charming	negative	median	not charming	negative	median	unreliable	negative	median
no charmingly	negative	median	not charmingly	negative	median	unsure	negative	median
no chaste	negative	median	not chaste	negative	median	untrustworthy	negative	median
no cheap	negative	median	not cheap	negative	median	vacillating	negative	median
no cheaply	negative	median	not cheaply	negative	median	variable	negative	median
no cheerful	negative	median	not cheerful	negative	median	wayward	negative	median
no cheerfully	negative	median	not cheerfully	negative	median	weak	negative	median
no cheerily	negative	median	not cheerily	negative	median	wilful	negative	median
no cheerless	positive	max	not cheerless	positive	max	willful	negative	median
no cheery	negative	median	not cheery	negative	median	firmer	positive	high
no childish	positive	median	not childish	positive	median	safer	positive	high
no childishly	positive	median	not childishly	positive	median	truer	positive	high
no chirpy	negative	high	not chirpy	negative	high	softer	negative	high
no chuffed	negative	median	not chuffed	negative	median	weaker	negative	high
no cinematic	negative	median	not cinematic	negative	median	bravest	positive	max
no circumspect	negative	median	not circumspect	negative	median	safest	positive	max
no classic	negative	median	not classic	negative	median	truest	positive	max
no classical	negative	median	not classical	negative	median	weakest	negative	max
no classically	negative	median	not classically	negative	median	carefully	positive	median
no clean	negative	median	not clean	negative	median	cautiously	positive	median
no cleaner	negative	high	not cleaner	negative	high	constantly	positive	median
no cleanest	negative	max	not cleanest	negative	max	courageously	positive	high
no cleanly	negative	median	not cleanly	negative	median	devotedly	positive	median
no clear	negative	median	not clear	negative	median	discreetly	positive	median
no clear	negative	median	not clear	negative	median	faithfully	positive	median
no clearer	negative	high	not clearer	negative	high	firmly	positive	median
no clever	negative	median	not clever	negative	median	honestly	positive	median
no cleverly	negative	median	not cleverly	negative	median	meticulously	positive	median
no clumsy	positive	median	not clumsy	positive	median	patiently	positive	median
no coarse	positive	median	not coarse	positive	median	reliably	positive	median
no coarsely	positive	median	not coarsely	positive	median	resolutely	positive	median
no cocky	positive	median	not cocky	positive	median	safely	positive	median
no coherent	negative	median	not coherent	negative	median	stalwartly	positive	median
no colorful	negative	median	not colorful	negative	median	staunchly	positive	high
no colourful	negative	median	not colourful	negative	median	steadfastly	positive	high
no comely	negative	median	not comely	negative	median	steadily	positive	median
no comfortable	negative	median	not comfortable	negative	median	tirelessly	positive	median
no comic	negative	median	not comic	negative	median	unfailingly	positive	median
no comically	negative	median	not comically	negative	median	uprightly	positive	median
no commendable	negative	median	not commendable	negative	median	valiantly	positive	median
no common	positive	median	not common	positive	median	erratically	negative	median
no commoner	positive	high	not commoner	positive	high	impatiently	negative	median
no commonly	positive	median	not commonly	positive	median	recklessly	negative	high
no commonplace	positive	median	not commonplace	positive	median	shakily	negative	median
no compatible	negative	median	not compatible	negative	median	softly	negative	median
no compelling	negative	median	not compelling	negative	median	spinelessly	negative	high
no competent	negative	low	not competent	negative	low	stubbornly	negative	median
no competently	negative	median	not competently	negative	median	weakly	negative	median
no competitive	negative	median	not competitive	negative	median	auspicious	positive	median
no completely	minimize	not completely	minimize	celebrated	positive	median		
no complex	positive	median	not complex	positive	median	charmed	positive	median
no complexly	positive	median	not complexly	positive	median	cool	positive	median
no comprehensive	negative	median	not comprehensive	negative	median	cooler	positive	high
no comprehensively	decrease	not comprehensively	decrease	customary	positive	median		
no compulsive	positive	median	not compulsive	positive	median	distinguished	positive	median
no conceited	positive	median	not conceited	positive	median	eminent	positive	median
no concise	negative	median	not concise	negative	median	familiar	positive	median
no confident	negative	median	not confident	negative	median	famous	positive	high
no confidently	negative	median	not confidently	negative	median	famous	positive	high
no confusing	positive	median	not confusing	positive	median	fashionable	positive	median
no congenial	negative	median	not congenial	negative	median	favored	positive	median

no conscientious	negative	median	not conscientious	negative	median	fortunate	positive	median
no considerably	decrease	not considerably	decrease	great	positive	median		
no considerate	negative	median	not considerate	negative	median	illustrious	positive	high
no considered	negative	median	not considered	negative	median	intimate	positive	median
no consistent	negative	median	not consistent	negative	median	lucky	positive	median
no consistently	negative	median	not consistently	negative	median	natural	positive	median
no consonant	negative	median	not consonant	negative	median	normal	positive	median
no conspicuous	positive	median	not conspicuous	positive	median	notable	positive	median
no constant	negative	median	not constant	negative	median	noted	positive	median
no constantly	negative	median	not constantly	negative	median	popular	positive	median
no contemporary	negative	median	not contemporary	negative	median	predictably	positive	median
no contented	negative	median	not contented	negative	median	predictable	positive	median
no contentedly	negative	median	not contentedly	negative	median	prominent	positive	high
no contradictory	positive	median	not contradictory	positive	median	propitious	positive	median
no controversial	positive	low	not controversial	positive	low	renowned	positive	median
no conventional	positive	median	not conventional	positive	median	stable	positive	median
no conventionally	positive	median	not conventionally	positive	median	typical	positive	median
no convincing	negative	median	not convincing	negative	median	unsung	positive	median
no convincingly	negative	median	not convincingly	negative	median	usual	positive	median
no convoluted	positive	high	not convoluted	positive	high	abnormal	negative	median
no cool	negative	median	not cool	negative	median	bizarre	negative	high
no cooler	negative	high	not cooler	negative	high	daggy	negative	median
no coolest	negative	max	not coolest	negative	max	dated	negative	median
no coolly	negative	median	not coolly	negative	median	eccentric	negative	median
no corniest	positive	max	not corniest	positive	max	freak	negative	median
no corny	positive	high	not corny	positive	high	freakish	negative	median
no correct	negative	median	not correct	negative	median	freaky	negative	median
no corrupt	positive	median	not corrupt	positive	median	hapless	negative	median
no courageous	negative	high	not courageous	negative	high	idiosyncratic	negative	median
no courageously	negative	high	not courageously	negative	high	irregular	negative	median
no courteous	negative	median	not courteous	negative	median	kooky	negative	high
no courteously	negative	median	not courteously	negative	median	luckless	negative	median
no covetous	positive	median	not covetous	positive	median	nameless	negative	median
no cowardly	positive	high	not cowardly	positive	high	obscure	negative	median
no crafty	positive	high	not crafty	positive	high	odd	negative	median
no cranky	positive	median	not cranky	positive	median	oddball	negative	median
no crazed	positive	median	not crazed	positive	median	outlandish	positive	high
no crazy	positive	median	not crazy	positive	median	peculiar	negative	median
no creative	negative	median	not creative	negative	median	queer	negative	high
no creatively	negative	median	not creatively	negative	median	retrograde	negative	median
no credible	negative	median	not credible	negative	median	strange	negative	median
no creepily	positive	high	not creepily	positive	high	stranger	negative	high
no creepy	positive	high	not creepy	positive	high	unidentified	negative	median
no crestfallen	positive	median	not crestfallen	positive	median	unilluminated	negative	median
no criminal	positive	median	not criminal	positive	median	unknown	negative	median
no crippled	positive	median	not crippled	positive	median	unlucky	negative	median
no critical	negative	median	not critical	negative	median	unrenowned	negative	median
no crooked	positive	median	not crooked	positive	median	unusual	negative	median
no cross	positive	median	not cross	positive	median	weird	negative	median
no crucial	negative	high	not crucial	negative	high	weirdo	negative	median
no crude	positive	median	not crude	positive	median	whacko	negative	high
no crudely	positive	median	not crudely	positive	median	whacky	negative	median
no cruel	positive	high	not cruel	positive	high	wretched	negative	median
no cruelly	positive	high	not cruelly	positive	high	greater	positive	high
no crummy	positive	median	not crummy	positive	median	coolest	positive	max
no cryptic	positive	median	not cryptic	positive	median	greatest	positive	max
no cumbersome	positive	median	not cumbersome	positive	median	farthest	negative	max
no cunning	positive	high	not cunning	positive	high	strangest	negative	max
no cunningly	positive	high	not cunningly	positive	high	coolly	positive	median
no curvaceous	negative	median	not curvaceous	negative	median	eminently	positive	median
no customary	negative	median	not customary	negative	median	fortunately	positive	median
no cute	negative	median	not cute	negative	median	intimately	positive	median
no cuter	negative	high	not cuter	negative	high	normally	positive	median
no cutest	negative	max	not cutest	negative	max	popularly	positive	median
no cynical	positive	median	not cynical	positive	median	prominently	positive	high
no daft	positive	high	not daft	positive	high	renownedly	positive	high
no daggy	positive	median	not daggy	positive	median	stably	positive	median
no damaging	positive	median	not damaging	positive	median	typically	positive	median
no darned	positive	median	not darned	positive	median	usually	positive	median
no dated	positive	median	not dated	positive	median	abnormally	negative	median
no dauntless	negative	high	not dauntless	negative	high	bizarrely	negative	high
no dazzling	negative	high	not dazzling	negative	high	eccentrically	negative	median
no deadly	negative	median	not deadly	negative	median	irregularly	negative	median
no dear	negative	median	not dear	negative	median	obscurely	negative	median
no deceitful	positive	high	not deceitful	positive	high	oddly	negative	median
no deceiving	positive	median	not deceiving	positive	median	outrageously	negative	high
no decent	negative	low	not decent	negative	low	outstandingly	positive	high
no decently	negative	low	not decently	negative	low	outwardly	negative	median
no deceptive	positive	median	not deceptive	positive	median	strangely	negative	median
no deceptively	positive	median	not deceptively	positive	median	weirdly	negative	median
no decisive	negative	median	not decisive	negative	median	able	positive	median
no decorative	negative	low	not decorative	negative	low	accomplished	positive	median
no deep	negative	median	not deep	negative	median	adept	positive	very high
no deeper	negative	high	not deeper	negative	high	adroit	positive	high
no deepest	negative	max	not deepest	negative	max	adult	positive	median
no deeply	negative	median	not deeply	negative	median	astute	positive	median
no defective	positive	median	not defective	positive	median	balanced	positive	median
no definitely	decrease	not definitely	decrease	capable	positive	median		
no definitive	negative	max	not definitive	negative	max	clever	positive	median
no degenerate	positive	median	not degenerate	positive	median	competent	positive	low
no dejected	positive	median	not dejected	positive	median	droll	positive	median
no delicate	negative	median	not delicate	negative	median	educated	positive	median
no delicious	negative	median	not delicious	negative	median	experienced	positive	median
no delighted	negative	median	not delighted	negative	median	expert	positive	high
no delightful	negative	median	not delightful	negative	median	fit	positive	median

no delightfully	negative	median	not delightfully	negative	median	gifted	positive	median
no deliriously	decrease	not deliriously	decrease	hale	positive	median		
no delusive	positive	median	not delusive	positive	median	healthy	positive	median
no delusory	positive	median	not delusory	positive	median	humorous	positive	median
no demented	positive	median	not demented	positive	median	insightful	positive	median
no dense	positive	high	not dense	positive	high	intelligent	positive	median
no densely	positive	high	not densely	positive	high	knowing	positive	median
no densest	positive	max	not densest	positive	max	learned	positive	median
no dependable	negative	median	not dependable	negative	median	literate	positive	median
no depraved	positive	high	not depraved	positive	high	masterly	positive	high
no depressed	positive	median	not depressed	positive	median	mature	positive	median
no depressing	positive	median	not depressing	positive	median	powerful	positive	median
no depressingly	positive	median	not depressingly	positive	median	productive	positive	median
no deranged	positive	median	not deranged	positive	median	proficient	positive	median
no derivative	positive	median	not derivative	positive	median	qualified	positive	low
no deserving	negative	median	not deserving	negative	median	robust	positive	median
no desirable	negative	median	not desirable	negative	median	sagacious	positive	high
no desolate	positive	median	not desolate	positive	median	sane	positive	median
no despairing	positive	high	not despairing	positive	high	sensible	positive	median
no desperate	positive	median	not desperate	positive	median	sharp	positive	median
no despicable	positive	median	not despicable	positive	median	shrewd	positive	median
no despondent	positive	median	not despondent	positive	median	skilled	positive	median
no destructive	positive	median	not destructive	positive	median	skillful	positive	median
no detailed	negative	median	not detailed	negative	median	successful	positive	median
no detailed	negative	median	not detailed	negative	median	successful	positive	median
no devilish	positive	median	not devilish	positive	median	together	positive	median
no devious	negative	high	not devious	negative	high	trained	positive	median
no devoted	negative	median	not devoted	negative	median	vigorous	positive	median
no devotedly	negative	median	not devotedly	negative	median	wise	positive	median
no diligent	negative	median	not diligent	negative	median	witty	positive	median
no dim	positive	median	not dim	positive	median	smart	positive	median
no dime-a-dozen	positive	median	not dime-a-dozen	positive	median	talented	positive	median
no dimly	positive	median	not dimly	positive	median	professional	positive	median
no dim-witted	positive	high	not dim-witted	positive	high	sophisticated	positive	median
no direct	negative	median	not direct	negative	median	ambitious	positive	median
no directly	negative	median	not directly	negative	median	masterful	positive	high
no dirty	positive	median	not dirty	positive	median	intellectual	positive	median
no disabled	positive	median	not disabled	positive	median	imaginative	positive	median
no disagreeable	positive	median	not disagreeable	positive	median	backward	negative	median
no disappointing	positive	median	not disappointing	positive	median	boring	negative	median
no disastrous	positive	median	not disastrous	positive	median	brainless	negative	median
no discerning	negative	median	not discerning	negative	median	callow	negative	median
no disconcerting	positive	median	not disconcerting	positive	median	childish	negative	median
no disconsolate	positive	very high	not disconsolate	positive	very high	crazed	negative	median
no discordant	positive	high	not discordant	positive	high	crazy	negative	median
no discouraged	positive	median	not discouraged	positive	median	crippled	negative	median
no discouraging	positive	median	not discouraging	positive	median	daft	negative	high
no discourteous	positive	median	not discourteous	positive	median	demented	negative	median
no discreetly	negative	median	not discreetly	negative	median	dense	negative	high
no discrete	negative	median	not discrete	negative	median	deranged	negative	median
no discriminating	negative	median	not discriminating	negative	median	dim	negative	median
no disgusting	positive	median	not disgusting	positive	median	doltish	negative	median
no disheartened	positive	high	not disheartened	positive	high	dreary	negative	median
no dishonest	positive	median	not dishonest	positive	median	dull	negative	median
no dishonorable	positive	median	not dishonorable	positive	median	dumb	negative	median
no dishy	negative	median	not dishy	negative	median	feeble	negative	median
no disloyal	positive	median	not disloyal	positive	median	flaky	negative	median
no dismal	positive	max	not dismal	positive	max	foolish	negative	median
no dismally	positive	median	not dismally	positive	median	grave	negative	median
no dismayed	positive	median	not dismayed	positive	median	helpless	negative	median
no dismayedly	positive	median	not dismayedly	positive	median	ignorant	negative	median
no disorganised	positive	median	not disorganised	positive	median	illiterate	negative	median
no dispirited	positive	median	not dispirited	positive	median	imbecilic	negative	high
no disproportionate	positive	median	not disproportionate	positive	median	immature	negative	median
no dissatisfied	positive	median	not dissatisfied	positive	median	impotent	negative	median
no dissonant	positive	high	not dissonant	positive	high	incapable	negative	median
no distasteful	positive	median	not distasteful	positive	median	incompetent	negative	median
no distinctive	negative	median	not distinctive	negative	median	inexperienced	negative	median
no distinguished	negative	median	not distinguished	negative	median	inexpert	negative	median
no distorted	positive	high	not distorted	positive	high	infantile	negative	median
no distracted	positive	low	not distracted	positive	low	infirm	negative	median
no distraught	positive	median	not distraught	positive	median	insane	negative	median
no distressed	positive	median	not distressed	positive	median	lunatic	negative	median
no disturbed	positive	median	not disturbed	positive	median	mild	negative	median
no disturbing	positive	median	not disturbing	positive	median	moronic	negative	high
no dizzy	positive	median	not dizzy	positive	median	naive	negative	median
no doleful	positive	high	not doleful	positive	high	neurotic	negative	median
no doltish	positive	median	not doltish	positive	median	numskulled	negative	high
no doting	negative	median	not doting	negative	median	obtuse	negative	median
no double-dealing	positive	high	not double-dealing	positive	high	puerile	negative	median
no down	positive	median	not down	positive	median	sick	negative	high
no downcast	positive	median	not downcast	positive	median	simple	negative	median
no downcastly	positive	median	not downcastly	positive	median	slack	negative	median
no downhearted	positive	median	not downhearted	positive	median	slow	negative	median
no downheartedly	positive	median	not downheartedly	positive	median	slow	negative	median
no downly	positive	median	not downly	positive	median	sluggish	negative	median
no drab	positive	median	not drab	positive	median	stupid	negative	median
no dramatic	negative	high	not dramatic	negative	high	stupider	negative	high
no dramatic	negative	median	not dramatic	negative	median	thick	negative	median
no dramatically	negative	median	not dramatically	negative	median	thickheaded	negative	median
no drastic	positive	median	not drastic	positive	median	unaccomplished	negative	median
no dreadful	positive	median	not dreadful	positive	median	unbalanced	negative	median
no dreary	positive	median	not dreary	positive	median	uncultured	negative	median
no droll	negative	median	not droll	negative	median	uneducated	negative	median
no dry	positive	median	not dry	positive	median	unenlightened	negative	median

no dubious	positive	median	not dubious	positive	median	unfit	negative	median
no dull	positive	median	not dull	positive	median	unhealthy	negative	median
no duller	positive	high	not duller	positive	high	unhinged	negative	median
no dullest	positive	max	not dullest	positive	max	uninstructed	negative	median
no dully	positive	median	not dully	positive	median	unintelligent	negative	median
no dumb	positive	median	not dumb	positive	median	unlearned	negative	median
no dumber	positive	high	not dumber	positive	high	unlettered	negative	high
no dumbest	positive	max	not dumbest	positive	max	unproductive	negative	median
no duplicitous	positive	median	not duplicitous	positive	median	unprotected	negative	median
no durable	negative	median	not durable	negative	median	unqualified	negative	median
no dutiful	negative	median	not dutiful	negative	median	unschooled	negative	median
no dutifully	negative	median	not dutifully	negative	median	unskilled	negative	median
no dynamic	negative	median	not dynamic	negative	median	unsophisticated	negative	median
no eager	negative	median	not eager	negative	median	unsound	negative	median
no eagerly	negative	median	not eagerly	negative	median	unsuccessful	negative	median
no earnest	negative	median	not earnest	negative	median	untaught	negative	median
no easier	negative	high	not easier	negative	high	untrained	negative	median
no easiest	negative	max	not easiest	negative	max	untutored	negative	median
no easily	negative	median	not easily	negative	median	whimpy	negative	median
no easy	negative	median	not easy	negative	median	witless	negative	median
no eccentric	positive	median	not eccentric	positive	median	idiotic	negative	high
no eccentrically	positive	median	not eccentrically	positive	median	mindless	negative	median
no eclectic	negative	low	not eclectic	negative	low	clumsy	negative	median
no ecstatic	negative	high	not ecstatic	negative	high	saner	positive	high
no edgy	negative	median	not edgy	negative	median	sharper	positive	high
no educated	negative	median	not educated	negative	median	wiser	positive	high
no effective	negative	median	not effective	negative	median	duller	negative	high
no effectively	negative	median	not effectively	negative	median	milder	negative	high
no effervescent	negative	high	not effervescent	negative	high	mildest	negative	max
no efficient	negative	median	not efficient	negative	median	sicker	negative	high
no effortless	negative	median	not effortless	negative	median	simpler	negative	high
no egoistic	positive	median	not egoistic	positive	median	slower	negative	high
no egotistic	positive	median	not egotistic	positive	median	thicker	negative	high
no elaborate	negative	median	not elaborate	negative	median	sharpest	positive	max
no elaborately	negative	median	not elaborately	negative	median	wisest	positive	max
no elated	negative	high	not elated	negative	high	densest	negative	max
no electrifying	negative	very high	not electrifying	negative	very high	dullest	negative	max
no elegant	negative	median	not elegant	negative	median	simplest	negative	max
no elegantly	negative	median	not elegantly	negative	median	slowest	negative	max
no elementary	positive	median	not elementary	positive	median	stupidest	negative	max
no eminent	negative	median	not eminent	negative	median	thickest	negative	max
no eminently	negative	median	not eminently	negative	median	cleverly	positive	median
no emotional	negative	median	not emotional	negative	median	competently	positive	median
no emotionally	negative	median	not emotionally	negative	median	expertly	positive	high
no emphatically	decrease	not emphatically	decrease	fitfully	positive	median		
no empty-headed	positive	high	not empty-headed	positive	high	fittingly	positive	median
no enchanting	negative	high	not enchanting	negative	high	imaginatively	positive	median
no endearing	negative	median	not endearing	negative	median	intellectually	positive	median
no energetic	negative	median	not energetic	negative	median	intelligently	positive	median
no energetically	negative	median	not energetically	negative	median	knowingly	positive	high
no engaged	negative	median	not engaged	negative	median	masterfully	positive	high
no engaging	negative	median	not engaging	negative	median	powerfully	positive	median
no engagingly	negative	median	not engagingly	negative	median	professionally	positive	median
no engrossed	negative	high	not engrossed	negative	high	sagaciously	positive	high
no engrossing	negative	high	not engrossing	negative	high	sanely	positive	median
no enigmatic	negative	median	not enigmatic	negative	median	sensibly	positive	median
no enjoyable	negative	median	not enjoyable	negative	median	sharply	positive	median
no enormously	decrease	not enormously	decrease	shrewdly	positive	median		
no entertained	negative	median	not entertained	negative	median	skillfully	positive	median
no enthralled	negative	high	not enthralled	negative	high	smartly	positive	median
no enthralling	negative	high	not enthralling	negative	high	sophisticatedly	positive	median
no enthusiastic	negative	median	not enthusiastic	negative	median	soundly	positive	median
no enthusiastically	negative	median	not enthusiastically	negative	median	successfully	positive	median
no entirely	decrease	not entirely	decrease	vigorously	positive	median		
no entranced	negative	high	not entranced	negative	high	wisely	positive	median
no entrancing	negative	high	not entrancing	negative	high	backwardly	negative	median
no equitable	negative	median	not equitable	negative	median	childishly	negative	median
no erotic	negative	median	not erotic	negative	median	densely	negative	high
no erratic	positive	median	not erratic	positive	median	dimly	negative	median
no erratically	positive	median	not erratically	positive	median	dully	negative	median
no essential	negative	high	not essential	negative	high	foolishly	negative	median
no ethical	negative	median	not ethical	negative	median	helplessly	negative	median
no euphoric	negative	median	not euphoric	negative	median	insanely	increase	
no evenly	negative	median	not evenly	negative	median	moronically	negative	high
no everyday	positive	median	not everyday	positive	median	naively	negative	median
no evil	positive	high	not evil	positive	high	sickly	negative	median
no evilly	positive	high	not evilly	positive	high	sickeningly	negative	high
no exact	negative	median	not exact	negative	median	slackly	negative	median
no exceedingly	decrease	not exceedingly	decrease	slowly	negative	median		
no excellent	negative	high	not excellent	negative	high	sluggishly	negative	median
no excellent	negative	high	not excellent	negative	high	stupidly	negative	median
no excellently	negative	high	not excellently	negative	high	unsuccessfully	negative	median
no exceptional	negative	very high	not exceptional	negative	very high	altruistic	positive	median
no exceptionally	decrease	not exceptionally	decrease	beneficial	positive	median		
no excessive	positive	median	not excessive	positive	median	benevolent	positive	median
no excessively	decrease	not excessively	decrease	blameless	positive	max		
no excited	negative	median	not excited	negative	median	caring	positive	median
no exciting	negative	median	not exciting	negative	median	charitable	positive	median
no exclusive	negative	median	not exclusive	negative	median	chaste	positive	median
no excruciating	positive	high	not excruciating	positive	high	courteous	positive	median
no exhaustive	negative	median	not exhaustive	negative	median	decent	positive	low
no exhilarating	negative	median	not exhilarating	negative	median	dutiful	positive	median
no exotic	negative	median	not exotic	negative	median	ethical	positive	median
no experienced	negative	median	not experienced	negative	median	fair	positive	median
no expert	negative	high	not expert	negative	high	friendly	positive	median

no expertly	negative	high	not expertly	negative	high	generous	positive	median
no explosive	negative	median	not explosive	negative	median	good	positive	median
no expressive	negative	median	not expressive	negative	median	gracious	positive	high
no exquisite	negative	high	not exquisite	negative	high	helpful	positive	median
no exquisitely	negative	high	not exquisitely	negative	high	humane	positive	low
no extensive	negative	median	not extensive	negative	median	humanitarian	positive	median
no extraordinarily	decrease	not extraordinarily	decrease	humble	positive	median		
no extraordinary	negative	very high	not extraordinary	negative	very high	kind	positive	median
no extravagant	positive	median	not extravagant	positive	median	kindly	positive	median
no extreme	negative	median	not extreme	negative	median	kindly	positive	median
no extremely	decrease	not extremely	decrease	magnanimous	positive	high		
no exuberant	negative	high	not exuberant	negative	high	meek	positive	median
no exultant	negative	high	not exultant	negative	high	modest	positive	median
no exulting	negative	high	not exulting	negative	high	moral	positive	median
no fabulous	negative	high	not fabulous	negative	high	noble	positive	high
no faint	positive	median	not faint	positive	median	nobler	positive	very high
no fair	negative	median	not fair	negative	median	obliging	positive	median
no fairly	increase	not fairly	increase	philanthropic	positive	median		
no faithful	negative	median	not faithful	negative	median	polite	positive	median
no faithfully	negative	median	not faithfully	negative	median	praiseworthy	positive	high
no fake	positive	median	not fake	positive	median	principled	positive	median
no fallacious	positive	median	not fallacious	positive	median	proper	positive	median
no famed	negative	median	not famed	negative	median	respectable	positive	median
no familiar	negative	median	not familiar	negative	median	respectful	positive	median
no famous	negative	high	not famous	negative	high	reverent	positive	median
no famous	negative	high	not famous	negative	high	right	positive	median
no fancy	positive	median	not fancy	positive	median	righteous	positive	high
no fantastic	negative	median	not fantastic	negative	median	sensitive	positive	median
no fantastically	negative	median	not fantastically	negative	median	sympathetic	positive	median
no far	decrease	not far	decrease	thoughtful	positive	median		
no farthest	positive	max	not farthest	positive	max	tolerant	positive	median
no fascinated	negative	median	not fascinated	negative	median	unassuming	positive	median
no fascinating	negative	median	not fascinating	negative	median	unostentatious	positive	median
no fashionable	negative	median	not fashionable	negative	median	unpretentious	positive	median
no favorable	negative	median	not favorable	negative	median	upstanding	positive	median
no favored	negative	median	not favored	negative	median	virtuous	positive	high
no favorite	negative	max	not favorite	negative	max	worthy	positive	median
no fearful	positive	median	not fearful	positive	median	arrogant	negative	median
no fearfully	positive	median	not fearfully	positive	median	atrocious	negative	high
no fearless	negative	high	not fearless	negative	high	avaricious	negative	median
no feasible	negative	low	not feasible	negative	low	bad	negative	median
no feeble	positive	median	not feeble	positive	median	barbarous	negative	median
no felicitous	negative	median	not felicitous	negative	median	base	negative	median
no ferocious	negative	median	not ferocious	negative	median	beggarly	negative	median
no fickle	positive	median	not fickle	positive	median	brutal	negative	high
no fictional	positive	median	not fictional	positive	median	callous	negative	median
no fidgety	positive	high	not fidgety	positive	high	conceited	negative	median
no fierce	negative	median	not fierce	negative	median	corrupt	negative	median
no filthy	positive	median	not filthy	positive	median	covetous	negative	median
no fine	negative	median	not fine	negative	median	crooked	negative	median
no finely	negative	median	not finely	negative	median	crude	negative	median
no finer	negative	high	not finer	negative	high	cruel	negative	high
no finest	negative	max	not finest	negative	max	degenerate	negative	median
no finicky	positive	median	not finicky	positive	median	depraved	negative	high
no firm	negative	median	not firm	negative	median	dirty	negative	median
no firmer	negative	high	not firmer	negative	high	discourteous	negative	median
no firmly	negative	median	not firmly	negative	median	dishonorable	negative	median
no first-class	negative	median	not first-class	negative	median	egoistic	negative	median
no first-rate	negative	median	not first-rate	negative	median	egotistic	negative	median
no fishy	positive	median	not fishy	positive	median	evil	negative	high
no fit	negative	median	not fit	negative	median	foul	negative	median
no fitfully	negative	median	not fitfully	negative	median	grabby	negative	median
no fitting	negative	low	not fitting	negative	low	grasping	negative	median
no fittingly	negative	median	not fittingly	negative	median	greedy	negative	median
no flaky	positive	median	not flaky	positive	median	harsh	negative	median
no flashy	positive	median	not flashy	positive	median	hateful	negative	median
no flat	positive	median	not flat	positive	median	haughty	negative	median
no flatly	positive	median	not flatly	positive	median	immoral	negative	median
no flawed	positive	median	not flawed	positive	median	impertinent	negative	median
no flawless	negative	max	not flawless	negative	max	impolite	negative	median
no flexible	negative	median	not flexible	negative	median	indecent	negative	high
no flexible	negative	median	not flexible	negative	median	inequitable	negative	median
no flimsily	positive	median	not flimsily	positive	median	iniquitous	negative	median
no flimsy	positive	median	not flimsy	positive	median	insensitive	negative	median
no float	positive	median	not float	positive	median	insolent	negative	median
no fly-by-night	positive	median	not fly-by-night	positive	median	irreverent	negative	median
no fond	negative	median	not fond	negative	median	low	negative	median
no fonder	negative	high	not fonder	negative	high	malevolent	negative	median
no fondest	negative	max	not fondest	negative	max	mean	negative	median
no fondly	negative	median	not fondly	negative	median	meaner	negative	high
no foolhardy	positive	high	not foolhardy	positive	high	mercenary	negative	median
no foolish	positive	median	not foolish	positive	median	miserly	negative	median
no foolishly	positive	median	not foolishly	positive	median	nefarious	negative	high
no foolproof	negative	median	not foolproof	negative	median	obscene	negative	median
no forbidding	positive	median	not forbidding	positive	median	parsimonious	negative	median
no foremost	negative	median	not foremost	negative	median	partial	negative	median
no forlorn	positive	high	not forlorn	positive	high	petty	negative	median
no formal	negative	median	not formal	negative	median	reprobate	negative	median
no formidable	negative	median	not formidable	negative	median	rotten	negative	high
no formless	positive	median	not formless	positive	median	rough	negative	median
no formulaic	positive	median	not formulaic	positive	median	rude	negative	median
no forthright	negative	median	not forthright	negative	median	savage	negative	median
no fortunate	negative	median	not fortunate	negative	median	selfish	negative	median
no fortunately	negative	median	not fortunately	negative	median	sinful	negative	median
no foul	positive	median	not foul	positive	median	snobby	negative	median

no foully	positive	median	not foully	positive	median	unchaste	negative	median
no foxy	positive	median	not foxy	positive	median	uncivil	negative	median
no fragile	positive	median	not fragile	positive	median	uncivilized	negative	median
no frank	negative	median	not frank	negative	median	unclean	negative	median
no franker	negative	high	not franker	negative	high	uncouth	negative	median
no frankly	negative	median	not frankly	negative	median	unethical	negative	median
no frantic	positive	median	not frantic	positive	median	unfair	negative	median
no fraudulent	positive	median	not fraudulent	positive	median	unfeeling	negative	median
no freak	positive	median	not freak	positive	median	ungracious	negative	median
no freakish	positive	median	not freakish	positive	median	unjust	negative	median
no freaky	positive	median	not freaky	positive	median	unkind	negative	median
no frenzied	positive	median	not frenzied	positive	median	unmannerly	negative	median
no fresh	negative	median	not fresh	negative	median	unpleasant	negative	median
no freshly	negative	median	not freshly	negative	median	unpolished	negative	median
no fretful	positive	median	not fretful	positive	median	unprincipled	negative	median
no friendly	negative	median	not friendly	negative	median	unrefined	negative	median
no frightened	positive	median	not frightened	positive	median	unrighteous	negative	median
no frightful	positive	median	not frightful	positive	median	unscrupulous	negative	median
no frightfully	positive	median	not frightfully	positive	median	vain	negative	median
no frivolous	positive	median	not frivolous	positive	median	vicious	negative	median
no fruitless	positive	median	not fruitless	positive	median	vile	negative	high
no frustrated	positive	median	not frustrated	positive	median	villainous	negative	median
no frustrating	positive	median	not frustrating	positive	median	vulgar	negative	high
no fully	decrease	not fully	decrease	wicked	negative	median		
no fuming	positive	median	not fuming	positive	median	wrongful	negative	median
no fun	negative	median	not fun	negative	median	kinder	positive	high
no functional	negative	low	not functional	negative	low	baser	negative	high
no functioning	negative	low	not functioning	negative	low	harsher	negative	high
no fundamental	negative	median	not fundamental	negative	median	lower	negative	high
no funnier	negative	high	not funnier	negative	high	rougher	negative	high
no funniest	negative	max	not funniest	negative	max	roughest	negative	max
no funnily	negative	median	not funnily	negative	median	kindest	positive	max
no funny	negative	median	not funny	negative	median	noblest	positive	max
no furious	positive	high	not furious	positive	high	lowest	negative	max
no furiously	positive	high	not furiously	positive	high	meanest	negative	max
no fussy	positive	median	not fussy	positive	median	courteously	positive	median
no gallant	negative	high	not gallant	negative	high	decently	positive	low
no gay	positive	median	not gay	positive	median	dutifully	positive	median
no generous	negative	median	not generous	negative	median	generously	positive	median
no generously	negative	median	not generously	negative	median	graciously	positive	high
no gentle	negative	median	not gentle	negative	median	helpfully	positive	median
no genuine	negative	high	not genuine	negative	high	justifiably	positive	median
no genuinely	negative	high	not genuinely	negative	high	most	maximize	
no ghastly	positive	high	not ghastly	positive	high	modestly	positive	median
no gifted	negative	median	not gifted	negative	median	morally	positive	median
no glad	negative	median	not glad	negative	median	obligingly	positive	median
no gladly	negative	median	not gladly	negative	median	politely	positive	median
no glamorous	negative	median	not glamorous	negative	median	properly	positive	median
no glaring	decrease	not glaring	decrease	respectably	positive	median		
no gleeful	negative	high	not gleeful	negative	high	respectfully	positive	median
no gleefully	negative	high	not gleefully	negative	high	reverentially	positive	median
no glitzy	positive	median	not glitzy	positive	median	righteously	positive	high
no gloomy	positive	high	not gloomy	positive	high	rightfully	positive	median
no glorious	negative	high	not glorious	negative	high	rightly	positive	median
no gloriously	negative	high	not gloriously	negative	high	sensitively	positive	median
no glum	positive	median	not glum	positive	median	sympathetically	positive	median
no god-awful	positive	high	not god-awful	positive	high	thoughtfully	positive	median
no golden	negative	median	not golden	negative	median	worthily	positive	median
no good	negative	median	not good	negative	median	arrogantly	negative	median
no good-for-nothing	positive	very high	not good-for-nothing	positive	very high	atrociously	negative	high
no good-looking	negative	median	not good-looking	negative	median	badly	negative	median
no goodly	negative	median	not goodly	negative	median	brutally	negative	high
no gorgeous	negative	median	not gorgeous	negative	median	crudely	negative	median
no gorgeously	negative	median	not gorgeously	negative	median	cruelly	negative	high
no gory	positive	median	not gory	positive	median	evilly	negative	high
no grabby	positive	median	not grabby	positive	median	foully	negative	median
no graceful	negative	median	not graceful	negative	median	harshly	negative	median
no gracious	negative	high	not gracious	negative	high	lowly	negative	median
no graciously	negative	high	not graciously	negative	high	lowly	negative	median
no grand	negative	median	not grand	negative	median	reprobately	negative	median
no grander	negative	high	not grander	negative	high	roughly	negative	median
no grandiosely	negative	high	not grandiosely	negative	high	rudely	negative	median
no grasping	positive	median	not grasping	positive	median	savagely	negative	median
no grateful	negative	median	not grateful	negative	median	selfishly	negative	median
no gratifying	negative	median	not gratifying	negative	median	sinfully	negative	median
no gratuitous	positive	median	not gratuitous	positive	median	snobbily	negative	median
no gratuitously	positive	median	not gratuitously	positive	median	unfairly	negative	median
no grave	positive	median	not grave	positive	median	unpleasantly	negative	median
no great	negative	median	not great	negative	median	vainly	negative	median
no greater	negative	high	not greater	negative	high	wickedly	negative	median
no greatest	negative	max	not greatest	negative	max	wrongfully	negative	median
no greatly	decrease	not greatly	decrease	believable	positive	median		
no greedy	positive	median	not greedy	positive	median	candid	positive	median
no grim	positive	median	not grim	positive	median	credible	positive	median
no grimly	positive	median	not grimly	positive	median	direct	positive	median
no gripping	negative	median	not gripping	negative	median	discrete	positive	median
no grody	positive	median	not grody	positive	median	equitable	positive	median
no gross	positive	median	not gross	positive	median	forthright	positive	median
no grossly	positive	median	not grossly	positive	median	frank	positive	median
no grotesque	positive	high	not grotesque	positive	high	guileless	positive	median
no grotesquely	positive	high	not grotesquely	positive	high	ingenuous	positive	median
no gruesome	positive	median	not gruesome	positive	median	open	positive	median
no gruff	positive	low	not gruff	positive	low	plain	positive	median
no grumpy	positive	median	not grumpy	positive	median	plainer	positive	high
no guileful	positive	median	not guileful	positive	median	plainspoken	positive	median

no guileless	negative	median	not guileless	negative	median	sincere	positive	median
no guilty	positive	median	not guilty	positive	median	straight	positive	median
no gutless	positive	median	not gutless	positive	median	tactful	positive	median
no gutsy	negative	high	not gutsy	negative	high	truthful	positive	median
no hackneyed	positive	median	not hackneyed	positive	median	unbiased	positive	median
no hair-raising	negative	median	not hair-raising	negative	median	beguiling	negative	median
no hale	negative	median	not hale	negative	median	blabbermouth	negative	median
no half-baked	positive	median	not half-baked	positive	median	blunt	negative	median
no half-witted	positive	median	not half-witted	positive	median	crafty	negative	high
no handsome	negative	median	not handsome	negative	median	cunning	negative	high
no handsomely	negative	median	not handsomely	negative	median	deceitful	negative	high
no handy	negative	median	not handy	negative	median	deceiving	negative	median
no hapless	positive	median	not hapless	positive	median	deceptive	negative	median
no happier	negative	high	not happier	negative	high	delusive	negative	median
no happily	negative	median	not happily	negative	median	delusory	negative	median
no happy	negative	median	not happy	negative	median	devious	positive	high
no hard	positive	median	not hard	positive	median	dishonest	negative	median
no hardly	increase	not hardly	increase	duplicitous	negative	median		
no hardy	negative	median	not hardy	negative	median	fallacious	negative	median
no harmless	negative	low	not harmless	negative	low	fishy	negative	median
no harmonious	negative	median	not harmonious	negative	median	foxy	negative	median
no harsh	positive	median	not harsh	positive	median	fraudulent	negative	median
no harsher	positive	high	not harsher	positive	high	guileful	negative	median
no harshly	positive	median	not harshly	positive	median	insidious	negative	high
no has-been	positive	median	not has-been	positive	median	knavish	negative	median
no hasty	positive	median	not hasty	positive	median	lying	negative	median
no hateful	positive	median	not hateful	positive	median	manipulative	negative	median
no haughty	positive	median	not haughty	positive	median	mendacious	negative	median
no haunting	negative	median	not haunting	negative	median	misleading	negative	median
no hazy	positive	median	not hazy	positive	median	oblique	negative	median
no headstrong	positive	median	not headstrong	positive	median	roguish	negative	median
no healthy	negative	median	not healthy	negative	median	roundabout	negative	median
no heartbroken	positive	median	not heartbroken	positive	median	scheming	negative	high
no heartsick	positive	median	not heartsick	positive	median	shady	negative	median
no heavenly	negative	median	not heavenly	negative	median	shifty	negative	median
no heavy-hearted	positive	median	not heavy-hearted	positive	median	slick	negative	median
no helpful	negative	median	not helpful	negative	median	slippery	negative	median
no helpfully	negative	median	not helpfully	negative	median	sly	negative	median
no helpless	positive	median	not helpless	positive	median	sneaky	negative	median
no helplessly	positive	median	not helplessly	positive	median	tricky	negative	median
no heroic	negative	high	not heroic	negative	high	underhanded	negative	high
no hideous	positive	high	not hideous	positive	high	unhonest	negative	median
no hideously	positive	median	not hideously	positive	median	untruthful	negative	median
no high	decrease	not high	decrease	wily	negative	median		
no highly	decrease	not highly	decrease	franker	positive	high		
no high-resolution	negative	median	not high-resolution	negative	median	plainest	positive	max
no high-strung	positive	median	not high-strung	positive	median	sincerest	positive	max
no hilarious	negative	high	not hilarious	negative	high	candidly	positive	median
no hilariously	negative	high	not hilariously	negative	high	directly	positive	median
no hohum	positive	median	not hohum	positive	median	frankly	positive	median
no ho-hum	positive	median	not ho-hum	positive	median	openly	positive	median
no homely	positive	median	not homely	positive	median	plainly	positive	median
no homicidal	positive	median	not homicidal	positive	median	sincerely	positive	median
no honest	negative	median	not honest	negative	median	straightly	positive	median
no honestly	negative	median	not honestly	negative	median	straightforwardly	positive	median
no honorable	negative	median	not honorable	negative	median	truthfully	positive	median
no hopeful	negative	median	not hopeful	negative	median	bluntly	negative	median
no hopeless	positive	high	not hopeless	positive	high	cunningly	negative	high
no hopelessly	positive	median	not hopelessly	positive	median	deceptively	negative	median
no horrible	positive	high	not horrible	positive	high	insidiously	negative	high
no horribly	decrease	not horribly	decrease	schemingly	negative	high		
no horrid	positive	high	not horrid	positive	high	shadily	negative	median
no horrifying	positive	high	not horrifying	positive	high	shiftily	negative	median
no hostile	positive	median	not hostile	positive	median	slickly	negative	median
no hot	negative	median	not hot	negative	median	slyly	negative	median
no huge	decrease	not huge	decrease	sneakily	negative	median		
no hugely	minimize	not hugely	minimize	trickily	negative	median		
no humane	negative	low	not humane	negative	low	cheerful	positive	median
no humanitarian	negative	median	not humanitarian	negative	median	comfortable	positive	median
no humble	negative	median	not humble	negative	median	impressed	positive	median
no humdrum	positive	median	not humdrum	positive	median	happy	positive	median
no humorous	negative	median	not humorous	negative	median	confident	positive	median
no hysterical	negative	median	not hysterical	negative	median	assured	positive	median
no icy	positive	median	not icy	positive	median	trusting	positive	median
no idea	negative	median	not idea	negative	median	involved	positive	median
no ideal	negative	max	not ideal	negative	max	absorbed	positive	high
no ideally	negative	median	not ideally	negative	median	engrossed	positive	high
no idiosyncratic	positive	median	not idiosyncratic	positive	median	satisfied	positive	high
no idiotic	positive	high	not idiotic	positive	high	pleased	positive	median
no idle	positive	median	not idle	positive	median	chuffed	positive	median
no idyllic	negative	median	not idyllic	negative	median	thrilled	positive	high
no ignorant	positive	median	not ignorant	positive	median	buoyant	positive	median
no illegal	positive	median	not illegal	positive	median	jubilant	positive	high
no ill-fated	positive	median	not ill-fated	positive	median	fond	positive	median
no illiterate	positive	median	not illiterate	positive	median	loving	positive	median
no ill-mannered	positive	median	not ill-mannered	positive	median	adoring	positive	high
no ill-starred	positive	median	not ill-starred	positive	median	attentive	positive	median
no ill-tempered	positive	median	not ill-tempered	positive	median	busy	positive	median
no illuminating	negative	median	not illuminating	negative	median	industrious	positive	median
no illustrious	negative	high	not illustrious	negative	high	upbeat	positive	median
no imaginative	negative	median	not imaginative	negative	median	intent	positive	median
no imaginatively	negative	median	not imaginatively	negative	median	secure	positive	median
no imbalanced	positive	median	not imbalanced	positive	median	sanguine	positive	median
no imbecilic	positive	high	not imbecilic	positive	high	captivated	positive	high
no immature	positive	median	not immature	positive	median	warm	positive	median

no immense	negative	median	not immense	negative	median	rapt	positive	high
no immensely	minimize	not immensely	minimize	perky	positive	median		
no immoral	positive	median	not immoral	positive	median	occupied	positive	median
no impatient	positive	median	not impatient	positive	median	lively	positive	median
no impatiently	positive	median	not impatiently	positive	median	joyful	positive	median
no impeccable	negative	high	not impeccable	negative	high	exultant	positive	high
no imperceptibly	increase	not imperceptibly	increase	engaged	positive	median		
no imperfect	positive	median	not imperfect	positive	median	doting	positive	median
no impertinent	positive	median	not impertinent	positive	median	diligent	positive	median
no impetuous	positive	high	not impetuous	positive	high	chirpy	positive	high
no impolite	positive	median	not impolite	positive	median	bouncy	positive	high
no important	negative	median	not important	negative	median	blithe	positive	median
no importantly	negative	median	not importantly	negative	median	assiduous	positive	median
no impotent	positive	median	not impotent	positive	median	wrapped	positive	median
no impractical	positive	median	not impractical	positive	median	triumphant	positive	median
no imprecise	positive	median	not imprecise	positive	median	triumphal	positive	median
no impressed	negative	median	not impressed	negative	median	tender	positive	median
no impressive	negative	median	not impressive	negative	median	positive	positive	median
no impressively	negative	median	not impressively	negative	median	pleasant	positive	median
no impulsive	positive	median	not impulsive	positive	median	peppy	positive	high
no inaccurate	positive	median	not inaccurate	positive	median	gleeful	positive	high
no inadequate	positive	median	not inadequate	positive	median	glad	positive	median
no inane	positive	median	not inane	positive	median	fascinated	positive	median
no inappropriate	positive	median	not inappropriate	positive	median	exulting	positive	high
no incapable	positive	median	not incapable	positive	median	contented	positive	median
no incessant	positive	median	not incessant	positive	median	affectionate	positive	median
no incisive	negative	high	not incisive	negative	high	zealous	positive	median
no incoherent	positive	median	not incoherent	positive	median	overjoyed	positive	high
no incomparably	minimize	not incomparably	minimize	merry	positive	median		
no incompatible	positive	median	not incompatible	positive	median	enthralled	positive	high
no incompetent	positive	median	not incompetent	positive	median	energetic	positive	median
no incomplete	positive	median	not incomplete	positive	median	elated	positive	high
no incomprehensible	positive	median	not incomprehensible	positive	median	effervescent	positive	high
no incongruent	positive	median	not incongruent	positive	median	animated	positive	median
no incongruous	positive	median	not incongruous	positive	median	joyous	positive	median
no inconsistent	positive	median	not inconsistent	positive	median	jolly	positive	median
no inconsonant	positive	median	not inconsonant	positive	median	vivacious	positive	median
no inconstant	positive	median	not inconstant	positive	median	delighted	positive	median
no inconvenient	positive	median	not inconvenient	positive	median	cheery	positive	median
no incorrect	positive	median	not incorrect	positive	median	jocund	positive	median
no incredible	negative	median	not incredible	negative	median	mirthful	positive	median
no incredibly	minimize	not incredibly	minimize	keen	positive	median		
no indecent	positive	high	not indecent	positive	high	jovial	positive	high
no indefatigable	negative	median	not indefatigable	negative	median	lighthearted	positive	median
no independent	negative	median	not independent	negative	median	enthusiastic	positive	median
no indifferent	positive	low	not indifferent	positive	low	ecstatic	positive	high
no indignant	positive	median	not indignant	positive	median	eager	positive	median
no indispensable	negative	max	not indispensable	negative	max	felicitous	positive	median
no indubitable	negative	high	not indubitable	negative	high	euphoric	positive	median
no industrious	negative	median	not industrious	negative	median	excited	positive	median
no ineffective	positive	median	not ineffective	positive	median	entranced	positive	high
no ineffectual	positive	median	not ineffectual	positive	median	sad	negative	median
no inept	positive	median	not inept	positive	median	anxious	negative	median
no ineptly	positive	median	not ineptly	positive	median	bored	negative	median
no inequitable	positive	median	not inequitable	positive	median	angry	negative	median
no inestimable	negative	high	not inestimable	negative	high	fearful	negative	median
no inexcusable	positive	high	not inexcusable	positive	high	afraid	negative	median
no inexperienced	positive	median	not inexperienced	positive	median	melancholy	negative	median
no inexpert	positive	median	not inexpert	positive	median	dejected	negative	median
no infamous	positive	median	not infamous	positive	median	joyless	negative	max
no infantile	positive	median	not infantile	positive	median	cheerless	negative	max
no infectious	negative	median	not infectious	negative	median	unhappy	negative	median
no inferior	positive	median	not inferior	positive	median	downcast	negative	median
no infirm	positive	median	not infirm	positive	median	depressed	negative	median
no influential	negative	median	not influential	negative	median	uneasy	negative	median
no infuriated	positive	median	not infuriated	positive	median	startled	negative	median
no infuriating	positive	high	not infuriating	positive	high	surprised	positive	high
no ingenious	negative	high	not ingenious	negative	high	astonished	negative	median
no ingeniously	negative	high	not ingeniously	negative	high	float	negative	median
no ingenuous	negative	median	not ingenuous	negative	median	stale	negative	median
no inharmonious	positive	median	not inharmonious	positive	median	jaded	negative	median
no inimitable	negative	median	not inimitable	negative	median	cross	negative	median
no iniquitous	positive	median	not iniquitous	positive	median	furious	negative	high
no innocent	negative	median	not innocent	negative	median	scared	negative	median
no innovational	negative	median	not innovational	negative	median	terrorised	negative	high
no innovative	negative	high	not innovative	negative	high	terrorized	negative	high
no inoperable	positive	median	not inoperable	positive	median	restless	negative	median
no inordinately	decrease	not inordinately	decrease	fidgety	negative	high		
no insane	positive	median	not insane	positive	median	nervous	negative	median
no insanely	decrease	not insanely	decrease	dismal	negative	max		
no insensitive	positive	median	not insensitive	positive	median	aghast	negative	very high
no insidious	positive	high	not insidious	positive	high	down	negative	median
no insidiously	positive	high	not insidiously	positive	high	blue	negative	high
no insightful	negative	median	not insightful	negative	median	jumpy	negative	median
no insignificant	positive	median	not insignificant	positive	median	gloomy	negative	high
no insignificantly	positive	median	not insignificantly	positive	median	frightened	negative	median
no insipid	positive	median	not insipid	positive	median	dispirited	negative	median
no insolent	positive	median	not insolent	positive	median	apprehensive	negative	median
no inspirational	negative	median	not inspirational	negative	median	unquiet	negative	median
no inspiring	negative	median	not inspiring	negative	median	troubled	negative	median
no insufficient	positive	median	not insufficient	positive	median	raging	negative	median
no insulting	positive	high	not insulting	positive	high	mournful	negative	high
no intellectual	negative	median	not intellectual	negative	median	miserable	negative	median
no intellectually	negative	median	not intellectually	negative	median	disconsolate	negative	very high
no intelligent	negative	median	not intelligent	negative	median	woebegone	negative	median

no intelligently	negative	median	not intelligently	negative	median	terrible	negative	very high
no intense	negative	median	not intense	negative	median	terrible	negative	high
no intensely	decrease	not intensely	decrease	somber	negative	median		
no intent	negative	median	not intent	negative	median	nervy	negative	median
no interesting	negative	median	not interesting	negative	median	jittery	negative	median
no interestingly	negative	median	not interestingly	negative	median	glum	negative	median
no interminable	positive	high	not interminable	positive	high	fretful	negative	median
no intimate	negative	median	not intimate	negative	median	downhearted	negative	median
no intimately	negative	median	not intimately	negative	median	doleful	negative	high
no intimidated	positive	median	not intimidated	positive	median	disturbed	negative	median
no intoxicating	negative	high	not intoxicating	negative	high	distressed	negative	median
no intricate	negative	median	not intricate	negative	median	bleak	negative	median
no intricately	negative	median	not intricately	negative	median	terrified	negative	high
no intriguing	negative	high	not intriguing	negative	high	sorrowful	negative	median
no intriguingly	negative	high	not intriguingly	negative	high	oppressive	negative	median
no intrusive	positive	median	not intrusive	positive	median	grim	negative	median
no intuitive	negative	median	not intuitive	negative	median	forlorn	negative	high
no invaluable	negative	high	not invaluable	negative	high	depressing	negative	median
no inventive	negative	median	not inventive	negative	median	wrathful	negative	median
no inventively	negative	median	not inventively	negative	median	worried	negative	median
no invincible	negative	max	not invincible	negative	max	uptight	negative	median
no inviting	negative	median	not inviting	negative	median	perturbed	negative	low
no involved	negative	median	not involved	negative	median	morose	negative	median
no irascible	positive	median	not irascible	positive	median	agitated	negative	median
no irate	positive	median	not irate	positive	median	sullen	negative	median
no ireful	positive	median	not ireful	positive	median	spooked	negative	median
no irksome	positive	median	not irksome	positive	median	moody	negative	median
no irregular	positive	median	not irregular	positive	median	irritable	negative	median
no irregularly	positive	median	not irregularly	positive	median	ireful	negative	median
no irrelevant	positive	median	not irrelevant	positive	median	irate	negative	median
no irresistible	negative	high	not irresistible	negative	high	irascible	negative	median
no irresolute	positive	median	not irresolute	positive	median	infuriated	negative	median
no irreverent	positive	median	not irreverent	positive	median	indignant	negative	median
no irritable	positive	median	not irritable	positive	median	disheartened	negative	high
no irritating	positive	high	not irritating	positive	high	discouraged	negative	median
no jaded	positive	median	not jaded	positive	median	awful	negative	high
no jealous	positive	median	not jealous	positive	median	unsettled	negative	median
no jerky	positive	median	not jerky	positive	median	tense	negative	median
no jittery	positive	median	not jittery	positive	median	panicky	negative	median
no jocund	negative	median	not jocund	negative	median	intimidated	negative	median
no jolly	negative	median	not jolly	negative	median	heartsick	negative	median
no jovial	negative	high	not jovial	negative	high	fuming	negative	median
no joyful	negative	median	not joyful	negative	median	dismayed	negative	median
no joyless	positive	max	not joyless	positive	max	crestfallen	negative	median
no joyous	negative	median	not joyous	negative	median	dreadful	negative	median
no jubilant	negative	high	not jubilant	negative	high	desolate	negative	median
no juicy	negative	median	not juicy	negative	median	surly	negative	median
no jumpy	positive	median	not jumpy	positive	median	sulky	negative	median
no justifiably	negative	median	not justifiably	negative	median	spiritless	negative	median
no keen	negative	median	not keen	negative	median	peevish	negative	median
no keenly	negative	median	not keenly	negative	median	hopeless	negative	high
no key	negative	median	not key	negative	median	despairing	negative	high
no kind	negative	median	not kind	negative	median	pessimistic	negative	median
no kinder	negative	high	not kinder	negative	high	pathetic	negative	median
no kindest	negative	max	not kindest	negative	max	heartbroken	negative	median
no kindly	negative	median	not kindly	negative	median	desperate	negative	median
no kindly	negative	median	not kindly	negative	median	cantankerous	negative	median
no knavish	positive	median	not knavish	positive	median	mirthless	negative	median
no knowing	negative	median	not knowing	negative	median	grumpy	negative	median
no knowingly	negative	high	not knowingly	negative	high	mopey	negative	median
no knowledgeable	negative	median	not knowledgeable	negative	median	happier	positive	high
no know-nothing	positive	high	not know-nothing	positive	high	livelier	positive	high
no kooky	positive	high	not kooky	positive	high	warmer	positive	high
no lame	positive	median	not lame	positive	median	fonder	positive	high
no landmark	negative	high	not landmark	negative	high	fondest	positive	max
no largely	decrease	not largely	decrease	affectionately	positive	median		
no lasting	negative	median	not lasting	negative	median	assuredly	positive	median
no late	positive	median	not late	positive	median	attentively	positive	median
no lavish	negative	median	not lavish	negative	median	cheerfully	positive	median
no law-abiding	negative	median	not law-abiding	negative	median	cheerily	positive	median
no lazy	positive	median	not lazy	positive	median	confidently	positive	median
no learned	negative	median	not learned	negative	median	contentedly	positive	median
no legal	negative	median	not legal	negative	median	eagerly	positive	median
no legendary	negative	max	not legendary	negative	max	energetically	positive	median
no legitimate	negative	low	not legitimate	negative	low	enthusiastically	positive	median
no lengthy	positive	median	not lengthy	positive	median	fondly	positive	median
no less	increase	not less	increase	gladly	positive	median		
no level	negative	median	not level	negative	median	gleefully	positive	high
no lifeless	positive	median	not lifeless	positive	median	happily	positive	median
no lifelike	negative	median	not lifelike	negative	median	keenly	positive	median
no lighthearted	negative	median	not lighthearted	negative	median	pleasantly	positive	median
no limitless	negative	median	not limitless	negative	median	sanguinely	positive	median
no literate	negative	median	not literate	negative	median	satisfiedly	positive	median
no livelier	negative	high	not livelier	negative	high	securely	positive	median
no lively	negative	median	not lively	negative	median	tenderly	positive	median
no loathsome	positive	high	not loathsome	positive	high	togetherly	positive	median
no logical	negative	median	not logical	negative	median	anxiously	negative	median
no logically	negative	median	not logically	negative	median	awfully	negative	high
no lonely	positive	median	not lonely	positive	median	depressingly	negative	median
no longawaited	negative	median	not longawaited	negative	median	dismally	negative	median
no loose	positive	median	not loose	positive	median	dismayedly	negative	median
no lousy	positive	median	not lousy	positive	median	downly	negative	median
no lovely	negative	median	not lovely	negative	median	downcastly	negative	median
no lovely	negative	median	not lovely	negative	median	downheartedly	negative	median
no loving	negative	median	not loving	negative	median	fearfully	negative	median

no low	positive	median	not low	positive	median	furiously	negative	high
no lower	positive	high	not lower	positive	high	grimly	negative	median
no lowest	positive	max	not lowest	positive	max	hopelessly	negative	median
no lowly	positive	median	not lowly	positive	median	morosely	negative	median
no lowly	positive	median	not lowly	positive	median	nervously	negative	median
no low-spirited	positive	median	not low-spirited	positive	median	restlessly	negative	median
no loyal	negative	median	not loyal	negative	median	sadly	negative	median
no lucid	negative	median	not lucid	negative	median	somberly	negative	median
no luckless	positive	median	not luckless	positive	median	sorrowfully	negative	median
no lucky	negative	median	not lucky	negative	median	spiritlessly	negative	median
no ludicrous	positive	median	not ludicrous	positive	median	spookily	negative	median
no lunatic	positive	median	not lunatic	positive	median	stalely	negative	median
no lurid	positive	median	not lurid	positive	median	startledly	negative	median
no lush	negative	median	not lush	negative	median	sulkily	negative	median
no lying	positive	median	not lying	positive	median	sullenly	negative	median
no mad	positive	median	not mad	positive	median	surlily	negative	median
no madcap	positive	median	not madcap	positive	median	surprisedly	positive	high
no maddening	positive	median	not maddening	positive	median	clean	positive	median
no magical	negative	median	not magical	negative	median	clear	positive	median
no magnanimous	negative	high	not magnanimous	negative	high	clear	positive	median
no magnificent	negative	high	not magnificent	negative	high	detailed	positive	median
no magnificently	negative	high	not magnificently	negative	high	detailed	positive	median
no mainstream	negative	median	not mainstream	negative	median	elaborate	positive	median
no major	negative	median	not major	negative	median	elegant	positive	median
no malevolent	positive	median	not malevolent	positive	median	intricate	positive	median
no manic	negative	median	not manic	negative	median	lucid	positive	median
no manipulative	positive	median	not manipulative	positive	median	precise	positive	median
no marvelous	negative	median	not marvelous	negative	median	refined	positive	median
no marvelously	negative	median	not marvelously	negative	median	refined	positive	median
no masterful	negative	high	not masterful	negative	high	rich	positive	median
no masterfully	negative	high	not masterfully	negative	high	uncomplicated	positive	median
no masterly	negative	high	not masterly	negative	high	classic	positive	median
no matchless	negative	median	not matchless	negative	median	easy	positive	median
no mature	negative	median	not mature	negative	median	easier	positive	high
no mean	positive	median	not mean	positive	median	arcane	negative	median
no meaner	positive	high	not meaner	positive	high	byzantine	negative	high
no meanest	positive	max	not meanest	positive	max	convoluted	negative	high
no meaningful	negative	median	not meaningful	negative	median	excessive	negative	median
no meaningless	positive	median	not meaningless	positive	median	extravagant	negative	median
no meaty	negative	median	not meaty	negative	median	fancy	negative	median
no mediocre	positive	median	not mediocre	positive	median	monolithic	negative	median
no meek	negative	median	not meek	negative	median	ornate	positive	median
no melancholy	positive	median	not melancholy	positive	median	overelaborate	negative	median
no melodramatic	positive	median	not melodramatic	positive	median	oversimplified	negative	median
no memorable	negative	median	not memorable	negative	median	simplistic	negative	median
no menacing	positive	median	not menacing	positive	median	unclear	negative	median
no mendacious	positive	median	not mendacious	positive	median	woolly	negative	median
no mercenary	positive	median	not mercenary	positive	median	confusing	negative	median
no merciless	positive	high	not merciless	positive	high	unnecessary	negative	median
no merry	negative	median	not merry	negative	median	complex	negative	median
no messy	positive	median	not messy	positive	median	cleaner	positive	high
no methodical	negative	median	not methodical	negative	median	cleanest	positive	max
no meticulous	negative	median	not meticulous	negative	median	clearer	positive	high
no meticulously	negative	median	not meticulously	negative	median	richer	positive	high
no mightily	decrease	not mightily	decrease	purest	positive	max		
no mighty	negative	high	not mighty	negative	high	richest	positive	max
no mild	positive	median	not mild	positive	median	easiest	positive	max
no milder	positive	high	not milder	positive	high	classically	positive	median
no mildest	positive	max	not mildest	positive	max	cleanly	positive	median
no mildly	increase	not mildly	increase	easily	positive	median		
no mind-blowing	negative	median	not mind-blowing	negative	median	elaborately	positive	median
no mindless	positive	median	not mindless	positive	median	elegantly	positive	median
no minimal	increase	not minimal	increase	intricately	positive	median		
no minor	increase	not minor	increase	precisely	positive	median		
no mirthful	negative	median	not mirthful	negative	median	richly	positive	median
no mirthless	positive	median	not mirthless	positive	median	complexly	negative	median
no miserable	positive	median	not miserable	positive	median	simplistically	negative	median
no miserly	positive	median	not miserly	positive	median	coherent	positive	median
no misguided	positive	median	not misguided	positive	median	considered	positive	median
no misleading	positive	median	not misleading	positive	median	consistent	positive	median
no misshapen	positive	median	not misshapen	positive	median	consonant	positive	median
no model	negative	median	not model	negative	median	curvaceous	positive	median
no moderate	increase	not moderate	increase	harmonious	positive	median		
no moderately	increase	not moderately	increase	logical	positive	median		
no modern	negative	median	not modern	negative	median	proportional	positive	median
no modest	negative	median	not modest	negative	median	proportioned	positive	median
no modestly	negative	median	not modestly	negative	median	shapely	positive	median
no monolithic	positive	median	not monolithic	positive	median	symmetrical	positive	median
no monotonous	positive	high	not monotonous	positive	high	unified	positive	median
no monstrous	positive	median	not monstrous	positive	median	willowy	positive	median
no moody	positive	median	not moody	positive	median	proportionate	positive	median
no mopey	positive	median	not mopey	positive	median	symmetric	positive	median
no moral	negative	median	not moral	negative	median	perfect	positive	max
no morally	negative	median	not morally	negative	median	flawless	positive	max
no more	decrease	not more	decrease	neat	positive	median		
no moronic	positive	high	not moronic	positive	high	aberrant	negative	median
no moronically	positive	high	not moronically	positive	high	amorphous	negative	median
no morose	positive	median	not morose	positive	median	asymmetric	negative	median
no morosely	positive	median	not morosely	positive	median	asymmetrical	negative	median
no most	minimize	not most	minimize	contradictory	negative	median		
no mostly	decrease	not mostly	decrease	discordant	negative	high		
no mournful	positive	high	not mournful	positive	high	disorganised	negative	median
no moving	negative	median	not moving	negative	median	disproportionate	negative	median
no mundane	positive	median	not mundane	positive	median	distorted	negative	high
no murky	positive	median	not murky	positive	median	flawed	negative	median

no musically	negative	median	not musically	negative	median	formless	negative	median
no mystical	negative	median	not mystical	negative	median	misshapen	negative	median
no naive	positive	median	not naive	positive	median	nonsymmetrical	negative	median
no naively	positive	median	not naively	positive	median	shapeless	negative	median
no nameless	positive	median	not nameless	positive	median	unequal	negative	median
no nastier	positive	high	not nastier	positive	high	uneven	negative	median
no nastiest	positive	max	not nastiest	positive	max	ungraceful	negative	median
no nastily	positive	median	not nastily	positive	median	unharmonious	negative	median
no nasty	positive	median	not nasty	positive	median	unnatural	negative	median
no natural	negative	median	not natural	negative	median	unsymmetrical	negative	median
no neat	negative	median	not neat	negative	median	inharmonious	negative	median
no neatly	negative	median	not neatly	negative	median	inconsonant	negative	median
no needless	positive	median	not needless	positive	median	incongruous	negative	median
no nefarious	positive	high	not nefarious	positive	high	imbalanced	negative	median
no nervous	positive	median	not nervous	positive	median	dissonant	negative	high
no nervously	positive	median	not nervously	positive	median	incongruent	negative	median
no nervy	positive	median	not nervy	positive	median	unmusical	negative	median
no neurotic	positive	median	not neurotic	positive	median	awkward	negative	median
no nice	negative	median	not nice	negative	median	sloppy	negative	median
no nicely	negative	median	not nicely	negative	median	consistently	positive	median
no nicer	negative	high	not nicer	negative	high	evenly	positive	median
no nicest	negative	max	not nicest	negative	max	logically	positive	median
no noble	negative	high	not noble	negative	high	musically	positive	median
no nobler	negative	very high	not nobler	negative	very high	neatly	positive	median
no noblest	negative	max	not noblest	negative	max	proportionately	positive	median
no noisome	positive	median	not noisome	positive	median	symmetrically	positive	median
no noisy	positive	median	not noisy	positive	median	awkwardly	negative	median
no nondescript	positive	median	not nondescript	positive	median	shapelessly	negative	median
no nonsymmetrical	positive	median	not nonsymmetrical	positive	median	sloppily	negative	median
no normal	negative	median	not normal	negative	median	admirable	positive	median
no normally	negative	median	not normally	negative	median	allright	positive	median
no notable	negative	median	not notable	negative	median	attractive	positive	median
no notably	decrease	not notably	decrease	beauteous	positive	median		
no noted	negative	median	not noted	negative	median	beautiful	positive	median
no noteworthy	negative	median	not noteworthy	negative	median	bewitching	positive	high
no noteworthy	negative	median	not noteworthy	negative	median	brilliant	positive	high
no noticeable	negative	low	not noticeable	negative	low	captivating	positive	high
no notorious	positive	median	not notorious	positive	median	charming	positive	median
no novel	negative	median	not novel	negative	median	comely	positive	median
no numskulled	positive	high	not numskulled	positive	high	congenial	positive	median
no obliging	negative	median	not obliging	negative	median	delightful	positive	median
no obligingly	negative	median	not obligingly	negative	median	dishy	positive	median
no oblique	positive	median	not oblique	positive	median	enchanting	positive	high
no obnoxious	positive	median	not obnoxious	positive	median	excellent	positive	high
no obscene	positive	median	not obscene	positive	median	excellent	positive	high
no obscure	positive	median	not obscure	positive	median	exceptional	positive	very high
no obscurely	positive	median	not obscurely	positive	median	exquisite	positive	high
no obstinate	positive	median	not obstinate	positive	median	favorable	positive	median
no obtuse	positive	median	not obtuse	positive	median	fine	positive	median
no obvious	negative	median	not obvious	negative	median	glorious	positive	high
no occupied	negative	median	not occupied	negative	median	gorgeous	positive	median
no odd	positive	median	not odd	positive	median	grand	positive	median
no oddball	positive	median	not oddball	positive	median	gratifying	positive	median
no oddly	positive	median	not oddly	positive	median	handsome	positive	median
no offbeat	negative	low	not offbeat	negative	low	lovely	positive	median
no offensive	positive	median	not offensive	positive	median	lovely	positive	median
no off-putting	positive	median	not off-putting	positive	median	magnificent	positive	high
no ok	negative	median	not ok	negative	median	marvelous	positive	median
no okay	negative	median	not okay	negative	median	nice	positive	median
no open	negative	median	not open	negative	median	ok	positive	median
no openly	negative	median	not openly	negative	median	okay	positive	median
no operational	negative	low	not operational	negative	low	pleasing	positive	median
no oppressive	positive	median	not oppressive	positive	median	pleasing	positive	median
no optimal	negative	median	not optimal	negative	median	pleasing	positive	median
no optimistic	negative	median	not optimistic	negative	median	radiant	positive	median
no optimum	negative	max	not optimum	negative	max	ravishing	positive	high
no ordinary	positive	median	not ordinary	positive	median	resplendent	positive	high
no original	negative	median	not original	negative	median	splendid	positive	median
no originally	negative	median	not originally	negative	median	splendiferous	positive	high
no originative	negative	median	not originative	negative	median	sterling	positive	median
no ornate	negative	median	not ornate	negative	median	stunning	positive	median
no outdated	positive	median	not outdated	positive	median	superb	positive	high
no outlandish	negative	high	not outlandish	negative	high	welcome	positive	median
no outrageous	positive	median	not outrageous	positive	median	wonderful	positive	median
no outrageously	positive	high	not outrageously	positive	high	tasty	positive	median
no outstanding	negative	median	not outstanding	negative	median	subtle	positive	median
no outstandingly	negative	high	not outstandingly	negative	high	rare	positive	median
no outwardly	positive	median	not outwardly	positive	median	fresh	positive	median
no overdue	positive	median	not overdue	positive	median	fun	positive	median
no overelaborate	positive	median	not overelaborate	positive	median	abhorrent	negative	median
no overjoyed	negative	high	not overjoyed	negative	high	abominable	negative	median
no oversimplified	positive	median	not oversimplified	positive	median	beastly	negative	median
no painful	positive	median	not painful	positive	median	despicable	negative	median
no painless	negative	median	not painless	negative	median	disagreeable	negative	median
no paltry	positive	median	not paltry	positive	median	disgusting	negative	median
no panicky	positive	median	not panicky	positive	median	distasteful	negative	median
no paranoid	positive	median	not paranoid	positive	median	filthy	negative	median
no parsimonious	positive	median	not parsimonious	positive	median	forbidding	negative	median
no partial	positive	median	not partial	positive	median	frightful	negative	median
no partially	increase	not partially	increase	grody	negative	median		
no partly	increase	not partly	increase	gross	negative	median		
no passably	increase	not passably	increase	grotesque	negative	high		
no pathetic	positive	median	not pathetic	positive	median	hideous	negative	high
no patient	negative	median	not patient	negative	median	homely	negative	median
no patiently	negative	median	not patiently	negative	median	horrible	negative	high

no peculiar	positive	median	not peculiar	positive	median	horrid	negative	high
no peculiarly	positive	median	not peculiarly	positive	median	loathsome	negative	high
no peculiarly	decrease	not peculiarly	decrease	monstrous	negative	median		
no pedestrian	positive	median	not pedestrian	positive	median	nasty	negative	median
no peevish	positive	median	not peevish	positive	median	nastier	negative	high
no penetrating	negative	median	not penetrating	negative	median	noisome	negative	median
no peppy	negative	high	not peppy	negative	high	obnoxious	negative	median
no perfect	negative	max	not perfect	negative	max	repellent	negative	median
no perfectly	minimize	not perfectly	minimize	repugnant	negative	high		
no perky	negative	median	not perky	negative	median	repulsive	negative	median
no persevering	negative	median	not persevering	negative	median	revolting	negative	high
no perturbed	positive	low	not perturbed	positive	low	sickening	negative	high
no pessimistic	positive	median	not pessimistic	positive	median	sleazy	negative	median
no petty	positive	median	not petty	positive	median	ugly	negative	median
no phenomenal	negative	very high	not phenomenal	negative	very high	uncomely	negative	median
no philanthropic	negative	median	not philanthropic	negative	median	undecorated	negative	median
no phoney	positive	high	not phoney	positive	high	unlovely	negative	median
no phony	positive	high	not phony	positive	high	unseemly	negative	median
no plain	negative	median	not plain	negative	median	unsightly	negative	median
no plainer	negative	high	not plainer	negative	high	wrong	negative	median
no plainest	negative	max	not plainest	negative	max	yucky	negative	median
no plainly	negative	median	not plainly	negative	median	yuk	negative	median
no plainspoken	negative	median	not plainspoken	negative	median	finer	positive	high
no pleasant	negative	median	not pleasant	negative	median	grander	positive	high
no pleasantly	negative	median	not pleasantly	negative	median	nicer	positive	high
no pleased	negative	median	not pleased	negative	median	finest	positive	max
no pleasing	negative	median	not pleasing	negative	median	nicest	positive	max
no pleasing	negative	median	not pleasing	negative	median	nastiest	negative	max
no pleasing	negative	median	not pleasing	negative	median	attractively	positive	median
no pleasingly	negative	median	not pleasingly	negative	median	beautifully	increase	
no plentiful	negative	median	not plentiful	negative	median	brilliantly	positive	high
no plucky	negative	median	not plucky	negative	median	charmingly	positive	median
no poetic	negative	median	not poetic	negative	median	delightfully	positive	median
no poignant	negative	median	not poignant	negative	median	excellently	positive	high
no pointless	positive	median	not pointless	positive	median	exquisitely	positive	high
no polite	negative	median	not polite	negative	median	finely	positive	median
no politely	negative	median	not politely	negative	median	freshly	positive	median
no poor	positive	median	not poor	positive	median	gloriously	positive	high
no poorer	positive	high	not poorer	positive	high	goodly	positive	median
no poorly	positive	median	not poorly	positive	median	gorgeously	positive	median
no popular	negative	median	not popular	negative	median	grandiosely	positive	high
no popularly	negative	median	not popularly	negative	median	handsomely	positive	median
no positive	negative	median	not positive	negative	median	magnificently	positive	high
no positively	decrease	not positively	decrease	marvelously	positive	median		
no potent	negative	median	not potent	negative	median	nicely	positive	median
no powerful	negative	median	not powerful	negative	median	pleasingly	positive	median
no powerfully	negative	median	not powerfully	negative	median	prettily	positive	median
no praiseworthy	negative	high	not praiseworthy	negative	high	resplendently	positive	high
no precious	negative	median	not precious	negative	median	splendidly	positive	median
no precise	negative	median	not precise	negative	median	splendiferously	positive	high
no precisely	negative	median	not precisely	negative	median	sterlingly	positive	median
no predictable	negative	median	not predictable	negative	median	stunningly	positive	median
no predictably	negative	median	not predictably	negative	median	subtlely	positive	median
no premium	negative	median	not premium	negative	median	superbly	positive	high
no presentable	negative	median	not presentable	negative	median	wonderfully	positive	median
no pretentious	positive	median	not pretentious	positive	median	frightfully	negative	median
no prettily	negative	median	not prettily	negative	median	grossly	negative	median
no pretty	negative	median	not pretty	negative	median	grotesquely	negative	high
no pretty	decrease	not pretty	decrease	hideously	negative	median		
no priceless	negative	very high	not priceless	negative	very high	horribly	increase	
no pricelessly	negative	high	not pricelessly	negative	high	nastily	negative	median
no pricey	positive	median	not pricey	positive	median	repellently	negative	median
no prime	negative	median	not prime	negative	median	repugnantly	negative	high
no principled	negative	median	not principled	negative	median	repulsively	negative	median
no prized	negative	median	not prized	negative	median	revoltingly	negative	high
no prodigiously	minimize	not prodigiously	minimize	sleazily	negative	median		
no productive	negative	median	not productive	negative	median	wrongly	negative	median
no professional	negative	median	not professional	negative	median	absorbing	positive	median
no professionally	negative	median	not professionally	negative	median	alluring	positive	high
no proficient	negative	median	not proficient	negative	median	amazing	positive	median
no profound	negative	median	not profound	negative	median	appealing	positive	median
no profoundly	minimize	not profoundly	minimize	arousing	positive	median		
no prominent	negative	high	not prominent	negative	high	arresting	positive	median
no prominently	negative	high	not prominently	negative	high	arrestive	positive	median
no promising	negative	median	not promising	negative	median	breathtaking	positive	high
no prompt	negative	median	not prompt	negative	median	compelling	positive	median
no proper	negative	median	not proper	negative	median	dramatic	positive	high
no properly	negative	median	not properly	negative	median	dramatic	positive	median
no propitious	negative	median	not propitious	negative	median	electrifying	positive	very high
no proportional	negative	median	not proportional	negative	median	engaging	positive	median
no proportionate	negative	median	not proportionate	negative	median	engrossing	positive	high
no proportionately	negative	median	not proportionately	negative	median	enthralling	positive	high
no proportioned	negative	median	not proportioned	negative	median	entrancing	positive	high
no prosaic	positive	median	not prosaic	positive	median	exciting	positive	median
no proud	negative	median	not proud	negative	median	fascinating	positive	median
no provocative	negative	high	not provocative	negative	high	gripping	positive	median
no provoking	negative	median	not provoking	negative	median	impressive	positive	median
no puerile	positive	median	not puerile	positive	median	incredible	positive	median
no pure	negative	median	not pure	negative	median	inspiring	positive	median
no purely	negative	median	not purely	negative	median	intense	positive	median
no purer	negative	high	not purer	negative	high	interesting	positive	median
no purest	negative	max	not purest	negative	max	intoxicating	positive	high
no qualified	negative	low	not qualified	negative	low	intriguing	positive	high
no quality	negative	median	not quality	negative	median	moving	positive	median
no queer	positive	high	not queer	positive	high	noteworthy	positive	median

no questionable	positive	median	not questionable	positive	median	noteworthy	positive	median
no quick	negative	median	not quick	negative	median	provocative	positive	high
no quickest	negative	max	not quickest	negative	max	provoking	positive	median
no quite	decrease	not quite	decrease	remarkable	positive	median		
no radiant	negative	median	not radiant	negative	median	riveting	positive	very high
no raging	positive	median	not raging	positive	median	rousing	positive	median
no rapt	negative	high	not rapt	negative	high	sensational	positive	high
no rare	negative	median	not rare	negative	median	spectacular	positive	high
no rash	positive	median	not rash	positive	median	staggering	positive	high
no rather	decrease	not rather	decrease	startling	positive	median		
no ravishing	negative	high	not ravishing	negative	high	stimulating	positive	median
no readable	negative	median	not readable	negative	median	stirring	positive	median
no real	negative	median	not real	negative	median	striking	positive	median
no realistic	negative	median	not realistic	negative	median	surprising	positive	median
no realistically	negative	median	not realistically	negative	median	thrilling	positive	high
no really	increase	not really	increase	titillating	positive	very high		
no reasonable	negative	low	not reasonable	negative	low	vivid	positive	high
no reasonably	decrease	not reasonably	decrease	stimulative	positive	median		
no reckless	positive	high	not reckless	positive	high	special	positive	median
no recklessly	positive	high	not recklessly	positive	high	funny	positive	median
no reductive	positive	median	not reductive	positive	median	funnier	positive	high
no reductively	positive	median	not reductively	positive	median	comic	positive	median
no redundant	positive	median	not redundant	positive	median	hilarious	positive	high
no refined	negative	median	not refined	negative	median	amusing	positive	median
no refined	negative	median	not refined	negative	median	cute	positive	median
no refreshing	negative	median	not refreshing	negative	median	cuter	positive	high
no refreshingly	negative	median	not refreshingly	negative	median	sexy	positive	median
no relatively	increase	not relatively	increase	terrific	positive	median		
no relevant	negative	median	not relevant	negative	median	fantastic	positive	median
no reliable	negative	median	not reliable	negative	median	satisfying	positive	median
no reliable	negative	median	not reliable	negative	median	outstanding	positive	median
no reliably	negative	median	not reliably	negative	median	refreshing	positive	median
no religious	negative	median	not religious	negative	median	stylish	positive	median
no reluctant	positive	median	not reluctant	positive	median	emotional	positive	median
no remarkable	negative	median	not remarkable	negative	median	catchy	positive	median
no remarkably	decrease	not remarkably	decrease	poetic	positive	median		
no renowned	negative	median	not renowned	negative	median	dynamic	positive	median
no renownedly	negative	high	not renownedly	negative	high	solid	positive	median
no repellent	positive	median	not repellent	positive	median	convincing	positive	median
no repellently	positive	median	not repellently	positive	median	unexpected	positive	median
no repetitive	positive	median	not repetitive	positive	median	haunting	positive	median
no repetitively	positive	median	not repetitively	positive	median	colorful	positive	median
no reprobate	positive	median	not reprobate	positive	median	colourful	positive	median
no reprobately	positive	median	not reprobately	positive	median	artistic	positive	median
no repugnant	positive	high	not repugnant	positive	high	super	positive	median
no repugnantly	positive	high	not repugnantly	positive	high	poignant	positive	median
no repulsive	positive	median	not repulsive	positive	median	explosive	positive	median
no repulsively	positive	median	not repulsively	positive	median	endearing	positive	median
no reputable	negative	median	not reputable	negative	median	astonishing	positive	median
no resistant	positive	median	not resistant	positive	median	hysterical	positive	median
no resolute	negative	high	not resolute	negative	high	ascetic	negative	median
no resolutely	negative	median	not resolutely	negative	median	blah	negative	median
no respectable	negative	median	not respectable	negative	median	dry	negative	median
no respectably	negative	median	not respectably	negative	median	flat	negative	median
no respectful	negative	median	not respectful	negative	median	hohum	negative	median
no respectfully	negative	median	not respectfully	negative	median	humdrum	negative	median
no resplendent	negative	high	not resplendent	negative	high	insipid	negative	median
no resplendently	negative	high	not resplendently	negative	high	irksome	negative	median
no responsible	negative	median	not responsible	negative	median	lifeless	negative	median
no restless	positive	median	not restless	positive	median	monotonous	negative	high
no restlessly	positive	median	not restlessly	positive	median	pedestrian	negative	median
no retrograde	positive	median	not retrograde	positive	median	prosaic	negative	median
no reverent	negative	median	not reverent	negative	median	routine	negative	median
no reverentially	negative	median	not reverentially	negative	median	soporific	negative	median
no revolting	positive	high	not revolting	positive	high	tedious	negative	median
no revoltingly	positive	high	not revoltingly	positive	high	tedious	negative	median
no revolutionary	negative	median	not revolutionary	negative	median	tiresome	negative	median
no rich	negative	median	not rich	negative	median	tiring	negative	median
no richer	negative	high	not richer	negative	high	unexciting	negative	median
no richest	negative	max	not richest	negative	max	unimaginative	negative	median
no richly	negative	median	not richly	negative	median	uninspiring	negative	median
no ridiculous	positive	high	not ridiculous	positive	high	uninteresting	negative	median
no ridiculously	positive	high	not ridiculously	positive	high	uninviting	negative	median
no right	negative	median	not right	negative	median	unremarkable	negative	median
no righteous	negative	high	not righteous	negative	high	vapid	negative	high
no righteously	negative	high	not righteously	negative	high	wearisome	negative	median
no rightfully	negative	median	not rightfully	negative	median	poor	negative	median
no rightly	negative	median	not rightly	negative	median	annoying	negative	high
no riveting	negative	very high	not riveting	negative	very high	ridiculous	negative	high
no robust	negative	median	not robust	negative	median	bland	negative	median
no roguish	positive	median	not roguish	positive	median	offensive	negative	median
no romantic	negative	median	not romantic	negative	median	creepy	negative	high
no rotten	positive	high	not rotten	positive	high	irritating	negative	high
no rough	positive	median	not rough	positive	median	disappointing	negative	median
no rougher	positive	high	not rougher	positive	high	lame	negative	median
no roughest	positive	max	not roughest	positive	max	unfunny	negative	median
no roughly	positive	median	not roughly	positive	median	pretentious	negative	median
no roundabout	positive	median	not roundabout	positive	median	formulaic	negative	median
no rousing	negative	median	not rousing	negative	median	corny	negative	high
no routine	positive	median	not routine	positive	median	inept	negative	median
no routinely	positive	median	not routinely	positive	median	silly	negative	median
no rude	positive	median	not rude	positive	median	insulting	negative	high
no rudely	positive	median	not rudely	positive	median	incomprehensible	negative	median
no rugged	negative	median	not rugged	negative	median	funniest	positive	max
no run-of-the-mill	positive	median	not run-of-the-mill	positive	median	cutest	positive	max

no ruthless	positive	median	not ruthless	positive	median	silliest	negative	max
no sad	positive	median	not sad	positive	median	corniest	negative	max
no sadistic	positive	median	not sadistic	positive	median	amusingly	positive	median
no sadly	positive	median	not sadly	positive	median	artistically	positive	median
no safe	negative	median	not safe	negative	median	breathtakingly	positive	high
no safely	negative	median	not safely	negative	median	comically	positive	median
no safer	negative	high	not safer	negative	high	convincingly	positive	median
no safest	negative	max	not safest	negative	max	dramatically	positive	median
no sagacious	negative	high	not sagacious	negative	high	emotionally	positive	median
no sagaciously	negative	high	not sagaciously	negative	high	engagingly	positive	median
no sane	negative	median	not sane	negative	median	fantastically	positive	median
no sanely	negative	median	not sanely	negative	median	funnily	positive	median
no saner	negative	high	not saner	negative	high	hilariously	positive	high
no sanguine	negative	median	not sanguine	negative	median	impressively	positive	median
no sanguinely	negative	median	not sanguinely	negative	median	interestingly	positive	median
no sappily	positive	median	not sappily	positive	median	intriguingly	positive	high
no sappy	positive	median	not sappy	positive	median	refreshingly	positive	median
no satisfactory	negative	low	not satisfactory	negative	low	satisfyingly	positive	median
no satisfied	negative	high	not satisfied	negative	high	sensationally	positive	high
no satisfiedly	negative	median	not satisfiedly	negative	median	sexily	positive	median
no satisfying	negative	median	not satisfying	negative	median	solidly	positive	median
no satisfyingly	negative	median	not satisfyingly	negative	median	spectacularly	positive	high
no savage	positive	median	not savage	positive	median	startlingly	positive	median
no savagely	positive	median	not savagely	positive	median	stimulatingly	positive	median
no scarcely	increase	not scarcely	increase	stimulatively	positive	median		
no scared	positive	median	not scared	positive	median	stirringly	positive	median
no scenic	negative	median	not scenic	negative	median	stylishly	positive	median
no scheming	positive	high	not scheming	positive	high	unexpectedly	positive	median
no schemingly	positive	high	not schemingly	positive	high	vividly	positive	high
no scrupulous	negative	median	not scrupulous	negative	median	annoyingly	negative	high
no scrupulous	negative	median	not scrupulous	negative	median	blandly	negative	median
no second-rate	positive	very high	not second-rate	positive	very high	creepily	negative	high
no secure	negative	median	not secure	negative	median	flatly	negative	median
no securely	negative	median	not securely	negative	median	ineptly	negative	median
no self-assured	negative	median	not self-assured	negative	median	poorly	negative	median
no self-assuredly	negative	median	not self-assuredly	negative	median	ridiculously	negative	high
no self-confident	negative	median	not self-confident	negative	median	routinely	negative	median
no self-confidently	negative	median	not self-confidently	negative	median	sillily	negative	median
no self-effacing	negative	high	not self-effacing	negative	high	soporifically	negative	median
no selfish	positive	median	not selfish	positive	median	tediously	negative	median
no selfishly	positive	median	not selfishly	positive	median	appropriate	positive	median
no self-righteous	positive	high	not self-righteous	positive	high	authentic	positive	median
no self-willed	positive	median	not self-willed	positive	median	bonafide	positive	median
no sensational	negative	high	not sensational	negative	high	creative	positive	median
no sensationally	negative	high	not sensationally	negative	high	deep	positive	median
no sensible	negative	median	not sensible	negative	median	effective	positive	median
no sensibly	negative	median	not sensibly	negative	median	important	positive	median
no sensitive	negative	median	not sensitive	negative	median	incisive	positive	high
no sensitively	negative	median	not sensitively	negative	median	indubitable	positive	high
no serious	negative	median	not serious	negative	median	inestimable	positive	high
no serviceable	negative	median	not serviceable	negative	median	ingenious	positive	high
no severe	decrease	not severe	decrease	inimitable	positive	median		
no sexily	negative	median	not sexily	negative	median	innovational	positive	median
no sexy	negative	median	not sexy	negative	median	innovative	positive	high
no shabby	positive	median	not shabby	positive	median	invaluable	positive	high
no shadily	positive	median	not shadily	positive	median	inventive	positive	median
no shady	positive	median	not shady	positive	median	landmark	positive	high
no shakily	positive	median	not shakily	positive	median	longawaited	positive	median
no shaky	positive	median	not shaky	positive	median	matchless	positive	median
no shallow	positive	high	not shallow	positive	high	original	positive	median
no shallower	positive	very high	not shallower	positive	very high	originative	positive	median
no shallowly	positive	high	not shallowly	positive	high	penetrating	positive	median
no sham	positive	high	not sham	positive	high	phenomenal	positive	very high
no shapeless	positive	median	not shapeless	positive	median	priceless	positive	very high
no shapelessly	positive	median	not shapelessly	positive	median	prized	positive	median
no shapely	negative	median	not shapely	negative	median	profound	positive	median
no sharp	negative	median	not sharp	negative	median	real	positive	median
no sharper	negative	high	not sharper	negative	high	singular	positive	median
no sharpest	negative	max	not sharpest	negative	max	timely	positive	median
no sharply	negative	median	not sharply	negative	median	unequalled	positive	median
no shiftily	positive	median	not shiftily	positive	median	unique	positive	median
no shifty	positive	median	not shifty	positive	median	unparalleled	positive	high
no shocking	positive	median	not shocking	positive	median	unprecedented	positive	high
no shoddily	positive	high	not shoddily	positive	high	useful	positive	median
no shoddy	positive	high	not shoddy	positive	high	valid	positive	median
no shrewd	negative	median	not shrewd	negative	median	valuable	positive	median
no shrewdly	negative	median	not shrewdly	negative	median	valued	positive	median
no sick	positive	high	not sick	positive	high	veritable	positive	median
no sickening	positive	high	not sickening	positive	high	worthwhile	positive	median
no sickeningly	positive	high	not sickeningly	positive	high	quick	positive	median
no sicker	positive	high	not sicker	positive	high	superior	positive	very high
no sickly	positive	median	not sickly	positive	median	promising	positive	median
no significant	negative	median	not significant	negative	median	significant	positive	median
no significantly	decrease	not significantly	decrease	extraordinary	positive	very high		
no silliest	positive	max	not silliest	positive	max	awesome	positive	high
no sillily	positive	median	not sillily	positive	median	tremendous	positive	high
no silly	positive	median	not silly	positive	median	idea	positive	median
no simple	positive	median	not simple	positive	median	gentle	positive	median
no simple-minded	positive	median	not simple-minded	positive	median	bogus	negative	median
no simpler	positive	high	not simpler	positive	high	coarse	negative	median
no simplest	positive	max	not simplest	positive	max	common	negative	median
no simplistic	positive	median	not simplistic	positive	median	commonplace	negative	median
no simplistically	positive	median	not simplistically	positive	median	conventional	negative	median
no sincere	negative	median	not sincere	negative	median	derivative	negative	median
no sincerely	negative	median	not sincerely	negative	median	everyday	negative	median

no sincerest	negative	max	not sincerest	negative	max	fake	negative	median
no sinful	positive	median	not sinful	positive	median	flimsy	negative	median
no sinfully	positive	median	not sinfully	positive	median	glitzy	negative	median
no singular	negative	median	not singular	negative	median	hackneyed	negative	median
no singularly	negative	median	not singularly	negative	median	ineffective	negative	median
no sinister	positive	median	not sinister	positive	median	ineffectual	negative	median
no skilled	negative	median	not skilled	negative	median	inferior	negative	median
no skillful	negative	median	not skillful	negative	median	insignificant	negative	median
no skillfully	negative	median	not skillfully	negative	median	meaningless	negative	median
no slack	positive	median	not slack	positive	median	mundane	negative	median
no slackly	positive	median	not slackly	positive	median	nondescript	negative	median
no sleazily	positive	median	not sleazily	positive	median	overdue	negative	median
no sleazy	positive	median	not sleazy	positive	median	paltry	negative	median
no sleek	negative	median	not sleek	negative	median	phoney	negative	high
no slick	positive	median	not slick	positive	median	phony	negative	high
no slickly	positive	median	not slickly	positive	median	pointless	negative	median
no slight	increase	not slight	increase	pricey	negative	median		
no slightly	increase	not slightly	increase	reductive	negative	median		
no slippery	positive	median	not slippery	positive	median	shallow	negative	high
no sloppily	positive	median	not sloppily	positive	median	sham	negative	high
no sloppy	positive	median	not sloppy	positive	median	shoddy	negative	high
no slow	positive	median	not slow	positive	median	stereotyped	negative	median
no slow	positive	median	not slow	positive	median	superficial	negative	median
no slower	positive	high	not slower	positive	high	trashy	negative	high
no slowest	positive	max	not slowest	positive	max	trifling	negative	median
no slowly	positive	median	not slowly	positive	median	trite	negative	median
no sluggish	positive	median	not sluggish	positive	median	trivial	negative	median
no sluggishly	positive	median	not sluggishly	positive	median	undistinguished	negative	median
no sly	positive	median	not sly	positive	median	unexceptional	negative	median
no slyly	positive	median	not slyly	positive	median	unimportant	negative	median
no smart	negative	median	not smart	negative	median	untimely	negative	median
no smartly	negative	median	not smartly	negative	median	unusable	negative	median
no sneakily	positive	median	not sneakily	positive	median	useless	negative	high
no sneaky	positive	median	not sneaky	positive	median	valueless	negative	high
no snobbily	positive	median	not snobbily	positive	median	worthless	negative	high
no snobby	positive	median	not snobby	positive	median	cheap	positive	median
no social	positive	median	not social	positive	median	average	negative	median
no soft	positive	median	not soft	positive	median	sorry	negative	median
no softer	positive	high	not softer	positive	high	mediocre	negative	median
no softly	positive	median	not softly	positive	median	gratuitous	negative	median
no solid	negative	median	not solid	negative	median	repetitive	negative	median
no solidly	negative	median	not solidly	negative	median	flashy	negative	median
no somber	positive	median	not somber	positive	median	absurd	negative	high
no somberly	positive	median	not somberly	positive	median	sappy	negative	median
no somewhat	increase	not somewhat	increase	deeper	positive	high		
no sophisticated	negative	median	not sophisticated	negative	median	commoner	negative	high
no sophisticatedly	negative	median	not sophisticatedly	negative	median	shallower	negative	very high
no soporific	positive	median	not soporific	positive	median	deepest	positive	max
no soporifically	positive	median	not soporifically	positive	median	quickest	positive	max
no sorriest	positive	max	not sorriest	positive	max	sorriest	negative	max
no sorrily	positive	median	not sorrily	positive	median	appropriately	positive	median
no sorrowful	positive	median	not sorrowful	positive	median	creatively	positive	median
no sorrowfully	positive	median	not sorrowfully	positive	median	deeply	positive	median
no sorry	positive	median	not sorry	positive	median	effectively	positive	median
no soundly	negative	median	not soundly	negative	median	ideally	positive	median
no special	negative	median	not special	negative	median	importantly	positive	median
no specially	negative	median	not specially	negative	median	ingeniously	positive	high
no spectacular	negative	high	not spectacular	negative	high	inventively	positive	median
no spectacularly	negative	high	not spectacularly	negative	high	originally	positive	median
no speedy	negative	median	not speedy	negative	median	pricelessly	positive	high
no spineless	positive	high	not spineless	positive	high	realistically	positive	median
no spinelessly	positive	high	not spinelessly	positive	high	singularly	positive	median
no spiritless	positive	median	not spiritless	positive	median	uniquely	positive	median
no spiritlessly	positive	median	not spiritlessly	positive	median	usefully	positive	median
no splendid	negative	median	not splendid	negative	median	absurdly	negative	high
no splendidly	negative	median	not splendidly	negative	median	cheaply	positive	median
no splendiferous	negative	high	not splendiferous	negative	high	coarsely	negative	median
no splendiferously	negative	high	not splendiferously	negative	high	commonly	negative	median
no spoiled	positive	median	not spoiled	positive	median	conventionally	negative	median
no spooked	positive	median	not spooked	positive	median	flimsily	negative	median
no spookily	positive	median	not spookily	positive	median	gratuitously	negative	median
no stable	negative	median	not stable	negative	median	insignificantly	negative	median
no stably	negative	median	not stably	negative	median	reductively	negative	median
no staggering	negative	high	not staggering	negative	high	repetitively	negative	median
no staggeringly	minimize	not staggeringly	minimize	sappily	negative	median		
no stale	positive	median	not stale	positive	median	shallowly	negative	high
no stalely	positive	median	not stalely	positive	median	shoddily	negative	high
no stalwart	negative	median	not stalwart	negative	median	sorrily	negative	median
no stalwartly	negative	median	not stalwartly	negative	median	superficially	negative	median
no star-crossed	positive	median	not star-crossed	positive	median	uselessly	negative	high
no startled	positive	median	not startled	positive	median	true-blue	positive	median
no startledly	positive	median	not startledly	positive	median	fly-by-night	negative	median
no startling	negative	median	not startling	negative	median	self-willed	negative	median
no startlingly	negative	median	not startlingly	negative	median	well-known	positive	median
no staunch	negative	high	not staunch	negative	high	also-ran	negative	median
no staunchly	negative	high	not staunchly	negative	high	has-been	negative	median
no steadfast	negative	high	not steadfast	negative	high	ill-fated	negative	median
no steadfastly	negative	high	not steadfastly	negative	high	ill-starred	negative	median
no steadily	negative	median	not steadily	negative	median	star-crossed	negative	median
no steady	negative	median	not steady	negative	median	unheard-of	negative	median
no stereotyped	positive	median	not stereotyped	positive	median	dim-witted	negative	high
no sterling	negative	median	not sterling	negative	median	empty-headed	negative	high
no sterlingly	negative	median	not sterlingly	negative	median	half-baked	negative	median
no stimulating	negative	median	not stimulating	negative	median	half-witted	negative	median
no stimulatingly	negative	median	not stimulatingly	negative	median	know-nothing	negative	high

no stimulative	negative	median	not stimulative	negative	median	simple-minded	negative	median
no stimulatively	negative	median	not stimulatively	negative	median	law-abiding	positive	median
no stirring	negative	median	not stirring	negative	median	self-righteous	negative	high
no stirringly	negative	median	not stirringly	negative	median	self-effacing	positive	high
no straight	negative	median	not straight	negative	median	well-mannered	positive	median
no straightforward	negative	median	not straightforward	negative	median	ill-mannered	negative	median
no straightforwardly	negative	median	not straightforwardly	negative	median	double-dealing	negative	high
no straightly	negative	median	not straightly	negative	median	two-faced	negative	high
no strange	positive	median	not strange	positive	median	self-assured	positive	median
no strangely	positive	median	not strangely	positive	median	well-off	positive	median
no stranger	positive	high	not stranger	positive	high	self-confident	positive	median
no strangest	positive	max	not strangest	positive	max	low-spirited	negative	median
no streamlined	negative	median	not streamlined	negative	median	ill-tempered	negative	median
no striking	negative	median	not striking	negative	median	heavy-hearted	negative	median
no strikingly	negative	median	not strikingly	negative	median	high-strung	negative	median
no strong	negative	median	not strong	negative	median	self-assuredly	positive	median
no strongly	decrease	not strongly	decrease	self-confidently	positive	median		
no stubborn	positive	median	not stubborn	positive	median	well-formed	positive	median
no stubbornly	positive	median	not stubbornly	positive	median	well-proportioned	positive	median
no stunning	negative	median	not stunning	negative	median	first-class	positive	median
no stunningly	negative	median	not stunningly	negative	median	first-rate	positive	median
no stupid	positive	median	not stupid	positive	median	good-looking	positive	median
no stupider	positive	high	not stupider	positive	high	god-awful	negative	high
no stupidest	positive	max	not stupidest	positive	max	off-putting	negative	median
no stupidly	positive	median	not stupidly	positive	median	hair-raising	positive	median
no sturdy	negative	median	not sturdy	negative	median	mind-blowing	positive	median
no stylish	negative	median	not stylish	negative	median	ho-hum	negative	median
no stylishly	negative	median	not stylishly	negative	median	dime-a-dozen	negative	median
no substantially	decrease	not substantially	decrease	good-for-nothing	negative	very high		
no subtle	negative	median	not subtle	negative	median	run-of-the-mill	negative	median
no subtlely	negative	median	not subtlely	negative	median	second-rate	negative	very high
no successful	negative	median	not successful	negative	median	write-off	negative	median
no successful	negative	median	not successful	negative	median	dumbest	negative	max
no successfully	negative	median	not successfully	negative	median	best	positive	max
no such	decrease	not such	decrease	better	positive	high		
no suitable	negative	low	not suitable	negative	low	high	increase	
no sulkily	positive	median	not sulkily	positive	median	big	positive	median
no sulky	positive	median	not sulky	positive	median	hard	negative	median
no sullen	positive	median	not sullen	positive	median	less	decrease	
no sullenly	positive	median	not sullenly	positive	median	basic	positive	median
no super	negative	median	not super	negative	median	huge	increase	
no superb	negative	high	not superb	negative	high	handy	positive	median
no superbly	negative	high	not superbly	negative	high	compatible	positive	median
no superficial	positive	median	not superficial	positive	median	accurate	positive	median
no superficially	positive	median	not superficially	positive	median	favorite	positive	max
no superior	negative	very high	not superior	negative	very high	adequate	positive	low
no supremely	minimize	not supremely	minimize	durable	positive	median		
no surlily	positive	median	not surlily	positive	median	bulky	negative	median
no surly	positive	median	not surly	positive	median	acceptable	positive	low
no surpassingly	minimize	not surpassingly	minimize	intuitive	positive	median		
no surprised	negative	high	not surprised	negative	high	hot	positive	median
no surprisedly	negative	high	not surprisedly	negative	high	ideal	positive	max
no surprising	negative	median	not surprising	negative	median	absolute	increase	
no surprisingly	decrease	not surprisingly	decrease	noticeable	positive	low		
no suspicious	positive	median	not suspicious	positive	median	extensive	positive	median
no symmetric	negative	median	not symmetric	negative	median	frustrating	negative	median
no symmetrical	negative	median	not symmetrical	negative	median	extreme	positive	median
no symmetrically	negative	median	not symmetrically	negative	median	minimal	decrease	
no sympathetic	negative	median	not sympathetic	negative	median	defective	negative	median
no sympathetically	negative	median	not sympathetically	negative	median	brief	positive	median
no tactful	negative	median	not tactful	negative	median	efficient	positive	median
no talented	negative	median	not talented	negative	median	needless	negative	median
no tasty	negative	median	not tasty	negative	median	enjoyable	positive	median
no technical	positive	median	not technical	positive	median	loose	negative	median
no tedious	positive	median	not tedious	positive	median	critical	positive	median
no tedious	positive	median	not tedious	positive	median	fabulous	positive	high
no tediously	positive	median	not tediously	positive	median	functional	positive	low
no tender	negative	median	not tender	negative	median	cumbersome	negative	median
no tenderly	negative	median	not tenderly	negative	median	memorable	positive	median
no tense	positive	median	not tense	positive	median	central	positive	median
no terrible	positive	high	not terrible	positive	high	disturbing	negative	median
no terrible	positive	very high	not terrible	positive	very high	essential	positive	high
no terribly	minimize	not terribly	minimize	frustrated	negative	median		
no terrific	negative	median	not terrific	negative	median	cinematic	positive	median
no terrifically	minimize	not terrifically	minimize	mad	negative	median		
no terrified	positive	high	not terrified	positive	high	optimal	positive	median
no terrorised	positive	high	not terrorised	positive	high	lousy	negative	median
no terrorized	positive	high	not terrorized	positive	high	fragile	negative	median
no thick	positive	median	not thick	positive	median	astounding	positive	high
no thicker	positive	high	not thicker	positive	high	innocent	positive	median
no thickest	positive	max	not thickest	positive	max	competitive	positive	median
no thickheaded	positive	median	not thickheaded	positive	median	independent	positive	median
no thorough	negative	median	not thorough	negative	median	inadequate	negative	median
no thoroughly	decrease	not thoroughly	decrease	inconvenient	negative	median		
no thoughtful	negative	median	not thoughtful	negative	median	comprehensive	positive	median
no thoughtfully	negative	median	not thoughtfully	negative	median	crucial	positive	high
no thrilled	negative	high	not thrilled	negative	high	outrageous	negative	median
no thrilling	negative	high	not thrilling	negative	high	lazy	negative	median
no timely	negative	median	not timely	negative	median	outdated	negative	median
no timid	positive	median	not timid	positive	median	lengthy	negative	median
no tireless	negative	median	not tireless	negative	median	artificial	negative	median
no tirelessly	negative	median	not tirelessly	negative	median	finicky	negative	median
no tiresome	positive	median	not tiresome	positive	median	deadly	positive	median
no tiring	positive	median	not tiring	positive	median	foremost	positive	median
no titillating	negative	very high	not titillating	negative	very high	jealous	negative	median

no together	negative	median	not together	negative	median	delicate	positive	median
no togetherly	negative	median	not togetherly	negative	median	challenging	positive	median
no tolerably	increase	not tolerably	increase	infamous	negative	median		
no tolerant	negative	median	not tolerant	negative	median	contemporary	positive	median
no top-quality	negative	median	not top-quality	negative	median	bold	positive	median
no totally	decrease	not totally	decrease	bitter	negative	median		
no tough	positive	median	not tough	positive	median	fitting	positive	low
no tragic	positive	median	not tragic	positive	median	legendary	positive	max
no trained	negative	median	not trained	negative	median	messy	negative	median
no traitorous	positive	median	not traitorous	positive	median	lonely	negative	median
no trashy	positive	high	not trashy	positive	high	ludicrous	negative	median
no treacherous	positive	high	not treacherous	positive	high	notorious	negative	median
no tremendous	negative	high	not tremendous	negative	high	glaring	increase	
no tremendously	minimize	not tremendously	minimize	definitive	positive	max		
no trickily	positive	median	not trickily	positive	median	desirable	positive	median
no tricky	positive	median	not tricky	positive	median	foolproof	positive	median
no trifling	positive	median	not trifling	positive	median	mighty	positive	high
no trite	positive	median	not trite	positive	median	exclusive	positive	median
no triumphal	negative	median	not triumphal	negative	median	knowledgeable	positive	median
no triumphant	negative	median	not triumphant	negative	median	jerky	negative	median
no trivial	positive	median	not trivial	positive	median	fussy	negative	median
no troubled	positive	median	not troubled	positive	median	inane	negative	median
no true-blue	negative	median	not true-blue	negative	median	aggressive	negative	median
no truer	negative	high	not truer	negative	high	optimum	positive	max
no truest	negative	max	not truest	negative	max	impractical	negative	median
no trusting	negative	median	not trusting	negative	median	cynical	negative	median
no trustworthy	negative	median	not trustworthy	negative	median	effortless	positive	median
no trusty	negative	median	not trusty	negative	median	cryptic	negative	median
no truthful	negative	median	not truthful	negative	median	concise	positive	median
no truthfully	negative	median	not truthfully	negative	median	dazzling	positive	high
no two-faced	positive	high	not two-faced	positive	high	exhaustive	positive	median
no typical	negative	median	not typical	negative	median	inaccurate	negative	median
no typically	negative	median	not typically	negative	median	dear	positive	median
no ugly	positive	median	not ugly	positive	median	exotic	positive	median
no ultimate	minimize	not ultimate	minimize	illegal	negative	median		
no ultimately	minimize	not ultimately	minimize	melodramatic	negative	median		
no ultra	decrease	not ultra	decrease	irrelevant	negative	median		
no unacceptable	positive	median	not unacceptable	positive	median	faint	negative	median
no unaccomplished	positive	median	not unaccomplished	positive	median	inconsistent	negative	median
no unafraid	negative	median	not unafraid	negative	median	horrifying	negative	high
no unassuming	negative	median	not unassuming	negative	median	meaningful	positive	median
no unbalanced	positive	median	not unbalanced	positive	median	damaging	negative	median
no unbearable	positive	median	not unbearable	positive	median	erotic	positive	median
no unbelievable	negative	median	not unbelievable	negative	median	dissatisfied	negative	median
no unbelievably	minimize	not unbelievably	minimize	incorrect	negative	median		
no unbiased	negative	median	not unbiased	negative	median	controversial	negative	low
no unchaste	positive	median	not unchaste	positive	median	lifelike	positive	median
no uncivil	positive	median	not uncivil	positive	median	blatant	negative	median
no uncivilized	positive	median	not uncivilized	positive	median	dubious	negative	median
no unclean	positive	median	not unclean	positive	median	menacing	negative	median
no unclear	positive	median	not unclear	positive	median	misguided	negative	median
no uncomely	positive	median	not uncomely	positive	median	distinctive	positive	median
no uncomplicated	negative	median	not uncomplicated	negative	median	insufficient	negative	median
no unconventional	negative	median	not unconventional	negative	median	calm	positive	median
no uncouth	positive	median	not uncouth	positive	median	grateful	positive	median
no uncultured	positive	median	not uncultured	positive	median	indispensable	positive	max
no undaunted	negative	median	not undaunted	negative	median	manic	positive	median
no undecorated	positive	median	not undecorated	positive	median	eclectic	positive	low
no undependable	positive	median	not undependable	positive	median	awry	negative	median
no underhanded	positive	high	not underhanded	positive	high	bulkier	negative	median
no understandable	negative	median	not understandable	negative	median	mystical	positive	median
no undistinguished	positive	median	not undistinguished	positive	median	commendable	positive	median
no uneasy	positive	median	not uneasy	positive	median	crummy	negative	median
no uneducated	positive	median	not uneducated	positive	median	gruesome	negative	median
no unenlightened	positive	median	not unenlightened	positive	median	fictional	negative	median
no unequal	positive	median	not unequal	positive	median	level	positive	median
no unequalled	negative	median	not unequalled	negative	median	formal	positive	median
no unethical	positive	median	not unethical	positive	median	inappropriate	negative	median
no uneven	positive	median	not uneven	positive	median	appalling	negative	median
no unexceptional	positive	median	not unexceptional	positive	median	impeccable	positive	high
no unexciting	positive	median	not unexciting	positive	median	frantic	negative	median
no unexpected	negative	median	not unexpected	negative	median	irresistible	positive	high
no unexpectedly	negative	median	not unexpectedly	negative	median	harmless	positive	low
no unfailing	negative	median	not unfailing	negative	median	discriminating	positive	median
no unfailingly	negative	median	not unfailingly	negative	median	intrusive	negative	median
no unfair	positive	median	not unfair	positive	median	offbeat	positive	low
no unfairly	positive	median	not unfairly	positive	median	cranky	negative	median
no unfaithful	positive	median	not unfaithful	positive	median	banal	negative	median
no unfeeling	positive	median	not unfeeling	positive	median	legitimate	positive	low
no unfit	positive	median	not unfit	positive	median	famed	positive	median
no unflagging	negative	high	not unflagging	negative	high	functioning	positive	low
no unforgettable	negative	median	not unforgettable	negative	median	edgy	positive	median
no unfortunate	positive	median	not unfortunate	positive	median	earnest	positive	median
no unfunny	positive	median	not unfunny	positive	median	classical	positive	median
no ungraceful	positive	median	not ungraceful	positive	median	inoperable	negative	median
no ungracious	positive	median	not ungracious	positive	median	lush	positive	median
no unhappy	positive	median	not unhappy	positive	median	fundamental	positive	median
no unharmonious	positive	median	not unharmonious	positive	median	adventurous	positive	median
no unhealthy	positive	median	not unhealthy	positive	median	inviting	positive	median
no unheard-of	positive	median	not unheard-of	positive	median	careless	negative	median
no unhinged	positive	median	not unhinged	positive	median	murky	negative	median
no unhonest	positive	median	not unhonest	positive	median	icy	negative	median
no unidentified	positive	median	not unidentified	positive	median	expressive	positive	median
no unified	negative	median	not unified	negative	median	hopeful	positive	median
no unilluminated	positive	median	not unilluminated	positive	median	illuminating	positive	median

no unimaginative	positive	median	not unimaginative	positive	median	disconcerting	negative	median
no unimportant	positive	median	not unimportant	positive	median	enigmatic	positive	median
no uninspiring	positive	median	not uninspiring	positive	median	inexcusable	negative	high
no uninstructed	positive	median	not uninstructed	positive	median	deserving	positive	median
no unintelligent	positive	median	not unintelligent	positive	median	glamorous	positive	median
no uninteresting	positive	median	not uninteresting	positive	median	drastic	negative	median
no uninviting	positive	median	not uninviting	positive	median	appreciable	positive	low
no unique	negative	median	not unique	negative	median	gory	negative	median
no uniquely	negative	median	not uniquely	negative	median	cocky	negative	median
no unjust	positive	median	not unjust	positive	median	juicy	positive	median
no unkind	positive	median	not unkind	positive	median	frivolous	negative	median
no unknown	positive	median	not unknown	positive	median	delicious	positive	median
no unlearned	positive	median	not unlearned	positive	median	drab	negative	median
no unlettered	positive	high	not unlettered	positive	high	incoherent	negative	median
no unlovely	positive	median	not unlovely	positive	median	elementary	negative	median
no unloyal	positive	median	not unloyal	positive	median	hostile	negative	median
no unlucky	positive	median	not unlucky	positive	median	lavish	positive	median
no unmannerly	positive	median	not unmannerly	positive	median	decisive	positive	median
no unmistakably	minimize	not unmistakably	minimize	arbitrary	negative	median		
no unmusical	positive	median	not unmusical	positive	median	hazy	negative	median
no unnatural	positive	median	not unnatural	positive	median	ghastly	negative	high
no unnecessary	positive	median	not unnecessary	positive	median	darned	negative	median
no unostentatious	negative	median	not unostentatious	negative	median	formidable	positive	median
no unparalleled	negative	high	not unparalleled	negative	high	articulate	positive	median
no unpleasant	positive	median	not unpleasant	positive	median	limitless	positive	median
no unpleasantly	positive	median	not unpleasantly	positive	median	indifferent	negative	low
no unpolished	positive	median	not unpolished	positive	median	lurid	negative	median
no unprecedented	negative	high	not unprecedented	negative	high	dumber	negative	high
no unpredictable	positive	median	not unpredictable	positive	median	incessant	negative	median
no unpretentious	negative	median	not unpretentious	negative	median	incomplete	negative	median
no unprincipled	positive	median	not unprincipled	positive	median	imperfect	negative	median
no unproductive	positive	median	not unproductive	positive	median	frenzied	negative	median
no unprotected	positive	median	not unprotected	positive	median	fierce	positive	median
no unqualified	positive	median	not unqualified	positive	median	maddening	negative	median
no unquiet	positive	median	not unquiet	positive	median	homicidal	negative	median
no unrefined	positive	median	not unrefined	positive	median	heavenly	positive	median
no unreliable	positive	median	not unreliable	positive	median	interminable	negative	high
no unremarkable	positive	median	not unremarkable	positive	median	idyllic	positive	median
no unrenowned	positive	median	not unrenowned	positive	median	gruff	negative	low
no unrighteous	positive	median	not unrighteous	positive	median	amoral	negative	median
no unschooled	positive	median	not unschooled	positive	median	infectious	positive	median
no unscrupulous	positive	median	not unscrupulous	positive	median	incompatible	negative	median
no unseemly	positive	median	not unseemly	positive	median	inspirational	positive	median
no unsettled	positive	median	not unsettled	positive	median	compulsive	negative	median
no unsightly	positive	median	not unsightly	positive	median	distraught	negative	median
no unskilled	positive	median	not unskilled	positive	median	hardy	positive	median
no unsophisticated	positive	median	not unsophisticated	positive	median	dizzy	negative	median
no unsound	positive	median	not unsound	positive	median	exhilarating	positive	median
no unsuccessful	positive	median	not unsuccessful	positive	median	chaotic	negative	median
no unsuccessfully	positive	median	not unsuccessfully	positive	median	borderline	negative	low
no unsung	negative	median	not unsung	negative	median	methodical	positive	median
no unsure	positive	median	not unsure	positive	median	devilish	negative	median
no unswerving	negative	high	not unswerving	negative	high	discerning	positive	median
no unsymmetrical	positive	median	not unsymmetrical	positive	median	lasting	positive	median
no untaught	positive	median	not untaught	positive	median	ferocious	positive	median
no untimely	positive	median	not untimely	positive	median	discouraging	negative	median
no untiring	negative	median	not untiring	negative	median	excruciating	negative	high
no untrained	positive	median	not untrained	positive	median	exuberant	positive	high
no untrustworthy	positive	median	not untrustworthy	positive	median	meaty	positive	median
no untruthful	positive	median	not untruthful	positive	median	imprecise	negative	median
no untutored	positive	median	not untutored	positive	median	decorative	positive	low
no unusable	positive	median	not unusable	positive	median	barren	negative	median
no unusual	positive	median	not unusual	positive	median	conspicuous	negative	median
no unusually	decrease	not unusually	decrease	destructive	negative	median		
no unvarying	negative	median	not unvarying	negative	median	influential	positive	median
no unwavering	negative	high	not unwavering	negative	high	feasible	positive	low
no unwilling	positive	median	not unwilling	positive	median	merciless	negative	high
no upbeat	negative	median	not upbeat	negative	median	agile	positive	median
no upright	negative	median	not upright	negative	median	novel	positive	median
no uprightly	negative	median	not uprightly	negative	median	invincible	positive	max
no upstanding	negative	median	not upstanding	negative	median	infuriating	negative	high
no uptight	positive	median	not uptight	positive	median	fruitless	negative	median
no usable	negative	low	not usable	negative	low	graceful	positive	median
no useable	negative	median	not useable	negative	median	quality	positive	median
no useful	negative	median	not useful	negative	median	unbearable	negative	median
no usefully	negative	median	not usefully	negative	median	paranoid	negative	median
no useless	positive	high	not useless	positive	high	tragic	negative	median
no uselessly	positive	high	not uselessly	positive	high	painful	negative	median
no usual	negative	median	not usual	negative	median	optimistic	positive	median
no usually	negative	median	not usually	negative	median	beloved	positive	median
no utilitarian	negative	median	not utilitarian	negative	median	painless	positive	median
no utterly	minimize	not utterly	minimize	proud	positive	median		
no vacillating	positive	median	not vacillating	positive	median	sleek	positive	median
no vague	positive	median	not vague	positive	median	sturdy	positive	median
no vain	positive	median	not vain	positive	median	disabled	negative	median
no vainly	positive	median	not vainly	positive	median	prime	positive	median
no valiant	negative	median	not valiant	negative	median	speedy	positive	median
no valiantly	negative	median	not valiantly	negative	median	high-resolution	positive	median
no valid	negative	median	not valid	negative	median	strong	positive	median
no valuable	negative	median	not valuable	negative	median	vague	negative	median
no valued	negative	median	not valued	negative	median	tough	negative	median
no valueless	positive	high	not valueless	positive	high	technical	negative	median
no vapid	positive	high	not vapid	positive	high	usable	positive	low
no variable	positive	median	not variable	positive	median	streamlined	positive	median
no vastly	decrease	not vastly	decrease	utilitarian	positive	median		

no veracious	negative	median	not veracious	negative	median	understandable	positive	median
no veritable	negative	median	not veritable	negative	median	readable	positive	median
no versatile	negative	median	not versatile	negative	median	useable	positive	median
no very	decrease	not very	decrease	straightforward	positive	median		
no vibrant	negative	median	not vibrant	negative	median	versatile	positive	median
no vicious	positive	median	not vicious	positive	median	obvious	positive	median
no vigilant	negative	median	not vigilant	negative	median	poorer	negative	high
no vigorous	negative	median	not vigorous	negative	median	disastrous	negative	median
no vigorously	negative	median	not vigorously	negative	median	shocking	negative	median
no vile	positive	high	not vile	positive	high	noisy	negative	median
no villainous	positive	median	not villainous	positive	median	immense	positive	median
no violent	positive	median	not violent	positive	median	unforgettable	positive	median
no virtuous	negative	high	not virtuous	negative	high	revolutionary	positive	median
no vivacious	negative	median	not vivacious	negative	median	magical	positive	median
no vivid	negative	high	not vivid	negative	high	wild	positive	median
no vividly	negative	high	not vividly	negative	high	unbelievable	positive	median
no vulgar	positive	high	not vulgar	positive	high	vibrant	positive	median
no wacky	negative	median	not wacky	negative	median	major	positive	median
no warm	negative	median	not warm	negative	median	redundant	negative	median
no warmer	negative	high	not warmer	negative	high	social	negative	median
no wary	negative	median	not wary	negative	median	alien	negative	median
no watchful	negative	median	not watchful	negative	median	ordinary	negative	median
no wayward	positive	median	not wayward	positive	median	mainstream	positive	median
no weak	positive	median	not weak	positive	median	unconventional	positive	median
no weaker	positive	high	not weaker	positive	high	plentiful	positive	median
no weakest	positive	max	not weakest	positive	max	wacky	positive	median
no weakly	positive	median	not weakly	positive	median	romantic	positive	median
no wearisome	positive	median	not wearisome	positive	median	worse	negative	high
no weird	positive	median	not weird	positive	median	worst	negative	max
no weirdly	positive	median	not weirdly	positive	median	sadistic	negative	median
no weirdo	positive	median	not weirdo	positive	median	ruthless	negative	median
no welcome	negative	median	not welcome	negative	median	sinister	negative	median
no well	negative	median	not well	negative	median	questionable	negative	median
no well-formed	negative	median	not well-formed	negative	median	suspicious	negative	median
no well-known	negative	median	not well-known	negative	median	criminal	negative	median
no well-made	negative	median	not well-made	negative	median	spoiled	negative	median
no well-mannered	negative	median	not well-mannered	negative	median	guilty	negative	median
no well-off	negative	median	not well-off	negative	median	violent	negative	median
no well-proportioned	negative	median	not well-proportioned	negative	median	religious	positive	median
no whacko	positive	high	not whacko	positive	high	reputable	positive	median
no whacky	positive	median	not whacky	positive	median	model	positive	median
no whimpy	positive	median	not whimpy	positive	median	responsible	positive	median
no wholly	minimize	not wholly	minimize	legal	positive	median		
no wicked	positive	median	not wicked	positive	median	correct	positive	median
no wickedly	positive	median	not wickedly	positive	median	shabby	negative	median
no wild	negative	median	not wild	negative	median	unacceptable	negative	median
no wilful	positive	median	not wilful	positive	median	operational	positive	low
no willful	positive	median	not willful	positive	median	satisfactory	positive	low
no willowy	negative	median	not willowy	negative	median	suitable	positive	low
no wily	positive	median	not wily	positive	median	reasonable	positive	low
no wise	negative	median	not wise	negative	median	serviceable	positive	median
no wisely	negative	median	not wisely	negative	median	top-quality	positive	median
no wiser	negative	high	not wiser	negative	high	presentable	positive	median
no wisest	negative	max	not wisest	negative	max	well-made	positive	median
no witless	positive	median	not witless	positive	median	scenic	positive	median
no witty	negative	median	not witty	negative	median	rugged	positive	median
no woebegone	positive	median	not woebegone	positive	median	well	positive	median
no wonderful	negative	median	not wonderful	negative	median	realistic	positive	median
no wonderfully	negative	median	not wonderfully	negative	median	unwilling	negative	median
no woolly	positive	median	not woolly	positive	median	idle	negative	median
no worried	positive	median	not worried	positive	median	reluctant	negative	median
no worse	positive	high	not worse	positive	high	resistant	negative	median
no worst	positive	max	not worst	positive	max	late	negative	median
no worthily	negative	median	not worthily	negative	median	potent	positive	median
no worthless	positive	high	not worthless	positive	high	prompt	positive	median
no worthwhile	negative	median	not worthwhile	negative	median	serious	positive	median
no worthy	negative	median	not worthy	negative	median	key	positive	median
no wrapped	negative	median	not wrapped	negative	median	gay	negative	median
no wrathful	positive	median	not wrathful	positive	median	premium	positive	median
no wretched	positive	median	not wretched	positive	median	relevant	positive	median
no write-off	positive	median	not write-off	positive	median	golden	positive	median
no wrong	positive	median	not wrong	positive	median	precious	positive	median
no wrongful	positive	median	not wrongful	positive	median	modern	positive	median
no wrongfully	positive	median	not wrongfully	positive	median	unfortunate	negative	median
no wrongly	positive	median	not wrongly	positive	median	moderate	decrease	
no yucky	positive	median	not yucky	positive	median	slight	decrease	
no yuk	positive	median	not yuk	positive	median	minor	decrease	
no zealous	negative	median	not zealous	negative	median	ultra	increase	
no FALSE	positive	median	not FALSE	positive	median	severe	increase	
no TRUE	negative	median	not TRUE	negative	median	ultimate	maximize	

Table B.1: Sentiment Lexicon (Argamon 2007)