This Is What a Feminist Tweets Like: "Women's Language" and Styling Activist Identities in a #YesAllWomen Twitter Corpus

Eleanor A. Morikawa
The Graduate Center, City University of New York
THIS IS WHAT A FEMINIST TWEETS LIKE:
‘WOMEN’S LANGUAGE’ AND STYLING ACTIVIST IDENTITIES IN A
#YESALLWOMEN TWITTER CORPUS

BY

ELEANOR A. MORIKAWA

A dissertation submitted to the Graduate Faculty in Linguistics in partial fulfillment of the requirements for the degree Doctor of Philosophy, The City University of New York

2019
This is what a feminist tweets like:
'Women's language' and styling activist identities in a #YesAllWomen Twitter corpus

by

Eleanor A. Morikawa

This manuscript has been read and accepted for the Graduate Faculty in Linguistics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Date

Cecelia Cutler

Chair of Examining Committee

Date

Gita Martohardjono

Executive Officer

Supervisory Committee:

Matthew Garley

Jillian Cavanaugh

Lauren Squires

THE CITY UNIVERSITY OF NEW YORK
Abstract

This dissertation presents results of a study of linguistic practice in the context of feminist activism on Twitter. Twitter has become a primary medium for social and political activism and a rich venue for study of the relationship between digitally mediated language and identity production. The focus of this study is the viral Twitter hashtag #YesAllWomen, a hashtag that rose in popularity following a misogyny-motivated terrorist attack in the spring of 2014. This dissertation treats the #YesAllWomen hashtag as an imagined space and a Discourse (Gee, 2015) where language serves as a site for the production of gender and feminist identity.

This investigation is conducted through three related studies. The first examines intra-speaker variation among a group of self-identified women who actively participated in the #YesAllWomen Discourse. The study tracks these women’s use of features of “women’s language” (Lakoff, 1975) to determine whether they emerge as linguistic resources that women recruit when performing feminist stances. The results of this study indicate that features of an online feminist style include an increase in vulgar language, a decrease in overt markers of politeness, a decrease in hedging strategies, and a decrease in stable nonstandard variants. These findings suggest that when taking feminist stances online, women reject certain features of stereotypically feminine language and enhance others, according to some theoretical paradigms. A second quantitative study examines the use of the same features among a group of male allies who tweeted with the #YesAllWomen hashtag in support of its feminist message. The results suggest that these men exhibit intra-speaker variation that mirrors that of the women in terms of average frequencies of each feature, but is less statistically robust. However, an investigation of linguistic practices not captured by the quantitative corpus study suggests that men deploy these linguistic resources differently when participating in the #YesAllWomen thread than in other Twitter interactions, showing potential influence of audience design (Bell,
or linguistic accommodation strategies (Giles, Coupland, & Coupland, 1991a, 1991b; Giles & Ogay, 2007).

The third study reported in this dissertation concerns speaker attitudes toward language, gender, and feminism. The data source for this component of the research is a survey completed by a group of active contributors to the #YesAllWomen Discourse. The results reveal demographic properties of the population, which were absent from the Twitter corpora, and the participants’ attitudes toward linguistic practice and its relationship to gender and feminism. These attitudes show evidence of language and gender ideology regarding “ideal” (Kiesling, 2007) masculinities and femininities that echoes some of the components of Lakoff’s theory of women’s language. The emergent patterns in the participants’ shared language and gender ideology shed light on the processes by which the observed language variation and resulting style shift become locally meaningful in context. This dissertation contributes to studies of computer-mediated communication and sociolinguistics and demonstrates the value of hybrid quantitative and qualitative research methods.
Acknowledgements

No graduate student is an island, or at least no graduate student should be. I owe this dissertation to so many people who have helped and supported me in innumerable ways. First, I express my deepest gratitude to Cece Cutler, my advisor and mentor who always made the time, met me in any borough, and supported and guided me through every stage of graduate school, from coursework to teaching to journal article submission, conference presentations, and this dissertation. I am equally indebted to Matt Garley, who was always a reliable source of practical suggestions, wisdom, and humor, and whose advice inspired me to recognize my own accomplishments, to click “submit”, to let go of a draft I could keep revising forever. I also thank Jillian Cavanaugh, for giving me the best teaching assignment I’ve ever had, and for shepherding me through anthropology literature that could have easily overwhelmed me.

I could not have completed this project without feedback from Lauren Squires, who may be the hardest working linguist I know but still agreed to serve on my dissertation committee, and Martin Chodorow, who generously coached me through the quantitative portions of this project even though his research is in a completely different subfield. To my fellow students in the GC Linguistics program, past and present, thank you all for your companionship as we’ve taken this long and bizarre trip together. Every study group, committee meeting, post-Colloquium happy hour, and impromptu catch-up session kept me tethered to this community and reminded me I am not doing this alone. To those who finished before me: thank you for showing me it can be done. I especially thank Chaya Nove, Ivana Đurović, Ben Macaulay, Andre Eliatamby, Kelsey Swift, Matt Stuck, Rachel Rakov, and Michelle McSweeney for always being willing to listen and provide valuable perspective when things were rocky, and eager to share in celebration of even the smallest triumph.

I extend a very special thank you to the members of the Onion Soup Writing Group, Pablo Gonzalez and Danielle Ronkos. You are the best support group and friends a doctoral student can ask for. Writing groups are nominally about holding one another to a writing
schedule and providing feedback on drafts, but I think pretty much anyone could have done that for me. Onion Soup was so much more than a writing group, and the fact that you deserve your own paragraph in my acknowledgements tells you everything you need to know. May we continue to help one another craft delicately worded emails for years to come.

I am also deeply grateful to Mom, Dad, Ben, Liz, all the Morikawas, and the rest of my family and friends for their unwavering support and for never asking how much longer I had until graduation. And above all, to Evan, who always believes in me.

Finally, I truly—literally—could not have done this without all of the citizens of the Twitterverse who tweeted with the #YesAllWomen hashtag hoping to make the world a better, safer place, especially those who took the time to complete a survey for this project. Besides providing the data, they continuously reminded me that the Internet still has room for humanity, generosity, and hope.
# Table of Contents

Chapter 1. Introduction ........................................................................................................ 1
1 Purpose ............................................................................................................................ 1
2 Background: Twitter and Twitter terminology ............................................................. 3
3 Origin of the #YesAllWomen movement ..................................................................... 5
4 Overview of chapters .................................................................................................... 10

Chapter 2. Theoretical background ................................................................................ 13
1 Discourse and discourse ............................................................................................... 13
2 Gender .......................................................................................................................... 14
  2.1 Gender as a sociolinguistic variable: three waves of sociolinguistics .................... 14
  2.2 Deficit, dominance, and difference ........................................................................... 17
  2.3 Gender in CMC ........................................................................................................ 19
  2.4 Queer theory, queer linguistics, and nonbinary approaches to gender ................... 21
  2.5 Gender as performance ............................................................................................ 22
3 Identity .......................................................................................................................... 23
4 Power ............................................................................................................................ 24
5 Feminism ....................................................................................................................... 25
  5.2 Feminist linguistics ................................................................................................... 27
  5.3 Waves of feminism .................................................................................................. 27
  5.4 Cyberfeminism ......................................................................................................... 29
6 Digital activism ............................................................................................................. 29
  6.1 Activism vs. “slacktivism” ......................................................................................... 29
  6.2 Trolls and trolling .................................................................................................... 31
7 Ideology ......................................................................................................................... 32
8 Stance and style ............................................................................................................ 34
  8.1 Stance ....................................................................................................................... 34
  8.2 Style .......................................................................................................................... 34
9 Audience design and participation roles .................................................................... 35
  9.1 Audience in CMC .................................................................................................... 36
10 Features ....................................................................................................................... 37
  10.1 Taboo language ....................................................................................................... 37
  10.2 Politeness ................................................................................................................ 39
  10.3 Hedging ................................................................................................................... 40
  10.4 Nonstandard variants ............................................................................................. 41
11 Research gap and motivation .................................................................................... 42
4.2 Understanding of gender ................................................................. 137
4.3 Additional questions and final thoughts ........................................ 138
Appendices ....................................................................................... 139
  Appendix A: Search terms by feature category .................................. 139
  Appendix B: Blank survey ................................................................ 140
References ....................................................................................... 169
List of Figures

Figure 1. Mean frequencies of feature per tweet, women ........................................... 50
Figure 2. Frequency of tweets by tweet length (abbreviations attested) ......................... 57
Figure 3. Frequency of tweets by tweet length (abbreviations not attested) ..................... 57
Figure 4. Mean frequency of feature per tweet, men .................................................... 70
Figure 5. Average feature frequencies per tweet, all corpora ....................................... 74
Figure 6. Survey respondents’ gender identity by age group ......................................... 94
Figure 7. Survey respondents’ education level by age group ......................................... 97
Figure 8. Likert-scale-style survey question .................................................................. 100
Figure 9. Pair of tweets .................................................................................................. 115
Figure 10. Likert scale accompanying pair of tweets .................................................... 116
Figure 11. Tweet ............................................................................................................ 117
Figure 12. Tweet ............................................................................................................ 117
Figure 13. Tweet ............................................................................................................ 117
Figure 14. Language in the workplace prompt and Likert scale ................................... 119

List of Tables

Table 1. Results of comparison between reduced corpora by women ............................ 72
Table 2. Education levels of survey participants ......................................................... 96
Chapter 1. Introduction

1 Purpose

Public discussion about feminism and gender equality have pervaded popular media in recent years, due in part to online communication tools allowing for the rapid spread of global feminist discourses and a spate of opinion pieces on what both women and men ought to do in the pursuit of gender equality. Numerous recent non-academic essays have asserted that women contribute to their own marginalization, especially in their professional lives, with the ways they purportedly use language. Essays like “Why women apologize and should stop” (Crosley, 2015) and “Google and Apple Alum says using this word can damage your credibility” (in which the author argues that women in the workplace too often use “just” to reduce the force of statements or requests) (Leanse, 2015) imply that the responsibility to minimize workplace gender inequality lies with women’s choices in their own communications. Furthermore, the underlying assumption of such essays is that women must converge to the linguistic styles associated with men, who establish the stylistic norms that women should target. This assumption is marketable: a downloadable plug-in for Google’s web browser Chrome called the “Just Not Sorry extension” was released as a tool to purge “weak” language from email drafts. The software does not explicitly target women in its marketing materials, but was covered by the press as a “plugin for women” (Cauterucci, 2015). Based on these media, it would appear that hedging and apologizing less frequently is all it takes for a woman to empower herself.

These prescriptivist essays inspired a handful of counter-point op-eds, including several blog posts by the influential feminist linguist Deborah Cameron¹. The counter arguments

---

¹ See, for example, “How to write a bullshit article about women’s language” on [www.debuk.wordpress.com](http://www.debuk.wordpress.com) (Cameron, 2015).
underline the reality it is not the linguistic choices of women that give rise to inequality, but much larger influences of systemic sexism. Negative evaluation of language is rooted in negative evaluation of people. Women’s language is scrutinized and corrected regardless of its linguistic form, but the pervasive belief that linguistic self-policing will elevate women in their professional lives is still attractive enough to be at the center of viral Internet articles every few years. The authors of those prescriptive essays—women, in the case of those mentioned here—are nominally advocating for gender equality. However, they do so by suggesting that language used by women is somehow deficient and to blame for the workplace gender gap, rather than structural or historical conditions, and that female empowerment is at least partially achieved by convergence with men’s unmarked linguistic style.

In the absence of scholarly work on the subject, meta-discourse on connections between linguistic practice and gender equality is dominated by these online opinion essays speculating about how women’s linguistic practices are counter-productive to feminism. Researchers in linguistics have extensively studied gender differences in language variation (Cheshire, 2004; Dixon & Foster, 1997; Eckert, 1989b; Holmes, 1997; Holmes & Schnurr, 2006; R. T. Lakoff, 1973; Schilling-Estes, 2002; Spender, 1990), performance of gender and sexual identity (Butler, 1988; Campbell-Kibler, Podesva, Roberts, & Wong, 2001; Podesva, 2007), and the performance of political stances (Hall-Lew, Coppock, & Starr, 2010), but no systematic research has been conducted on how feminists—both women and their non-female allies—use language and construct identities as activists and assert their stances in support of gender equality. The role of linguistic variation in women’s assertion of authority in a patriarchal system is often discussed but not studied with academic rigor. Serious linguistic research on the relationship between language use and gender activism is necessary in order to provide a scholarly response to these concerns.
To that end, this dissertation investigates the relationships between language use, gender activism, and language ideology in digital media. The research seeks to answer the following guiding questions:

1. What linguistic features characterize participation in an overtly feminist digitally-mediated Discourse?

2. Are features of a conceptualization of “women’s language” salient for people when discussing contemporary gender and power politics?

3. In what ways are linguistic practices in online feminist discourse informed by or in dialogue with language ideologies?

To address these questions, I present an analysis of data from two main sources: a corpus of Twitter data authored by people who actively tweeted with the hashtag #YesAllWomen (fully explained below) during the peak of its popularity in 2014, and a set of surveys completed by a smaller group of #YesAllWomen contributors. The survey elicited the participants’ attitudes toward feminism, language use, and the relationship between language and gender. A combination of quantitative and qualitative results shed light onto the relationship between linguistic practice and performance of gender and a socio-political stance and how that relationship reflects and reproduces ideological processes.

2 Background: Twitter and Twitter terminology

Online social media have emerged as a site for political and social activism and advocacy (Bonilla & Rosa, 2015; Gerbaudo, 2012; Gerbaudo & Treré, 2017; Hands, 2011; Lokot, 2018; Pond & Lewis, 2019; Rotman et al., 2011). Campaigns for civil rights, equality, and legal reform have been established online both formally and informally. The discourse surrounding issues of political activism provides a rich resource for the study into the role of language in projections of stance and identity in online forms of protest. Twitter is arguably the preferred social medium
for online activists and as such is a particularly valuable resource for investigations into these issues.

Twitter is a popular micro-blogging website on which members publish brief messages limited to a certain number of characters. They may also include hyperlinks, videos, or images, but all of these forms of content count toward the character limit. A Twitter user’s posts (called ‘tweets’) are posted to their profile, which is subject to each user’s own privacy preferences, though they are entirely public by default. When users log onto Twitter, they are greeted by a homepage displaying a stream of tweets published by the users that they have chosen to follow. Twitter users may address one another directly in their posts by using another user’s Twitter username (or “handle”) in a message. Such messages are known as instances of “atting” or “at-replies”, derived from the “@” that begins every Twitter handle. Users may also “retweet” another user’s post to their own followers, either by reproducing the tweet exactly or by embedding it in another tweet with their own commentary. If a tweet is retweeted by many people in a short period of time, it is said to have “gone viral”, though the precise point at which a tweet becomes viral is impossible to identify (Alhabash & Mcalister, 2015). Definitions of virality vary, and it is generally evaluated with respect to a Twitter user’s baseline level of influence; a user with 500 followers whose tweets are usually retweeted fewer than ten times may consider a tweet viral if it is retweeted one thousand times, but a major celebrity whose tweets are routinely retweeted thousands of times would consider the threshold for virality far higher.

The hashtag is an integral component of Twitter as a platform for the spread of ideas. Hashtags are user-generated strings of alpha-numeric characters preceded by a pound sign (#) and serve multiple functions. Twitter users can place a hash symbol before any word in a tweet

---

2 As of 2019, Twitter posts are limited to 280 characters, but prior to October, 2018, the character limit was 140. The data analyzed in this study were all authored before the change in character limit.
(though they often occur at the end of the message), transforming that term into an index or keyword that is then searchable by anyone on Twitter. For example, (1) from Bruns and Burgess (2011) tags several keywords in a message about the 2011 Tsunami in Japan, presumably so anyone searching for those keywords may see it:

(1) #japan #tsunami is the real killer. #sendai #earthquake PGA only 0.82g. 2011 #chch #eqnz 2.2g http://j.mp/ecy39r

If a Twitter user searches for a hashtag using the Twitter search bar, any unprotected tweet containing that hashtag may be returned in the results. Hashtags may provide additional commentary on the content of the tweet, as in (2) below (Evans, 2016); they may tag a tweet as part of a Twitter-wide running joke or meme, as in (3); or they may mark a tweet as a contribution to a socio-political debate or movement, as in (4):

(2) My daughter will be making up school days until July 4. My senior loves the cancellations because no makeup for him. #nomoresnowdays

(3) The sound was coming...from the FOURTH FLOOR #SpookyTalesForLinguists

(4) Today marks the 2nd anniversary of the murder of 12-year-old #TamirRice Rest In Power brother #BlackLivesMatter

The use of Twitter as a venue for social or political activism (of which #BlackLivesMatter as seen in (4) is among the most famous examples), is the subject of this study. Specifically, this study concerns the hashtag #YesAllWomen, which marked a Twitter discourse about gender equality. The hashtag is still in use but is not used nearly as frequently as it was in the spring of 2014 when it first appeared.

3 Origin of the #YesAllWomen movement

3 Tweets are protected if a Twitter user chooses to make their profile only visible to people who request and are granted access.
On May 24, 2014, a man named Elliot Rodger went on a shooting rampage in Isla Vista, CA, near the campus of the University of California at Santa Barbara, killing 6 before shooting and killing himself. It became apparent from Rodger’s online diary entries, YouTube videos, and a “manifesto” declaring his hatred of women that his act of terrorism was motivated by rage toward all women, fueled by what he perceived as lifelong romantic rejection. Much of the public response to the shooting focused on the dangers of misogyny and systemic sexism (Pachal, 2014; Weiss, 2014). Many argued that American men are socialized to feel entitled to control over women’s bodies and actions and subsequently feel entitled to revenge when their desires are not met. In response to this argument, other writers, primarily men, reminded the former group that the vast majority of men are socialized in the a similar environment as Rodger and do not grow up to treat women badly and then resort to violence when their behavior is not rewarded (Klingbiel, 2016; Zimmerman, 2014). The refrain of “not all men are like that” became a common thread in the online discourse about feminism and misogyny, and the shorthand trope “not all men” was taken up as a satirical slogan by feminist writers on the Internet. Example (5) is a tweet making typical use of the phrase:

(5) #Notallmen understand that it’s not all about them.

Shortly thereafter, the Twitter hashtag #YesAllWomen was coined. The first known use occurred on May 24, 2014. The main thrust of the hashtag was that while not all men are guilty of violence or even disrespect toward women, all women face the threat of harassment and sexism on a daily basis. This underscores that the point of the discussion of systemic sexism was not to generalize about male behavior, but to generalize about the female experience. (6) shows an early example of the tag:
(6) #notallmen practice violence against women but #yesallwomen live with the threat of male violence. Every. Single. Day. All over the world.

Starting in May 2014, women began sharing their stories of assault, harassment, or other instances of misogynistic speech or behavior and tagging them with #YesAllWomen. The movement sought to draw attention not only to the frequency of violence and abuse toward women, but also the more quotidian manifestations of sexism. Even if women are spared violence or other abuse at the hands of men, it is impossible to exist as a woman without being subject to frequent harassment or damaging but non-violent gender discrimination. In the earliest examples, the hashtag was syntactically incorporated into the tweet:

(7) #YesAllWomen know what it’s to be constantly interrupted, mocked for voicing legitimate concerns, called crazy for resisting oppression

It quickly became more common to see the tag operating as an independent clause, either at the end of a tweet or preceding because:

(8) Because every woman I know has experienced some form of sexual harassment, abuse or assault, myself included. #YesAllWomen

(9) #YesAllWomen because the media will mourn the lives of ruined high school football players, but not of the girls they assaulted

(10) “I have a boyfriend” is the easiest way to get a man to leave you alone. Because he respects another man more than you. #YesAllWomen

The #YesAllWomen hashtag gained momentum rapidly and became an ad hoc public⁴ (Bruns & Burgess, 2011) that dominated Twitter feeds around the world. It was tweeted over one million times in two days (Pachal, 2014). Even non-Twitter users were aware of the #YesAllWomen

---

⁴ Ad hoc publics are online communities of people that center around an issue or event and organize very quickly following an unplanned inciting incident. Ad hoc public formation is facilitated by hashtag use.
movement through the extensive media coverage it received. The movement constituted perhaps highest profile instance of cyberfeminism (Daniels, 2009; Hall, 1996; Wilding, 1998a) at that point in the social media age.

This research arrives at a moment when activism on social media, particularly on Twitter, is highly prevalent and has proven to have significant offline (or “real world”) consequences. For example, Twitter was an indispensable tool in the 2011 series of political revolutions known as the Arab Spring (Lotan et al., 2011). The protests in response to the shooting of Michael Brown, an unarmed black teenager, adopted the mantra “Black Lives Matter”, which became a viral hashtag on Twitter and gave rise to a wave of protests across the country and an enduring, established organization calling itself a “global network and movement”.

Online gender activism has emerged as another highly visible application of social networking for social causes. In the past few years, two hashtags have permeated public conversation about gender equality. The first, the #TimesUp movement, was founded in early 2018 by female celebrities in the American film industry as a response to revelations about the producer Harvey Weinstein, who had a history of sexual harassment and abuse and whose patterns of behavior were apparently not unique among powerful men in Hollywood. The catchphrase “Time’s up”, referring to the end of an era when such behavior is tolerated or willfully ignored, became the name of an associated nonprofit organization seeking to ensure safe working environments for women.

The other hashtag, #MeToo, has reshaped public discourse about sexual harassment and abuse in a way no other hashtag has. The phrase in this context originated on the social networking site MySpace in 2006 in a post by activist Tarana Burke and became newly viral on Twitter in 2017 following the public revelations about Harvey Weinstein and other powerful

---

5 See BlackLivesMatter.com
6 Timesupnow.com
figures in the entertainment industry who had abused female colleagues. The refrain of #MeToo spread from the film industry to other fields and was taken up in offline conversation in grammatically innovative ways. It is common to see references to an industry having a “Me Too moment” (Chamallas, 2018) or “the Me Too era” (Baum, 2019) or a male celebrity being “Me Too’d” (See Gaillot (2018) for a discussion of the implications of this passive construction).

The hashtag at the center of this study, #YesAllWomen, predates these other movements and in many ways set the stage for them. It was the first gender-focused viral hashtag focusing on the ubiquity of mistreatment of women in contemporary American society. It is also unique among these online trends in a number of measures. First of all, unlike #MeToo, it does not implicate the speaker/author as a victim in its linguistic structure. Anyone can tweet #YesAllWomen, whereas #MeToo was conceived as a message to be uttered by people who have been victims of sexual assault or harassment, as evidenced by its first-person singular pronoun. It is also the most universalist of the three. #TimesUp and #MeToo became popular as a response to allegations of widespread sexual abuse in Hollywood, and then were appropriated for use by other industries, but the reckonings were generally siloed within a professional community. #YesAllWomen, by contrast, was explicitly used as a universal truth for all women in all parts of the world and did not specifically focus on workplace discrimination. Finally, the most important distinguishing feature of #YesAllWomen is the presence of the word “women” in the hashtag. The other hashtags were originated and popularized by women, and women are undoubtedly the targets of the majority of sexual harassment and abuse, but the hashtags #MeToo and #TimesUp do not themselves draw attention to the gendered component of the movements. The #YesAllWomen discourse is therefore an even more productive space to conduct research on language and gender practices online, because gender is even more salient in the #YesAllWomen conversation than in other similar discussions organized around hashtags.
Contributors to the #YesAllWomen discourse appeared to be overwhelmingly female, but readers of the thread will encounter sympathetic tweets from male allies and the occasional inflammatory troll\(^7\). Due to the salience of gender in the discussion and the dominance of women in the participating population, the archive of tweets tagged with #YesAllWomen provides an invaluable resource for the study of the role of language in cyberfeminism and the general relationship between gender and computer-mediated communication (CMC). The women in the dataset are knowingly tweeting from a position of structural disempowerment, but with the goal of criticizing and subverting the patriarchal power structures at work in contemporary society. Given the complicated gender and power dynamics at play, it is not obvious what kinds of effects one would expect to see on the language used in the #YesAllWomen corpus. This study examines the effect of participation in online feminist activism on features of language often associated with women or disempowerment in linguistics literature, as well as speakers’ perceptions and ideologies surrounding those features and their social meaning. The findings offer insight into how gender and political stances are performed and interpreted in a digital medium.

4 Overview of chapters

The remainder of this dissertation presents a study in three stages and contextualizes it in several intellectual disciplines. The three sections report findings of 1) an intra-speaker variation study among women with respect to linguistic features that have historically been treated as indexing gender in some way, 2) a similar study among men and comparison between the two populations, and 3) an attitudes and demographics survey that situates the variation data in the context of language ideology.

\(^7\) See chapter 2 for an exploration of what constitutes “trolling”.
Chapter 2 summarizes intellectual traditions and theoretical frameworks relevant to this study. These include the history of language and gender research in three waves of sociolinguistic research, the major paradigm shifts in feminism and feminist theory, accommodation theory, audience design theory, and various approaches to computer-mediated communication research. This chapter also includes an overview of previous work on the specific linguistic features studied in the subsequent chapters: politeness markers, vulgar language, hedging, and stable nonstandard variants. These features have historically been linked to gender and power throughout the history of sociolinguistic research, and the prior research contextualizes this dissertation’s motivating questions.

Chapters 3 through 5 present the results of the three-part study. Chapter 3 shows results of an intra-speaker variation study among self-identified women who actively contributed to the #YesAllWomen corpus. The set of chosen linguistic features are quantitatively tracked in two corpora: a collection of tweets tagged with #YesAllWomen authored by self-identified women, and one authored by the same group of women but not filtered by hashtag or subject, representing a fuller cross-section of the authors’ Twitter posts. A comparison between the two corpora is followed by a closer look at how particular features are used to perform a feminist stance on Twitter. The results of this chapter illustrate how certain linguistic features contribute to an apparent style-shift triggered by participation in a feminist discourse.

Chapter 4 follows a similar structure to Chapter 3 and presents an intra-speaker variation study among self-identified men who participate in the #YesAllWomen corpus as self-described allies. This chapter explores the question of whether feminist allyship is performed through convergence to a set of feminist linguistic norms established through women’s linguistic practice in this context. The results suggest that audience design serves as a mechanism by which the discourse’s delicate politics are navigated by well-intentioned men and that the resulting variation is more complex than identifiable convergence or divergence.
The third results chapter explores the language attitudes and ideologies of the research population through the analysis of an online survey completed by a subset of the Twitter users represented in the #YesAllWomen corpus. The findings reported in Chapter 5 bring more nuanced meaning and context to the quantitative results and provide insight into how the population in question derives meaning from the linguistic variants studied. The survey also provides demographic information that is otherwise lacking from this and most other Twitter corpora.

The final chapter summarizes the main findings of the overall study, their implications for the fields of study on which this dissertation builds, and potential future directions for this line of inquiry. It also addresses inevitable limitations in the methodology and how these may be mitigated in future research.
Chapter 2. Theoretical background

This dissertation sits at the intersection of several long intellectual histories and builds on work spanning multiple disciplines and traditions. It is organized around key concepts that will anchor the analysis throughout this study. Previous studies in sociolinguistics, linguistic anthropology, discourse analysis, feminist and gender studies, and digital media studies provide foundations for this research. This chapter provides an overview of these overlapping but distinct lines of inquiry and previous findings about the particular linguistic features I focus on in the subsequent chapters, and finally concludes with a summary of how this dissertation contributes to the various intellectual traditions that inform it.

1 Discourse and discourse

Throughout the dissertation, I treat the conversation marked with the hashtag #YesAllWomen and its participants as a “Discourse” in the sense of Gee (1989, 2004, 2007, 2015). Gee uses the term “big ‘D’ Discourses” to refer groups of people with shared systems of meaning and the indexical links that constitute those systems. This differs from “small-d ‘discourse’”, which Gee uses to refer to a stretch of language in use. I will use “discourse” to refer to a series of utterances that may take place within the context of a Discourse. Small-d discourses are strictly linguistic, whereas Discourses extend beyond language to other semiotic systems and the conditions that permit them to exist and be meaningful to participants.

When capitalized, the term refers to both the participants in an interaction (broadly defined) and the processes that give rise to social meaning that those participants bring to the interaction. This makes it a more general term than similar concepts that will be explored in greater detail later in this chapter, including “community of practice” (Eckert & McConnell-Ginet, 1999; Holmes & Meyerhoff, 1999) or “indexicality” (Eckert, 2008, 2012), which may refer to the participants or the systems of meaning, respectively, but not both. The term also overlaps with the process by which stance produces style through continuous deployment over time.
(Bucholtz, 2009; Johnstone, 2009; Kiesling, 2009). A Discourse is constructed through the same processes, but it includes the human players themselves in addition to the abstract categories used to analyze their social practices, including their language. This study presents an analysis of the #YesAllWomen Discourse and includes analyses of the linguistic practices that produce an identifiable style and the participants’ understandings of their own identity-building practices. The #YesAllWomen Discourse may also be nested within other Discourses (e.g. feminist Discourse, Twitter Discourse, etc.). I treat #YesAllWomen both as a Discourse and as a virtual space (King, 2011) where stancetaking strategies build localized meaning in the process of negotiating gender and political identities on the Internet.

2 Gender

It is necessary to provide theoretical background on what is meant by “gender”, as it has been theorized through various lenses. Sociolinguists, anthropologists, and gender studies experts view the progress of language and gender research slightly differently with slightly different terminologies for overlapping concepts. This section provides an overview of these traditions.

2.1 Gender as a sociolinguistic variable: three waves of sociolinguistics

As many language and gender studies do, this one begins with Lakoff’s (1973; 1975) theory of “women’s language”. I do not introduce women’s language as a given assumption or a lens through which to interpret my results, but as a framework that still drives some popular assumptions about how women speak. These popular assumptions inform some of the methods of the study, as explained later in this chapter. Lakoff argued that the language used by and about women is grounded in the presupposition that women are marginal and associated with trivial or unserious pursuits and abilities. The features characterizing women’s language included, in Lakoff’s view, the tendency of women to hedge their statements, to use vacuous
modifiers and intensifiers (“nice”, “pretty”), to avoid profanity, vulgarity, or other strong language, to exhibit a greater degree of politeness or deference, and to possess a wider lexicon of terms for superficial features like colors. Language used to describe women also reinforced a power divide, she argued, famously citing gendered word pairs like *master/mistress* or *bachelor/spinster*, among other examples, in which the feminine form of a word that was once semantically equivalent to its masculine counterpart takes on negative connotations consistent with negative assumptions about women (see also Shulz (1975)). Lakoff’s conclusions were based largely on impressionistic or anecdotal data drawn from interactions with a limited demographic of women (i.e. white, middle-class, educated). Scholars immediately began pointing out these issues and investigating the language-gender relationship more systematically, leading to several quantitative first-wave variation studies of language and gender that treated gender as a predictive variable in formal linguistic variation (Labov; Macaulay & Trevelyan, 1977; Trudgill, 1972; Wolfram, 1969). These methodologies abstracted over all aspects of the speakers’ identities except outwardly apparent gender or sex when drawing conclusions, often ignoring the countless other factors that influence how, when, and why people speak the way they do.

Sociolinguists and linguistic anthropologists shortly began re-theorizing language and gender in ways that moved away from conclusions that generalized about the differences between how men speak and how women speak. First of all, at least in English, all variants are attested in production by both men and women, and that linguistic variation occurs within gender groups and within individual speakers (Cameron, 2008). It is therefore neither useful nor accurate to make generalizations following the template “Men do X; women do Y”. Second-wave studies sought to remedy this through longer-term, ethnography-based methods,

---

8 Waves of sociolinguistic methodologies are not chronological; focus in influential works of theory has shifted from one method to another over time, but it is not the case that “second wave” studies completely supplanted “first wave” studies, and so on. The trends certainly overlap and all three theoretical traditions still influence work conducted today.
conducting smaller studies in terms of number of speaker-subjects, but incorporating biographical data that may have an effect on the variables studied (Eckert, 1989a; Gal, 1978). However, both first- and second-wave sociolinguistic studies treated macro social categories like gender, race, and class as pre-existing ones and superimposed them on speakers. Third-wave studies attempted to correct this by arguing that socially meaningful categories emerge through interaction. These studies held that social identities were constructed through, rather than deterministic of, stylistic choices. These studies focused heavily on communities of practice (Bucholtz & Hall, 2003; Eckert, 2000; Eckert & Mcconnell-Ginet, 1992; Eckert & McConnell-Ginet, 1992), groups of people united by shared practices, not by macro social categories projected onto the subjects by researchers in a top-down fashion. This approach dictates that researchers must attend to social categories that are meaningful to the population and that the relevance of gender to a study must be demonstrated by a population’s orientations (Schegloff, 1997). This position was critiqued by some who found it constricting to suggest researchers should not discuss gender until their subjects do, as often social categories are relevant to interaction unbeknownst to or unmentioned by participants (Billig, 1999). These issues are particularly salient to conversation analysts (Stokoe & Weatherall, 2002), but are also relevant to other branches of communication research.

Third-wave sociolinguistic tradition imagines the link between language variation and gender as Ochs (1992) theorized it, challenging the directness of the indexical relationship between linguistic variants and gender. Ochs argued that the relationship between language and gender is “non-exclusive”, “constitutive”, and “temporally transcendent”. This is to say that linguistic variants are by and large used by people of any gender, that features may carry social meaning that in turn partially constitutes constructions of gender as a social category, and that the meaning-making processes of language use transcend time of utterance. These properties inform the way language is related to gender, not by direct indexical link but locally meaningful social constructs (including, in the case of the current study, gender, feminism, and feminist
Twitter as a genre) that are constituted through practice.

2.2 Deficit, dominance, and difference

Lakoff’s characterization of women’s language carried an assumption that patterns in how women speak perpetuate or exacerbate structural sexism, another idea that attracted significant criticism. Talbot (2010) summarizes this basic assumption as a “deficit” model, which theorizes that women’s linguistic practices are a display of disempowerment and that women would benefit from self-policing with the intention of leveling their differences from men’s language. A deficit model of gendered language no longer serves as the basis of language and gender scholarship, but still persists in non-academic domains, as evidenced the recent Internet thinkpieces about how women can earn more respect in their professional lives by censoring their more “feminine” practices like hedging or apologizing.

A reanalysis of gendered language practices as reflective of unequal power structures that favor men constituted a “dominance” framework, in which men’s and women’s language production is theorized as a manifestation of patriarchal forces that dictate expectations (Spender, 1985). In a dominance model, women’s language may be reanalyzed as “powerless language” (O’Barr & Atkins, 1980). As Talbot points out, this theory presents gendered power disparity as monolithic, when in reality power relationships are dynamic and constituted interactionally, and it is reductive to attribute language difference to power differences, when it is not the case that every man has authority over every woman.

The third analysis of the gender-language relationship theorizes gender as a cultural difference. The difference model contends that gender differences are learned and that communication between genders is akin to intercultural communication (Maltz & Borker, 1982; Mulac, Bradac, & Gibbons, 2001; Tannen, 1990, 1994, 1997). A central component of the difference model is emphasis on the differences between a “rapport” or “report” approach to interaction, which theorizes that men generally prioritize information transfer and the epistemic
or “reporting” function of interaction, while women are more likely to foreground affective or social functions of language, appealing more to interpersonal relation than to efficiency of message delivery. These differences give rise to difficulty in cross-gender communication according to the difference model. Critics of the model argue that structural inequalities are ignored or erased in such a framework and that gendered language patterns cannot be researched in isolation from the power differentials between them. Furthermore, they argue that approaching gender as inherent differences, culturally informed as they are, reifies gender and reinforces gender stereotypes and essentialism (Aries, 1997; Cameron, 1992a; Crawford, 1995; Talbot, 2010).

Talbot (2010) calls for frameworks beyond these deficit, dominance, and difference, promoting poststructuralist models that treat language as “the site of the cultural production of gender identity”. Rather than treat gender as a fixed property that people carry throughout their lives that conditions their language use, she (and many others, as will be discussed later in this chapter) argues that language positions men and women differently and that gender identity is constructed via language use, not a condition on language use.

Talbot’s proposed framework is echoed in Levon and Mendes’s (2015) description of contemporary language and gender research as subscribing to an “emergentist” approach. They contrast this approach with previous “correlational” and “constructionist” frameworks, which roughly correspond to first- and second-wave sociolinguistic theories. Emergentist approaches reject construction of a reified, static identity and are instead concerned with how language-users exploit resources to build locally meaningful stances. This approach serves as a basis for this study’s analysis. The site for stancetaking and related meaning-building processes is in the current study a digitally mediated Discourse. The differences in language use and how the speakers make sense of them are interpreted as partially constitutive of the Discourse and all of the indexical relationships it contains.
2.3 Gender in CMC

In research on computer-mediated communication (CMC), numerous scholars have pursued questions of whether gender-based patterns of language use persist in digital media, but gender was late to appear in studies of online communication. Early CMC research ignored gender entirely. This may have been due to the assumption pervasive in the early days of the Internet that the new technology was a communicative equalizer. Online text-based communication was thought to democratize communication because the anonymity in Internet communication erased the baggage brought to interactions in which a participant’s gender, race, status, age, or education level is known (Androutsopoulos, 2006; Herring, 2000). Furthermore, the Web theoretically allowed any marginalized group to self-publish or organize their own social and political advocacy without needing to break into media dominated by gatekeepers of the ruling class. This incorrectly assumed equal access to the Internet, which for years was not at all the case. Though women composed only 5% of Internet users in the early 90s (Herring, 2003; Taylor, Kramarae, & Ebben, 1993), the Internet use gender gap had closed by the early twenty-first century and Internet use has remained evenly divided between men and women (Perrin & Duggan, 2015). However, even with relatively equal participation, the Web did not emerge as the democratizing force it was expected to be. Even when online presence was evenly divided along gender lines, men continued to be overrepresented in online editorial and mediating roles, exercising more control than women over online content, preventing female voices from being heard. This led some feminist writers to lament the lack of feminist activism online, despite the access that women theoretically had to free and open platforms for political organization (Herring, 2003).

---

9 Closely related to the concept of Computer-Mediated Discourse (CMD). The two are often used interchangeably. I use CMC here in the interest of consistency.
In addition to the assumption that Internet access would democratize political activism or organization, early CMC researchers expected the potential anonymity of online interaction to homogenize communication styles across gender. However, several studies of gender and CMC revealed gender patterns in online communication to mirror those of spoken communication, with women and men exhibiting the same linguistic differences observed in other communicative media. Herring’s research on online message boards suggested that men and women employed different communication styles when participating in online discussion, with men’s language being more adversarial and argumentative and women’s more oriented toward positive politeness strategies like explicit agreement, thanks, and questions (Herring, 1994, 1992, 1996). Despite the expectation that supposed anonymity would neutralize gender differences in online communication, the communication styles of the participants often conformed to expectations grounded in early work on language and gender. Interestingly, these gender differences in online communication also extend to blogs, a medium in which speakers are not in direct interaction with one another (Herring & Paolillo, 2006; Pedersen & MacAfee, 2007). Through these early studies of gender and Internet language in the 90s and 2000s, the hypothesis that the inherent anonymity of Internet language prevents or obscures a connection between gender and language use was repeatedly undermined.

Since the 1990s, studies of identity and linguistic variation online have found that digitally mediated language use affords people innumerable resources with which to construct identity, of which gender is one facet. These linguistic resources may be different from those afforded by unmediated talk, and may include orthographic choices (Miller, 2011), color or other visual stylization (Vaisman, 2014), and emoji or emoticon use (Huffaker & Calvert, 2006). The particular resources may be different or new, but linguistic variation and performance of gender and other identity-building processes are no less prevalent in mediated platforms than in co-present talk.
2.4 Queer theory, queer linguistics, and nonbinary approaches to gender

A full discussion of language and gender research necessitates some measure of the body of work on language and sexuality research and queer linguistics. Queer theory (Butler, 1990, 1993b) was instrumental in rebuilding gender studies such that identity is the result of symbolic practice, not a predictor or cause of it (Levon & Mendes, 2015). Cameron (2005) points out that research on language and sexuality is not a parallel line of inquiry to gender studies, but an intersecting one, because gender and sexual identities are in part mutually constitutive. The current study acknowledges the significance of sexuality in the subject matter, especially regarding issues of consent and desire, but the #YesAllWomen Discourse is not as rich a site for discussion of these issues, and such topics will remain relatively marginal to the analysis.

Recent work in sociolinguistics, and social sciences in general, has moved away from gender-based variation studies that are vulnerable to criticism for essentialism or over-abstraction of social categories. Treating gender as a predetermined binary is an essentialist framework, as gender is not a discretely bounded quality that fits one of two molds that people are assigned and carry through their lives. There has been recent work re-evaluating gender and sexual binaries in linguistics and promoting gender identity in more deliberately inclusive ways in language research (Gratton, 2016; Zimman, Davis, & Raclaw, 2014).

Overall, linguistics as a field is currently reckoning with how it approaches gender. It is valuable to bring theories from anthropology and feminist studies to bear on linguistic research approaches, which I attempt to do here. The dataset used in this study largely relies on a binary gender framework with assumptions about men and women as social categories, but I do not mean for this study to represent a move backward to reductive gender frameworks in the field. My research methods allow me to interpret data from which a binary understanding of gender emerges as a meaningful social construct among the population I am studying, in keeping with the tradition of third-wave sociolinguistics studies. The #YesAllWomen Discourse is predicated on conceptualizations of two mutually exclusive gender identities that carry a set of cultural
expectations and shared experiences. The internal ideological workings of these constructs, whether they are essentialist or oversimplified, are not the focus of this project. I elaborate on the limitations of the dataset and my methods in chapter 6.

2.5 Gender as performance

Gender in sociolinguistic research, particularly in variation studies, has frequently been given very essentialist and reified treatment; that is, it is discussed as a static attribute that people have. Additionally, it is repeatedly treated as inextricably tied to biological sex and heterosexuality, and is studied only where most salient (McElhenny, 2003). This study may be vulnerable to criticism on these counts, but my intention is to conduct and report gender and variation research with the awareness that this treatment of gender can lead researchers to incomplete or misleading conclusions about the role of gender in their data. My use of survey data as a supplement to quantitative variation research provides insight into the ideologies of gender, language, and feminism among the people represented in my data set. Furthermore, I incorporate Judith Butler's framework of gender as performance when approaching my research questions.

Butler (1988, 1990, 1993a, 2004a) famously challenged the idea of gender-as-attribute in favor of gender-as-performance, using an idea of “performativity” similar to Austin’s (1962). In other words, gender is not something a person is or has, but something one does, the result of repeated representation of gendered traits. Butler also crucially challenged the assumption that sex and gender were differentiated by the process of social construction; that is, the idea that gender is socially constructed while sex is not. Butler argued instead that our understanding of biological sex is as mediated by ideology as our understanding of gender, and naturalizing sex is not less problematic than naturalizing gender, race, or any other social category. Performance of gender is socially conditioned throughout a human’s life, and normative male or female gender performances are easily naturalized because they are so deeply internalized. This
dissertation builds on understandings of gender in previous variation studies, as well as a framework of performative gender, to explore how a certain kind of gendered stance is performed. The language analyzed here is produced by people expressing a righteous anger at perceived sexism and declaring their gender identity in the same public forum in which they express their feminist outrage. This project is not structured as investigation into how men and women talk about feminism in comparison with one another, in which the categories of “men” and “women” are constant underlying attributes of the speakers, but how people on the Internet perform a political stance that is closely tied with their simultaneously performed gender identity.

3 Identity

Recent research on language and gender, while consistently moving away from methodologies that impose social categories on its subjects in favor of more nuanced or dynamic frameworks of indexicality, has seen some debate regarding the extent to which certain ideas should be the focus of the field. One prominent camp foregrounds identity (Bucholtz & Hall, 2003, 2004), or the self-determined constellations of traits that speakers construct through practice, linguistic and otherwise.

This dissertation builds on theories of identity consistent with Bucholtz and Hall’s (2003, 2004) work, making particular use of their description of tactics of intersubjectivity. Bucholtz and Hall are deliberate in their use of the term “intersubjectivity” over “identity” because it emphasizes the both passive and agentive nature of identity construction; one is simultaneously the subject of social processes and subject to social processes, rendering identity relational rather than a static property. In this research, the term “identity” is assumed to be intersubjective and not imagined as a reified item that a speaker can possess.

The tactics of intersubjectivity that Bucholtz and Hall describe form three continua, each with a positive and negative pole. The first, “adequation vs. distinction”, concerns social
sameness or difference. The second, “authentication vs. denaturalization”, refers to the authenticity or artificiality of a subject’s construction. Lastly, “authorization vs. illegitimation” concerns institutional recognition or marginalization. These three axes—How alike is the subject? How authentic or genuine is the subject? What authority does the subject have?—provide a practical framework for understanding how identity is meaningfully performed in the datasets analyzed here.

Some scholars have argued that gender and sexuality research should not limit itself to matters of identity on the grounds that it unnecessarily limits the field’s scope (Cameron, 2005; Cameron & Kulick, 2005). Cameron and Kulick acknowledge that identity is a relevant concept for language and gender research but is not in and of itself the entirety of the field’s concern. They submit that desire be more central in gender and sexuality research as a supplement to theory accounting for identity production. The current study is more concerned with issues of identity than with desire, but the limitations of an identity-focused study are certainly relevant here. This dissertation examines how people construct gender and political stances online in a mediated context where those identities intersect and inform one another. However, the population’s language use achieves more beyond identity construction, though the particulars of non-identity work falls beyond my scope here.

4 Power

Research on the relationships between language use and the role of power in interaction evolved alongside research on language and gender, often intersecting it. Early critiques of women’s language suggested perception of power dynamics in interaction is a confound in gender-based variation studies, partially because institutional power structures typically favor men. O’Barr & Atkins (1980) proposed a re-analysis of the set of linguistic features previously characterized as feminine instead as “powerless language”. Their argument was based on
analysis of courtroom interactions, where power disparity is relatively institutionalized—a witness has less power than an attorney while being cross-examined, for example. Their results suggested that relative power was a better predictor of women’s language features than gender and proposed an alternative paradigm. Production and perception studies have since supported some of O’Barr’s and Atkins’s arguments, suggesting that political power inversely predicts frequency of hedging, a feature of “women’s” or “powerless language” and that hedged are perceived as less authoritative than their hedgeless counterparts (Bradac & Mulac, 1984; Hosman, 1989; Jalilifar & Alavi, 2011).

These studies, while effectively demonstrating that power structures must not be ignored in gender-based variation research, are subject to a critique leveled against many first-wave sociolinguistic studies: that an independent variable is treated as a fixed attribute of a speaker, rather than something that is negotiated in interaction. Neither gender nor power should be reified and projected onto speakers in a study without sufficient attention paid to context and attitudes brought by the participants. Though these studies acknowledge that power structures are mutable and typically more fluid than gender, race, or class had previously been treated in linguistic research, these studies of language and power remain somewhat simplistic and deterministic in their treatment of power as a variable. Later studies, particularly discourse analyses of workplace dynamics, demonstrated a more circumspect approach to understanding the role of power in communication (Holmes & Schnurr, 2005, 2006; Mills & Mullany, 2011; Mullany, 2011).

5  Feminism

5.1 Who’s a feminist?

In the early stages of this research, I faced the question of how to define “feminist Discourse” and whether projecting the title of Feminist onto the speakers in the study is problematic for the same reasons projecting any socially meaningful category onto a speaker is.
Feminism itself is far from monolithic and contains an unknowable number of feminisms (Lane, 2015). It follows from the undefinable nature of feminism that the term “feminist” or “feminist identity” as it is used here is a shorthand meant to comprise diverse personae associated with a Discourse seeking to expose and critique gender inequality. A single feminist identity is impossible to define or identify, and for my purposes it is not a productive or interesting venture to attempt to redefine what it means to be a feminist.

For the purpose of this research, apparently good-faith participation in the #YesAllWomen Discourse served as evidence of feminism on the part of the participant. In expressing support for the ideals of the #YesAllWomen discussion, a person is performing a sort of sociopolitical ideology that I consider feminist. One does not need to self-identify as a feminist to be one (Valenti, 2014). If someone clearly supports the central tenets of mainstream feminism but prefers not to label oneself as a “feminist”, this study will label them as one anyway. It is immaterial to the research whether someone wishing to challenge gender-based double standards, discrimination, and harassment self-identifies as a feminist. The speakers in this study align themselves with a public discussion denouncing sexism, which qualifies as feminist for my purposes.

Similarly, I am not concerned with whether men are eligible to qualify as feminists. Some argue that a feminist is by definition a woman, and that men can be feminist allies but feminism is reserved for female-identifying people (Duelli Klein, 1983). Forums designated as feminist spaces are traditionally intended to offer venues where men are unable to dominate the proceedings (Mills & Mullany, 2011). If feminist spaces are meant to provide insulation from men, then men cannot, by definition, actively participate within them, and if men do infiltrate such a space, it is by that token no longer a feminist one. I treat #YesAllWomen as a feminist space here, but not one that is insulated from anyone—indeed, it was a popular target for internet
s, male and female, as many online feminist forums are (Herring, Scheckler, & Barab, 2002). Some founders of feminist forums may intend for them to be protected from male interlopers, but public online media are not. If men participate in an overtly feminist forum, like the one studied here and demarcated by use of a hashtag, are they undermining their own intentions by intruding on a discussion that emerged as a metaphorical space safe from men? This issue is further addressed in section 6 of this chapter as part of a discussion of theories of audience design, as well as throughout the results chapters.

5.2 Feminist linguistics

This dissertation is not only an addition to the body of work on language and gender, but to that on language and feminism. It may also constitute an example of feminist linguistics, although not explicitly. Feminist linguistics is a discipline focused on studies of language and gender with the expressed objective of exposing how language use perpetuates misogynistic and anti-LGBT discrimination (Mills and Mullany, 2011). The findings of this project include evidence of anti-woman/LGBT linguistic discrimination, but the research itself is not exclusively motivated by a desire to prove its existence. Language and gender studies are not inherently feminist (Cameron, 1992b)10, and while this study is partially rooted in feminist linguistics, it does not have an overt political agenda.

5.3 Waves of feminism

10The most striking counter-example is perhaps Otto Jespersen’s chapter “The Woman”, in which anecdotal gender differences in language include women’s inherently smaller vocabularies and tendencies to think less than men before speaking (Jespersen, 1922). Many subsequent variation studies of gender are no more feminism-motivated than Jespersen’s, despite explicitly dealing with gender issues.
Like sociolinguistic research, feminism has undergone three often-cited waves (Cameron, 2005; Gillis, Howie, & Munford, 2007). The chronological progression between these frameworks was also gradual, the paradigm shifts not marked by identifiable moments but rather by a slow progression from one focus to the other. First-wave feminism in the United States was characterized by demands for suffrage, property rights, and education for women in the late 19th and early 20th century. A second wave in the 1960s and 1970s followed, commonly known as the “women’s liberation” movement, focusing on liberating women from subservience to men. Cameron (2005) likens second-wave feminism to “modernism” in the trajectory of feminist thought. The dominant themes of women’s liberation are still central to contemporary feminism, but the third wave of feminism that emerged in the late 20th and early 21st century—“postmodernism” in Cameron’s framing—focuses more on individual experience, diversity, and intersectionality than either of the previous movements. Intersectionality refers to the experience of people marginalized in multiple dimensions. Crenshaw (1989), the first to use the term in the context of feminism, argued that black women are marginalized both by feminism that focused on the experience of white women, and by antiracist rhetoric that focused on the experience of black men. The foregrounding of multidimensional exclusion of people plays a significant role in the Twitter Discourse analyzed in this dissertation. The #YesAllWomen hashtag began as a declaration of the universality of sexism and harassment among all women, unifying all women’s experiences. The generalization of the female experience drew criticism for ignoring the intersectional realities of women’s lives; in unifying women under a common lived experience, the Discourse abstracted over the overlapping prejudices experienced by women of color, LGBT women, disabled women, and any combination of marginalized identities (Heideman, 2014; Rodino-Colocino, 2014). The #YesAllWomen conversation was a site for both second-wave-esque, women’s liberation-inflected sentiment (“Women are systemically marginalized and rendered subservient to men”) and distinctly third-wave, intersectional,
individual-focused feminism ("Women as a social category is not meaningful for social justice when ignoring the full spectrum of the ways humans are dehumanized").

5.4 Cyberfeminism

The data studied here constitute a recent high-profile example of cyberfeminism (da Rimini, Starrs, Pierce, & Barratt, 1991; Daniels, 2009; Hall, 1996; Plant, 2000; Wilding, 1998b), a movement dating from the early 1990s seeking to harness digital technology as a tool to challenge patriarchal norms\textsuperscript{11}. The term cyberfeminism has fallen relatively out of use compared to the 1990s and early 2000s but serves as a useful descriptor for the observed phenomena of feminist Discourses on social media in recent years, including #YesAllWomen and #MeToo.

Cyberfeminism as a phenomenon exists outside of the three discernable waves of feminism or feminist theory. Third-wave feminism, while it accounts for realities of being a woman in more nuanced ways than previous feminist movements, does not explicitly attend to major shifts in communicative modalities. The rise of social media and their influence in social restructuring both on- and offline are integral to contemporary feminism. So much of contemporary feminism is performed and negotiated on the Internet and removing the digital modality from the theorization of feminism ignores significant shifts in feminist activism.

6 Digital activism

6.1 Activism vs. "slacktivism"

\textsuperscript{11} This admittedly reductive definition suffices for the purposes of this research, but in reality cyberfeminism, especially at the peak of its prominence in feminist studies in the 1990s, had no single definition and was perpetually contested by self-proclaimed cyberfeminists (Wilding, 1998)
Though cyberfeminism as a theoretical framework has receded in prominence in recent years, the rhetoric of digital activist spaces has attracted significant scholarly attention, including instances of online feminist movements. Online activism has been maligned and given the pejorative nicknames of “slacktivism” or “clicktivism” by critics who contend that it serves only to create the illusion of involvement for people wishing to feel politically engaged, but does not lead to material offline change (Chen, Pain, & Barner, 2018; Rotman et al., 2011; Vie, 2014). Digital media as activist venues are often evaluated in contrast with “boots on the ground” activism, which is defined by the physical presence of activists in offline, “real-world” spaces. This may be a partly generational effect, whereby people for whom the Internet has been a central component of life during their formative years view online activism as no less “real” than online dating or online shopping, two extremely popular applications of the Internet with clear implications for a person’s offline existence (Goodling, 2015). The value of online activism, argue those in defense of it, is in exchange of information and change of educational status that can transcend offline structural barriers that previously impeded such movement of ideas, and that increased awareness of grassroots movements leads to offline action (Goodling, 2015; Vie, 2014).

The #YesAllWomen movement, for example, sought to increase awareness of everyday sexism and the constant threat of violence against women, in the wake of an extreme case of misogynistic violence that attracted widespread public attention. The early users of the hashtag almost certainly hoped for offline effects of the awareness campaign, in both men’s and women’s attitudes or behavior. However, the hashtag was not explicitly tied to any offline organizations or events. This contrasts it with other high-profile online activist campaigns like #BlackLivesMatter or the online component of the Arab Spring, two cases in which online political activity corresponded to offline demonstrations or other action. #MeToo, the later viral hashtag with a similar message to #YesAllWomen, was similarly not explicitly associated with offline demonstrations, but has been credited (or blamed) for offline consequences including the
public accusation of sexual misconduct leveled at prominent men in various professional fields, and their consequential loss of professional status. #MeToo is arguably the most famous (and, in terms of direct offline effects, most successful) instance of an online awareness campaign for systemic sexism, but can be seen as the culmination of several campaigns with similar goals (Baer, 2016; Clark, 2014, 2016; Horeck, 2014; Kim, 2017; Mendes, Ringrose, & Keller, 2018; Novak & Khazraee, 2014; Rentschler, 2015; Williams, 2015).

The question of “real world” effects of digital activism extends to the possibility of “real world” risk to participants. Part of the critique of online activism is the idea that those who engage in it are not putting themselves at the same bodily risks as those who march in streets, sit at segregated lunch counters, or confront armed police. While the nature of risk may be different online, the attention online activists draw to themselves can result in offline harassment or danger. Online feminist activists are frequently targeted by online trolls or harassers, the psychological effects of which are as “real” as any endured from offline bullying. Furthermore, some online trolling incites “doxxing”, or the sharing of someone’s home address or other offline contact information, enabling harassers to target them or their families in the physical world (Ruitenberg, 2018). Doxxing is an illustrative example of how online activism carries offline risks, further challenging the assumption that online and offline activism function in entirely discrete, siloed spheres—the physical and the digital.

6.2 Trolls and trolling

While trolls and trolling are not the central objects of this study, the concept merits some explanation here because online activism cannot be thoroughly discussed without mention of backlash, including trolling. The term “trolling” has been used with regard to Internet communication at least since the mid 1990s, when the word “troll” was included in the online dictionary of Internet terminology NetLingo (Jansen and James, 1995). The editors of NetLingo claim that “trolling” in Internet parlance originally meant posting a comment on a blog or public
forum that exaggerates a particular aspect of the post or discussion, in an attempt to lure inexperienced participants into naively correcting it. The definition has since evolved to include any inflammatory commentary intended to provoke outrage or incite argument for one’s own entertainment (Bishop, 2014). The definition is likely to have been inspired by fishing, where “trolling” refers to a method by which multiple baited lines are drawn through the water. A metaphorical sense of “trolling” to mean baiting or provoking an adversary into a fight may have been used in the US military before being adopted into digital terminology (Bishop, 2014).

Trolling behavior of various flavors is documented across online communities (Bishop, 2014; Buckels, Trapnell, & Paulhus, 2014; Grieser, 2019; Hardaker & Mcglashan, 2016; S. Herring et al., 2002; Johnson, 2018; Poland, 2016). The common defining feature of the Internet troll is a superficial apparent sincerity, masking an intention not to engage in good-faith debate, but to lure sincere participants into futile argument they have no hope of winning or even constructively contributing to (Grieser, 2019). Internet trolls make incendiary remarks only as agents as chaos who revel in the genuine reaction they elicit, not to persuade their target to change their position on a given issue. This differentiates trolling from other forms of verbal abuse or harassment, which may share similar intentions of provoking a reaction, but is not conducted under the same falsely sincere guise.

7 Ideology

This project incorporates notions of language ideology taken from Irvine and Gal (2000). They use the term language ideologies to refer to the ideas that people use to frame their understanding of linguistic differences and how they connect them to ideas about speakers, groups, places, events, or other meaningful entities. In the study of linguistic variation and its relationship to performance of gender and feminism, language ideologies are central to the analysis. Language ideologies also allow the framework of “women’s language” to be a more helpful concept here. “Women’s language” has multiple meanings in this dissertation. On one
hand, as originally framed by Lakoff, it constitutes a theoretical framework of gendered language variation. On the other, it can be interpreted as an ideological framework rather than an empirical fact of language use. That is, aspects of “women’s language” emerge as a system of understanding the relationship between the concept of gender (itself an ideological construct) and linguistic practice. The frequently expressed idea that women hedge, apologize, or self-efface too frequently suggests an ideology that attributes these practices to an “ideal” (Kiesling, 2007) female speaker. Whether the features of women’s language are disproportionately used by women or not, women’s language remains a salient component of language and gender ideologization and—crucially for this study—a potential resource for identity production processes. As Hall (1995) argued in a study of phone sex operators—women who construct a particular interpretation of gender identity through a mediated channel, not unlike the women in the current study—the features of women’s language remain meaningful linguistic resources in performance and stylization of gender. The empirical value of women’s language as a theory of linguistic variation is less relevant here than its role in ideologically-mediated identity production.

The semiotic processes that Irvine and Gal define as means of language ideology development are also relevant here. They used concepts of iconization (the process by which perceptions of linguistic features are transferred to people), fractal recursivity (the projection of a binary opposition salient at one scale onto another), and erasure (the perceptual disappearing of internal diversity of a group) to theorize the processes by which ideology of linguistic difference is constructed. In the case of this study, erasure emerges as the most relevant, as it underlies assumptions like “men speak one way, women speak another”. However, iconization is also involved in construction of women’s language ideology as it gives rise to assumptions that language perceived as weak or tentative (hedges, apologies, etc.) is associated with people who are weak or tentative (women).
8 Stance and style

8.1 Stance

A large component of this study is the examination of stancetaking resources in the #YesAllWomen Discourse. Stances are ways speakers position themselves vis-à-vis other speakers, ideas, or an interaction (Jaffe, 2009). Stances that are continually useful in a particular sociolinguistic context come to be associated with particular social identities or situations, which in turn establish meaningful styles (Bucholtz, 2015; Johnstone, 2009; Kiesling, 2009). Repeated stancetaking strategies have been shown to contribute to identity construction on Twitter (Evans, 2016). In this case, a set of linguistic features constitute some of the resources speakers use to perform stances that, over time, coalesce into stylistic patterns.

As Ochs (1992) argues, linguistic variables may index stances or social activities directly, and these stances or activities in turn index broader social categories like gender, but the variables themselves do not directly index gender. Among the examples she cites are tag questions in English, which can index a hesitant stance, which may then index gender through distribution of the variable. In the case of this study, the linguistic variables studied were selected because they have been assumed to have some indexical relationship to gender, but this relationship exists by way of stances or activities with a more direct semiotic relationship to linguistic form. For example, vulgar or taboo lexical items may index aggressive, angry, or subversive stances, which in turn may index masculinity. Deferential or face-affirming stances may index femininity, and are in turn indexed by markers of overt politeness or hedges, among other linguistic variants. In the data analyzed here, the speakers exploit certain linguistic variables as stancetaking resources, and the resulting stances are tied to conceptualizations of gender or power.

8.2 Style
I also use conceptualizations of style (Bucholtz, 2015; Eckert, 2003; Squires, 2016) and indexical field (Eckert, 2008) when making sense of the identity work done by the linguistic practices of speakers in my data set. Specifically, Bucholtz’s conceptualization of style as a “way of doing things” or a “system of sociocultural positioning through modes of semiotic action” is useful here. This definition is intentionally broad enough to comprise not only linguistic style but all semiotic practices, though this study focuses on linguistic variation and the resulting construction of stances that constitute an online feminist linguistic style through repeated use. Bucholtz refers to stances as “the building blocks of identity” in that they produce discernible styles by virtue of sustained, recurring deployment.

The stances and styles contextualizing use of linguistic variants narrow their indexical fields. The concept of “indexical field” is of value here because social meanings are not fixed. I do not frame the linguistic variables I am examining as tethered in an immutable way to social meanings. In other words: profanity does not equal aggression or rudeness; hedging does not equal tentativeness. These are possible social meanings within the indexical field of these linguistic features, but the styles with which the features are deployed influence what social meanings are available to interlocutors.

9 Audience design and participation roles

The role of audience is a recurring theme in the body of literature on language, gender and power. Bell’s (1984) and Goffman’s (1981) frameworks of participation allow analyses of interaction with respect to each participant’s role in the Discourse and their effect on a given speaker’s linguistic choices. Audience as an influence on linguistic style-shifting or style-matching is well documented (Gonzales, Hancock, & Pennebaker, 2010; Niederhoffer & Pennebaker, 2002; Taylor & Thomas, 2008) and plays a significant role in the analyses presented here. Audience design theory has been particularly useful in problematizing
assumptions about language and gender. For example, though men are frequently presumed to be more prolific users of profanity, vulgarity, and taboo language, both men and women have been found to use more such linguistic forms in single-sex interactions than mixed-gender ones, especially with respect to informal anatomical terminology (Jay & Janschewitz, 2008; Wells, 1989, 1990). It is not sufficient to find significantly different rates of a given linguistic feature between men and women or between a superior and subordinate without careful attention to the audience, observed or imagined, that may influence communicative practice.

9.1 Audience in CMC

A difficult challenge of conducting Twitter research, as with all CMC research, is determining the precise roles of the participants. If Internet communication is exceptionally free, public, open, and anonymous, it is difficult to determine who exactly the speakers and audience are. This is especially true of Twitter as a medium, where individual addressees are not often mentioned by name. Bell’s audience design framework is useful to untangle the possible configurations of audience members here. There is evidence for audience design influencing style shifting on Twitter evidenced by at-replies and by speculation on imagined audience (Danescu-Niculescu-Mizil, Gamon, & Dumais, 2011; Marwick & Boyd, 2011; Nguyen, Trieschnigg, & Cornips, 2015; Pavalanathan & Eisenstein, 2015). The #YesAllWomen corpus, the primary source of feminist Discourse data in this study, contains evidence that many contributors imagine a female, like-minded addressee, but the stated objective of the #YesAllWomen movement was to raise awareness about the ubiquity of misogyny (Grinberg, 2014), suggesting that men are intended to be at least exposed to the Discourse, even if they are not the direct addressee. Considering the importance of men to the Discourse, although they do

---

12 The term speaker here is not intended to mean only one who produces spoken language. In some forms of CMC, there is a true speaker (as in video blogs), but many CMC frameworks use the term speaker to refer to the producer of language, so that is the term employed here.
not appear to be the explicit addressees in most cases, my analysis also considers the possibility of referee design, in which a speaker does not respond to the speech of a present interlocutor or audience member but to an absent party, and which Bell theorized as being especially prevalent in mass communication (1984). Furthermore, harassers and trolls constitute unratted but known sectors of the imagined audience, potentially affecting stylistic choices. Previous quantitative studies suggest referee design as motivation for style-shifting in both public and personal communication (Hay, Jannedy, & Mendoza-Denton, 1999; Rickford & McNair-Knox, 1994). Issues of audience and referee design with respect to the #YesAllWomen corpus and the observed linguistic practices therein are addressed more thoroughly in chapters 3-5.

10 Features

This study explores certain linguistic features that have been heavily studied with respect to gender and power, including extensive research in online communication. This section of the literature review details previous findings in each of the chosen linguistic feature categories, with particular attention to studies that challenge early assumptions about language and gender. The features most closely examined in this study are use of taboo language, overt politeness strategies, hedging, and use of nonstandard or markedly informal variants. These features were selected for being computationally tractable as well the subject of a large body of literature in the fields of language and gender, language and power, and computer-mediated communication.

10.1 Taboo language

Among Lakoff’s claims about women’s language is that women are less likely to use profanity or swear, instead substituting an avoidance lexicon. Later research on gender and swearing reinforce the generalization that women are assumed to use less profanity overall, and
that swearing and vulgar terms are associated with masculinity and the lower classes. There is some evidence that men use more taboo language on Twitter than women (Bamman, Eisenstein, & Schnoebelen, 2014). However, close examination reveals the more complicated relationship between gender and taboo language use. It appears, for example, that both men and women are more likely to use profanity, especially vulgar terms for physical anatomy or sex acts, in single-sex interactions than in mixed company (Jay & Janschewitz, 2008; Wells, 1989, 1990). Even when controlling for audience, men have been shown to use language judged more offensive, even if overall frequency of profanity is equal between genders (Jay, 1992). Stapleton (2003) found that men are more likely to use profane terms related to anatomy than women, claiming this class of offensive language indexes masculinity.

While men are generally perceived to use more offensive language both in frequency and in level of vulgarity, the setting of the interaction and the relative power or social status of the speakers influences perceptions of appropriateness of profanity. Jay and Janschewitz (2008) show that in general, people in more respected positions were less expected to swear than those of lower status, but that men were more expected to swear if profession was held constant. Setting was also a significant influence on participants’ judgments of likelihood or offensiveness of profanity. A hypothetical college dean was judged less likely in general to swear than a janitor or a student, but was more likely to swear in the dean’s office than elsewhere on campus. According to the survey results, it is more offensive for a student or janitor to swear in the dean’s office than for the dean to swear in the dean’s office, even though students or janitors were more expected to swear overall. When an interaction takes place between speakers of unequal status, the higher-status individual appears to carry a greater degree of freedom to use offensive language, even if that person is expected to do so rarely in their overall habits. This claim is also supported by findings by White (2002), which show that Australian police officers regularly arrest young Aboriginals for using offensive language that the officers frequently use in public
without consequence. These studies reiterate the assumption that profanity use is more associated with masculinity and with relative power.

10.2 Politeness

The connection between politeness and gender was also originally over-simplified in the literature and has since been more deeply explored. Beginning with Lakoff, it was assumed that women exhibit more politeness strategies than men, especially in male-female conversation. This was at one point presumed to be a result of the systemic power disparity between men and women, resulting in women’s behaving deferentially toward men because of their vulnerable position. However, power is constructed and negotiated in interaction, and it is not the case that every man holds institutional power over every woman. Holmes (1988) re-analyzed the gender-politeness relationship as reflective of the affective or referential priorities of speakers. She argued that women employ politeness strategies that serve the interpersonal relationship between interlocutors, and apparent links between gender and politeness are complicated by different discursive goals between interlocutors. In CMC research, Herring (1994; 1992) found that men and women employ different communication styles online, with women’s styles indeed serving rapport between participants in online discourses. She argued that women’s conversational style was characterized by more positive politeness strategies and men’s was more adversarial. Furthermore, she found men tend to accommodate (Giles et al., 1991a) women’s conversational styles in female-dominated discourses, and vice-versa (Herring, 1996). A study of the Enron email corpus (Prabhakaran, Reid, & Rambow, 2014) also found gendered audience effects in overt politeness strategies, but they heavily interacted with the hierarchical structures within the organization of a given email’s author. It was found that subordinates addressing female-dominated conversation groups via email used the most conventional dialog acts (i.e., rapport-building discourse markers and features like salutations or small talk), whereas superiors addressing female-dominated groups used the fewest. These results suggest
that politeness is not best predicted by power or gender, but by a constellation of factors including the assumptions of the discourse surrounding gender and relative authority of the audience.

For the current study, it is important to keep in mind that when overt politeness markers like *please* and *thank you* are tracked, it should not be assumed that the utterances containing them are polite. Literature has distanced from Brown and Levinson’s (1987) conceptualization of politeness as a reified entity that is inherent to an utterance. Politeness is not contained in an isolated utterance but rather emerges through interaction (Arundale, 2010; Mills, 2002) and the study of Twitter data is no exception. The quantitative portions of this study use corpus-analysis methods to count instances of lexical items and makes no claims about the nature of the discourse at a micro level. The subsequent qualitative analyses will supplement the quantitative results to form a more complete picture of the role of politeness strategies in online feminist discourse.

10.3 Hedging

Early studies of gender and hedging presupposed, per Lakoff’s theory of women’s language, that women use language more tentatively, bleaching their statements of force or assertiveness. Later work crucially differentiated between affective hedging and epistemic hedging, pointing out that a hedge may serve different discursive goals, and that these differing goals are more likely to have gender effects than the superficial linguistic variation (Holmes, 1984, 1986; Mohajer & Jan, 2015). Hedging can constitute a positive or negative politeness strategy or express a lack of certainty. Holmes found that the phrase *you know*, for example, was found to be used more often in its epistemic capacity—reducing the certainty of a statement—by men, and in its affective capacity—inviting participation from an interlocutor—by women. However, a follow-up study did not consistently support Holmes’s findings (Dixon & Foster, 1997). Little work has been done on hedging and gender in CMC, but Bamman,
Eisenstein, and Schnoebelen (2014) found a stronger statistical association between “hesitation” and female Twitter users than with male Twitter users, but it is unclear what exactly constituted hesitation in their methodology.

While the connection between gender and hedging appears tenuous, frequent hedging remains associated with powerlessness in perception studies (Bradac & Mulac, 1984; Hosman, 1989). Several production studies have also suggested a correlation between hedging and relative powerlessness. O’Barr and Atkins (1980) found large effects for powerlessness and no gender effects when tracking frequency of hedges in courtrooms. In general, they found that high social status was inversely related to the features of “women’s language” that they were examining, leading them to advocate for re-analysis of women’s language as “powerless language”. Jalilifar and Alavi (2011) also found in a study of politicians’ speech that the political power wielded by the speaker negatively predicted use of hedging. The nature of the speech act itself also seems to be relevant to the relationship between hedging and power. While Brooke and Ng (1986) found no connection between hedges and low influence in small-group interactions, in settings where power disparity is more institutionalized, as in O’Barr and Atkins’s or Jalilifar and Alavi’s work, the inverse relationship between hedging and power is more predictable. As these studies point out, as well as those by Holmes and Dixon and Foster, institutional power often favors men, interfering with study of gender and hedging.

10.4 Nonstandard variants

Another perennial claim in language and gender research is that women are less likely to use nonstandard varieties, orienting their language instead to a prescriptive standard that affords them overt prestige. Labov (1990) argued that women lead some linguistic change, but are less likely to use stable nonstandard variants. Similar findings emerged throughout decades of sociolinguistic research (Cheshire, 2004; Schilling-Estes, 2002). The proposed explanations for this pattern vary, and many are generally compatible with Lakoff’s conceptualization of
women’s language. One prominent argument holds that the disempowered position of women leads them to orient their language toward upward mobility (Eckert & McConnell-Ginet, 1999), which often requires use of a prestige variety. If women hold less financial or institutionalized power than men, they are forced to rely on symbolic power to be taken seriously in interaction. Furthermore, Trudgill (1972) argues that men endow working-class linguistic features with covert prestige while women do not, offering explanation for why women avoid variants that carry stigma or working-class associations.

However, recent work on variation in CMC is not perfectly aligned with the prediction that women generally use more prestigious language. Bamman, Eisenstein, and Schoebelen (2014) found that in a large-scale corpus study, “non-dictionary” words were mathematically more closely associated with female Twitter users than with males. This finding held for both non-pronounceable nonstandard forms (e.g., *omg, lmao*) and pronounceable ones like *lol* or *haha*, including contractions without apostrophes. Furthermore, Squires (2010) argues that internet language is enregistered in opposition to a concept of a “standard” English, with language ideologies surrounding CMC characterized by an “imperative of containment” to appropriate contexts. While this language ideology is not directly tied to gender, Squires found that the participants in her data set attributed internet language to four social categories: youth, females, the lazy or inexperienced, or the uneducated or unintelligent. Enregistered internet lexical items, therefore, are relatively stable nonstandard variants but appear to be ideologically associated with women, and young women in particular, challenging established assumptions about prestige language and gender.

11 Research gap and motivation

Very little scholarly work currently exists at the intersection of the research threads discussed here. While the theory and social implications of feminism have been explored in philosophy, political science, sociology, and anthropology, little research has been done on the
linguistic aspects of feminism. Performance of gender and sexual identity is fairly well researched, as well as linguistic features associated with performance of political ideology (Hall-Lew et al., 2010), but performance of intersecting gender and political identities is far less studied. This gap is especially glaring considering the way feminism and gender equality have dominated popular discourse in recent years.

As discussed above, several non-academic essays have focused on the ways in which language use by women has contributed to their own marginalization, especially in their professional lives. The suggestion that women too often use “just” or apologize and thereby hold themselves back in their careers implies that the responsibility to minimize workplace gender inequality lies with women’s choices in their professional communications. Furthermore, the underlying assumption of such essays is that women must converge to linguistic styles associated with men; if women apologize or hedge their statements, then men necessarily do not, and therefore set the stylistic baseline that women should target. If women want to be treated equally to men at work, these arguments claim, they must use language like men at work. The authors of these essays are advocates for gender equality in the workplace, but partially blame women’s language use for the unsatisfactory rate of progress. This raises one of the primary research questions of this dissertation: when women explicitly aim to empower themselves, do they do so by converging to what are perceived to be men’s linguistic styles, or is the language of contemporary feminism characterized by features other than those that index masculinity? Moreover, when men wish to support gender equality, how do they perform their identities as feminist allies?

The performance of solidarity with another group is an especially relevant issue in the contemporary political climate in the United States. Public Discourses and demonstrations on systemic inequality has led to an expectation for anyone interested in social justice activism to demonstrate their allyship with marginalized groups, whether they identify as members of those groups or not. For example, on January 21, 2017, an estimated five million people marched for
gender equality in cities across the United States, with hundreds of thousands participating in “sister marches” around the world (Chenoweth, 2017). Counts by gender are unknown, but anecdotal reports suggest that a significant portion of the attendees were men (Shavin, 2017). Similarly, the Black Lives Matter movement, born out of criticism of police violence against people of color, is likewise supported by millions of non-black Americans. In early 2017, when the American presidential administration announced intentions to ban residents of certain countries, including refugees, from immigrating to the United States, American citizens rallied across the country to express outrage. When undocumented immigrants who were brought to the United States as children were threatened with deportation, American citizens joined them in protest. In short, it is expected that privileged populations will leverage their privilege to support the efforts of marginalized populations to receive equal treatment. This extends to beneficiaries of male privilege in a misogynistic society.

Crucially for this project, passively agreeing with these movements for equality does not constitute allyship in the eyes of the marginalized communities; deliberate, vocal activism is necessary to identify as an ally of the cause. Furthermore, there appears to be correct and incorrect methods of displaying one’s allyship. If a member of a privileged group is seeking advice on how to support a marginalized one, it is not difficult to find it: a Google search for “how to be a white ally” yields approximately 37 million results; “how to be a male ally” returns more than 21 million. And yet, while there is abundant prescriptive work on the performance of allyship, there is no descriptive scholarly work on the subject. This dissertation examines the linguistic patterns of gender activism, both from a group seeking empowerment for themselves, and one seeking empowerment for others. It draws on prior work on feminism, computer-mediated communication, linguistic accommodation, and perceptions of language’s relationship to gender and power, but presents a new systematic investigation of the linguistic performance of gender social activism against the very complicated backdrop of contemporary politics.
Chapter 3. #YesAllWomen’s language: Women’s style shifting on Twitter

1 Introduction

The recent wave of non-academic metalinguistic essays published online have insisted that the way women speak works against them professionally and socially, including accusations of excessive hedging and apology which supposedly cost a female-presenting speaker authority that should otherwise be afforded to her (Crosley, 2015; Leanse, 2015). Whether women actually engage in these practices more frequently than their male counterparts is unclear, but these pieces do raise an interesting question of the role of language in performed self-empowerment among women seeking gender equality. This section of the study seeks to address this question using data from Twitter to identify linguistic features of the performance of a feminist identity. Crucially, this does not mean determining whether feminists use language differently from other people. Rather, the project approaches the question of feminist linguistic performance with the assumption that feminist identity is negotiated through language and other semiotic systems rather than a static entity that speakers possess and that determines their linguistic choices.

The understanding of identity in this study is akin to what Bucholtz and Hall (2003; 2004) theorize as “intersubjectivity”. The subjects (here, the Twitter users) are both agents and patients of identity construction processes as they use linguistic resources to perform feminist stances that over time develop into linguistic styles (Bucholtz, 2009; Johnstone, 2009; Kiesling, 2009). This chapter attempts to identify how certain linguistic features that have historically been treated as gender-related in linguistics literature play a role in stancetaking and identity production in a feminist Twitter Discourse.

This chapter presents the results of a quantitative study that aims to identify how a single group of speakers\textsuperscript{13} alters their linguistic choices when participating in the Twitter Discourse.

\textsuperscript{13} Speaker here is a shorthand for “language user”; the data are written.
tagged with #YesAllWomen, a marker of an explicitly feminist online conversation that became very popular in April and May 2014. The particular variables examined here are computationally tractable features which have been associated with gender in sociolinguistics literature, including ‘women’s language’ (Lakoff, 1975): politeness markers, hedges, stable nonstandard variants, and profanities. General trends in variationist linguistics suggest that women (or, in some analyses, disempowered speakers), when other factors are ignored, are more likely to use overt affective markers indexing politeness, hedge their statements, use fewer stable nonstandard variants, and use less frequent taboo language.

The women whose language is the subject of this study are knowingly tweeting from a position of overall structural disempowerment (such is the premise of the #YesAllWomen hashtag), but seek to empower themselves and other women with the ultimate goal of subverting a system that perpetuates sexism in contemporary society. Furthermore, their own gender identity is given enhanced salience in this context because their experiences as women are the motivation for their participation in the Discourse. This chapter examines how language use by women on Twitter interacts with these gender and power dynamics. Following an analysis of tweets by a group of 140 women, the findings suggest that the authors use these features differently when participating in this particular feminist Discourse on Twitter, as compared with their contributions to other conversations in the same medium. This intra-speaker variation suggests that these features contribute to a style indexing an online feminist identity.

2 Methods

This chapter presents a quantitative intra-speaker variation study of Twitter data. Phillips (2014) archived over 2.7 million tweets tagged with #YesAllWomen as they were

---

14 See section 9 of the previous chapter for a more complete discussion of these generalizations, as well as relevant citations.

15 The full archive can be found at https://digital.library.unt.edu/ark:/67531/metadc304853/
published between May 25 and June 8, 2014, and published their numerical tweet IDs, making it possible to “rehydrate”\textsuperscript{16} and download the tweets and their metadata through the Twitter API using the open-source script Twarc (Summers, 2014).

For the present study, the #YesAllWomen corpus was reduced to 470,000 unique tweets authored by 203,658 individual accounts after retweets and duplicates were removed\textsuperscript{17}. The Twitter users in the corpus were sorted by the number of tweets they contributed to the dataset. The 200 most active participants (i.e., those who had authored the most tweets in the corpus) were selected for analysis. Each author’s Twitter profile was examined manually to determine that the user was a) still active on Twitter, b) an individual human rather than an organization or automated bot, to the best of my judgment, and c) identified as female. Twitter users who referred to themselves as a “girl” or “woman” or a feminine-marked kinship term (e.g., “wife”, “mother”), or declared “she” and “her” as their personal pronouns in the biography sections of their Twitter profiles were considered self-identified women for the purposes of this study\textsuperscript{18}.

\textsuperscript{16}The Twitter API generally does not permit the downloading of tweet data older than one week, but any publicly available tweet from any time can be downloaded via the rehydration process, by which a numerical tweet ID is passed to the API, which returns the full rehydrated tweet. A rehydrated tweet is in JSON format and includes the text of the tweet itself, when it was published, how many times it was liked or retweeted, and all information included in the author’s profile, among numerous other metadata. See https://developer.twitter.com/en/docs/tweets/data-dictionary/overview/intro-to-tweet-json.html - tweetobject for complete documentation on twitter data. Phillips’s archive of tweet ID numbers, captured as the tweets were first published in 2014, and his subsequent publishing of the IDs allow researchers to access 4-year-old data which would otherwise be very difficult to capture with Twitter developer tools. Of the more than 2.7 million tweet ID numbers Phillips published, approximately 2.5 million of them were still available for rehydration at the time of data collection in Spring 2016.
\textsuperscript{17}Retweets were removed to ensure as much as possible that a tweet was in its corresponding author’s own words.
\textsuperscript{18}At the time of data collection, I made no assumptions regarding the gender identity or sexual orientation of the Twitter users. Trans or nonbinary people who identified as feminine-marked kinship terms or third-person pronouns, if identifiable, were included. It is possible that some of the self-identified women are queer, trans, or nonbinary, and do not indicate as such in their public Twitter profiles. At the time the dataset was archived, 11 of the 140 women in the study included some indication of queerness in their profiles. A more nuanced study of queer identities and their relationship to the # YesAllWomen hashtag is, unfortunately, beyond the scope of this paper.
Profile photos were not considered in the classification process. A total of 141 of the 200 speakers met the specified criteria, one of whom was disqualified because her contributions to the #YesAllWomen discourse consistently argued against the premise of the hashtag, instead insisting that men were the true victims of any systemic sexism and that feminists needed to stop “whining”. Among the remaining group of 140, the number of individually contributed tweets ranged from 29 to 296 per person (M=70.21). These 140 most prolific #YesAllWomen contributors accounted for 9,866 of the 470,000-tweet corpus. This reduced corpus of 9,866 tweets served as one of two experimental corpora for intra-speaker analysis.

The control corpus was composed of tweets posted by the same 140 women. At the time of data collection (April 2016), the 3,200 most recent tweets authored by each speaker were collected through the Twitter API’s “GET statuses/user_timeline” method (this number is limited by Twitter's terms of service) and retweets were removed. This method yielded between 111 and 3,113 unique, original tweets per speaker (M=1932.48), for a total corpus of 272,211 tweets. The control corpus tweets were not screened for subject or hashtag, producing a collection of tweets on a wide variety of topics.

---

19 I hesitate to label this user as a troll, as the precise definition of that term difficult to determine, partially because it is so closely tied to intention (See chapter 2): is a provocative remark made to persuade an antagonist to change their point of view, or to elicit a genuine reaction for one’s amusement, unbeknownst to the antagonized party? It is impossible to know for certain the desired outcome of a given tweet that appears to subvert the general ethos of the #YesAllWomen hashtag. For the purposes of this study, any participant challenging the premise of #YesAllWomen or making statements attacking feminism or feminists in general (rather than a particular point raised by an individual) were dismissed as bad-faith participants. This is not to say such participants were “trolls”, per se, but that their contributions did not meet the criteria for linguistic analysis because of the confounding factors they introduce. One cannot know their intentions for doing so, and rather than speculate on who is challenging feminism as a movement because they enjoy the outraged reactions of self-described feminists or for other reasons, their data were disqualified on the grounds that support for the tenets of the #YesAllWomen movement is a controlled variable for the study.

The frequencies of the linguistic features described above were counted in each corpus, which entailed searching for manually-defined lexical items or phrases. Politeness markers included variations on *please, thank you,* and *sorry.* Profanity search terms included seventeen of the most common vulgar or swear words in English on Facebook (Kirk, 2013). Hedges included *I guess, I think, it seems,* and *sort of,* among others. Finally, nonstandard contractions included items like *wanna* and *gonna,* as well as instances of apostrophe absence in contracted forms like *dont* or *cant.* Common variations in spelling (*gonna* and *gunna,* for example) were included, but cases of typos, unenregistered innovative spellings, or nonstandard capitalization were not. Full lists of lexical items counted for the study can be found in Appendix A.

Once all frequencies-per-tweet of each feature were calculated for each speaker in both corpora, paired-samples t-tests were conducted to compare variation in these four dimensions (politeness markers, hedges, profanity, and nonstandard variants) within speakers. The two corpora were very different in tweet count and the features themselves are all relatively low-frequency, raising concern about sample size as a potential confound in the results. To address this, resampling procedures were employed to ensure that any effects were not due to the differences in sample sizes between the #YesAllWomen corpus and the control corpus.

3 Results

The results of this study show significant intra-speaker variation in all dimensions measured. When participating in the #YesAllWomen Discourse, the female Twitter users use fewer politeness markers, fewer hedges, more taboo lexical items, and fewer nonstandard abbreviations (both colloquial contractions like *gonna* and null apostrophes). The mean occurrences per tweet of each feature category are shown in figure 1. Results by feature are discussed in detail in sections 3.1-3.4.
3.1 Politeness

A paired-samples t-test showed the rate of politeness markers per tweet was significantly lower in the #YesAllWomen corpus (M = .029, SD = .060), than in the control corpus (M = .046, SD = .035), t(139) = -2.98, p < .01, 95% CI (-.028, -.006), d = -.34. Under an assumption that politeness is a marker of deference and relative powerlessness, the women appear to use less deferential language when participating to the #YesAllWomen thread, though earnest markers of positive politeness can be found in the same corpus, as in the following examples:

(1) **Thank you** to all the positive voices and supporters of #YesAllWomen But it’s not just a 1 time hashtag, #ConfrontMisogynyEveryDay

(2) #YesAllWomen **thanks** to all wonderful brave women still sharing their stories.

(3) This hashtag is truly heartbreaking. **I’m sorry** ladies. I offer hugs and support. #YesAllWomen
Example (3) represents the sense in which “sorry” is often used in the #YesAllWomen corpus. When not sarcastic, it is nearly always an expression of sympathy rather than an apology for one’s own actions. “Sorry” in an apologetic or contrite sense was more likely to be found in the control corpus, as in (4):

(4) **Sorry** I’ve been gone so much. Started a new job and it’s been keeping me busy.

Negated sorry (“I’m not sorry” or “#sorrynotsorry”) and sarcastic “sorry” were common in both corpora.

3.2 Hedges

Results showed significantly lower rates of hedges per tweet in the #YesAllWomen corpus (M = .012, SD = .018) than in the control corpus (M = .028, SD = .016), t(139) = -8.259, p < .001, 95% CI (-.019, -.012), d = -.91. This result mirrors the politeness result, in that in a women’s (or powerless) language paradigm hedges are also assumed to mark relative disempowerment and are less frequent in the #YesAllWomen corpus. Consider the following examples from the same Twitter user, the first two of which are from the control corpus and the third, containing emphatic language, is taken from the #YesAllWomen corpus (emphases added):

(5) **Maybe it's just me** but Obama having a pinterest **seems** silly. How to relate with women obsessed with clothes, recipes, funny cat pictures etc

(6) Hey twitter, it’s **kind of** BS that I can block someone and they can still tweet about me. #fthat @twitter (emphasis added)

(7) #YesAllWomen because I’m *REALLY* glad my walk to my parking garage is during daylight/next to the police station. Would b scared otherwise
In fact, this speaker did not use any of the hedging search terms in the #YesAllWomen corpus, but hedged several times in her control sample, illustrating a typical pattern in the results.

3.3 Taboo language

A t-test showed that rates of profanity or vulgarity per tweet were significantly higher in the #YesAllWomen corpus (M = .069, SD = .065) than in the women’s control sample (M = 0.053, SD = .058), t(139) = 2.969, p < .01, 95% CI (.004, .025), d = .25. In light of Stapleton’s (2003) finding that women are more hesitant to use profanity containing anatomical terminology, the two corpora were also compared in that dimension. That test revealed that anatomy-related profanity rates were also significantly higher in the #YesAllWomen corpus (M = .007, SD = .015) than in the comparison corpus (M = .004, SD = .007), t(139) = 2.708, p < .01, 95% CI (0.001, 0.005), d = .24.

However, it is important to note that many of the instances of profanity in the #YesAllWomen corpus, especially in the anatomy and sexism-related terms, are found in directly or indirectly quoted speech. Taboo words may be single or double voiced (Bakhtin, 1981), where single-voiced messages originate with the author, and double-voiced messages are delivered by the author but are implicitly or explicitly another person’s words or messaging. The terms “animator”, “author” 21, and “principle” (Goffman, 1981) are also useful categories here. In Goffman’s framework, the animator is the person uttering the words; the author is the person or people whose words they are, and the principle is the person or people whose views or position is being represented. A single person may constitute one, two, or all three of these roles for a given utterance. Consider the following examples:

21 In this case, I mean “author” in terms of Goffman’s participation roles. Throughout this study, “author” is also used to refer to the Twitter user who posted a given tweet. Where necessary, I disambiguate when I mean it in Goffman’s very specific sense.
(8) To the guy in 8th grade who trapped me on the bus me and said, "Someday, whether you want to or not, we're going to fuck." #YesAllWomen

(9) The default insult for women is "sluts;" they are punished if they choose sex & are punished for refusing to have sex. #YesAllWomen

(10) Because if I don’t smile at you, I’m a bitch. If I smile at someone else, I’m a slut. #YesAllWomen

In (8) above, the quote is direct, and the animator (the person whose Twitter account posted the tweet) is unambiguously not the author or principle, in Goffman’s framing. More frequently, however, as in (9) and (10), the profanity is not embedded in a direct quote, but is still not in the author’s own voice. The person who posted (9) made the point of putting the word sluts in quotation marks, to emphasize that the word is not in her own voice and to distance herself from people who use it. In (10), on the other hand, quotation marks or attribution is not necessary to show that the author uses the terms bitch and slut facetiously. The implication is that men thrust these derogatory labels onto women. The imagined speaker of the insults is not the Twitter user herself, rendering the tweet double-voiced. These are distinctly different uses of profanity than those in (11)-(13) below:

(11) We can't always tell if a man values us or his ability to fuck us, must be prepared for violence if we decline fucking. #YesAllWomen

(12) Because in our family my father made all the rules, and that was a BAD fucking policy #YesAllWomen

(13) #YesAllWomen because being a feminist should be considered merely knowing your facts and recognizing our society’s bullshit.

In these samples, the profanity is in the author’s voice. The use of profanity does not constitute a “sideways glance toward someone else’s hostile word” (Bakhtin 1981), but is the author’s own hostile word, thereby altering the author’s stance toward the subject matter and audience. One should not assume that these two uses of obscenity or profanity (double-voiced and single-voiced) are equivalent or indistinguishable simply because they are treated as such in this
quantitative analysis. It is an important distinction, and the following chapter includes further discussion of the issue. For the purposes of this chapter, it is notable that the occurrence of profanity is statistically higher in the #YesAllWomen corpus. Even in quoted speech, the authors chose not to censor themselves or use a less offensive euphemism. They chose to retain the shock value or emotional weight of the profanities for effect. For that reason, the offensive language in quoted speech was counted in the analysis.

The patterns of self-censorship are equally illuminating here. Consider the following tweets from the control corpus from the same user:

(14) Are you f**king kidding me?? Are we seriously going to lose to a damn 1-4 team, @AtlantaFalcons?? #RiseUp #ATLvsNO

(15) If you don’t want to wear a bra, f**k the patriarchy and don’t wear one. But please don’t use "cancer awareness" as an excuse.

Even when she tweets about gender inequality in (15), this author omits a hashtag and censors her language. This supports an audience design (Bell, 1984) interpretation of the variation, an idea explored further later in this chapter. The hashtag alters the imagined audience of the tweet, which informs stylistic choices. This particular author self-censored in this exact fashion five times in the control corpus, but never did in the #YesAllWomen corpus. In fact, only one instance of “f**k” appeared in the feminist corpus at all, and four of the nine women who used the censored version in the control corpus opted for the uncensored version when tweeting with #YesAllWomen. Other patterns of self-censorship (“sh*t”, “c*nt”, etc.) exhibit similar patterns. The women appear to construct the #YesAllWomen space as more permissive than other metaphorical venues on their Twitter feeds, at least with regard to profane or taboo language. This is a somewhat unexpected finding, as previous work has found that frequency of nonstandard variants is inversely proportional to size of imagined audience on Twitter, and that hashtags are believed to broaden a tweet’s imagined audience (Pavalanathan & Eisenstein,
One may assume that presence of profanity undermines the “standardness” of a tweet and would therefore decrease with a larger imagined audience, but in this case the opposite is true.

As Jay and Janschewitz (2008) and White (2002) found, profanity is less marked when used by a speaker with apparent authority within the particular venue in which talk occurs. The increase in profanity among women here suggests an increased sense of control or power in the imagined #YesAllWomen space. This is made explicit in tweets like the following:

(16) Thanks to the men who only said "#YesAllWomen" with no commentary & had the ability to realize this is not their space. #TakeBackSpace

The stance taken towards men as participants in the conversation—the idea that “this is not [men’s] space”—affords the women linguistic opportunities that may be more marked in digital spaces that are not expressly “women’s”.

3.4 Nonstandard abbreviations

Comparison of nonstandard abbreviations was separated into two separate statistical tests. The first compared frequency of nonstandard contractions like gonna, wanna, and ima. Again, the average rate of instances per tweet was calculated per speaker. These rates were compared in a paired t-test, which revealed significantly less frequent use in the #YesAllWomen corpus (M = .007, SD = .016), than in the control corpus (M = .018, SD = .015), t(139) = -5.795, p < .001, 95% CI (-.014, -.007), d = -.69.

The second paired t-test compared rates of null apostrophes in contracted negative forms like dont, cant, or wont, expressed as a proportion of all negative contractions used, rather than as a rate per tweet. The results were also significant, showing less frequent apostrophe absence in the #YesAllWomen corpus (M = .052, SD = .17) than in the control corpus (M = .156, SD = .14), t(139) = -4.750, p < .001, 95% CI (-.068, -.018), d = -.28. Both
measures of nonstandard abbreviations suggest a closer approximation of standard written English in the contributions to feminist discourse than in non-topic-specific utterances on Twitter. This is at odds with the other results under a canonical women’s language paradigm: women use fewer hedges and politeness markers, more profanity (less stereotypically feminine according to a theory women’s language), and more standard punctuation (more stereotypically feminine) in the #YesAllWomen corpus. The combination of increased profanity and decreased nonstandardness is especially surprising and leads to interesting inferences about the relationship between language and performance of an empowered stance, explored more fully in the discussion section below.

The issue of abbreviation and apostrophe-dropping on Twitter is complicated by Twitter’s character limits. At the time the tweets in this study were published, a tweet could not contain more than 140 characters, so brevity may have been given more consideration that it would in other media. One may be tempted assume that use of nonstandard abbreviations or null-apostrophes would serve to keep a tweet within the character limit and not necessarily reflect the level of formality or standardization in the author’s stylistic choices, and that perhaps the control tweets were simply longer than the #YesAllWomen tweets. To investigate this possible confound, an analysis was conducted to determine how influential character count was on abbreviation strategies. All tweets containing one of the abbreviations (including apostrophe-free contractions) were examined for character count. All tweets not containing any of the abbreviations in question also had their lengths recorded for comparison. The distribution of character counts (ranging from 1-140+) was then plotted for each collection of tweets. The results are shown in figures 2 and 3.

22 Tweets that appear to contain more than 140 characters contained emojis or other non-alphanumeric graphemes that were rendered using multiple characters when imported into the text editor in which the analysis was performed.
These distributions were not significantly different according to a two-tailed t-test. If the tweets containing abbreviations were significantly longer than the tweets not containing any, it
would have supported the prediction that space constraints predict these variants. One may expect an apostrophe to be deleted if that character were needed to complete a thought. However, the tweets containing abbreviations were not approaching the 140-character limit with significantly greater frequency than the tweets with no abbreviations. Of course, this does not prove that use of nonstandard abbreviations is not influenced by Twitter's character limit or that formality of register is the primary predictor of their frequency, but the similarity in the two distributions in figures 1 and 2 suggest that character limit is not the primary motivator for their presence in the data set.

4 Discussion

11.1 Identity tactics

The findings overall suggest a systematic stylistic variation triggered by participation in the #YesAllWomen Discourse. The language in this space is marked by less politeness, more profanity, less hedging, and fewer nonstandard forms. If these results were analyzed within a “women’s language” or “powerless language” framework, the results may be understood as inconsistent. On one hand, when declaring alliance with the feminist movement, the female speakers conform to stereotypically masculine or empowered language expectations with regard to profanity, hedging, and politeness markers. On the other hand, they simultaneously avoid nonstandard forms, which is often assumed to be more characteristic of women (Cheshire, 2004; Labov, 1990; Schilling-Estes, 2002). The results therefore do not neatly conform to a women’s/powerless language paradigm, but do provide an idea of which features correlate with the performance of a feminist persona online. I argue that the women engage in tactics of intersubjectivity in Bucholtz and Hall’s (2004) framing of identity: they engage in differentiation from stereotypical femininity, authentication of their narratives and experiences, and authorization as people entitled to speak publicly on these issues.
The combination of fewer nonstandard forms and more instances of profanity in these findings offers an interesting dissonance. Use of profanity or vulgarity is a stancetaking resource that indirectly indexes gender. Popular opinion holds that it is “unladylike” to use profane language (Gray, 2015; Tyler, 1977), though the observed distribution of taboo language is not clearly gendered. The indirect indexical relationship between swearing and masculinity makes the dramatic increase observed here a possible rejection of conventional expectations for feminine behavior. Furthermore, use of potentially offensive language is often treated as a face-threatening act (Brown & Levinson, 1987; Jay & Janschewitz, 2008). Frequent profanity, therefore, suggests disregard for an audience’s face needs, although a communicative style characteristic of women has been shown to prioritize interlocutors’ face needs, both in person and in computer-mediated discourse (Herring, 1994; 1992; Holmes, 1986, 1989). The variation observed here implies a shift away from a face-affirming stance and style that have been found to be more characteristic of women in past research. This practice differentiates the speakers from stereotypical femininity, indexing a confrontational stance, and thereby constructing an identity in opposition to conventional norms of womanhood. At the same time, vulgar language is a resource for authentication as survivors of sexual violence or harassment. However, the coupling of a significant increase in profanity with a significant decrease in nonstandard variants suggests an unapologetic brazenness that is legitimized by adherence to prescriptive linguistic standards associated with education, high socioeconomic status, and structural power. Frequent use of offensive language serves a provocative, attention-grabbing affective purpose (“my outrage is authentic”), while a standard variety is the style of an educated elite, lending authority to the statements of righteous indignation (“and I’m entitled to it”). This practice contributes to the tactic of authorization, by which a speaker’s identity is constructed as authoritative within a discursive context.

As narrators of their own experiences, the women’s first-hand knowledge should provide them with sufficient authority to be heard and validated, and yet a frequent theme in the thread
is the dismissal of women who have reported harassment and abuse in the workplace and elsewhere. If an institution disadvantages female victims of harassment, often by protecting male perpetrators, it becomes more crucial for women to capitalize on other forms of symbolic power, including language, in order to be taken seriously. This is an important dimension to the observed trend that women use more standard variants than men and is useful for interpreting the results here. Language indirectly indexing upward mobility can serve a speaker who is disenfranchised in some way and may serve to preserve the authority of a speaker who may be violating other taboos.

11.2 Audience

Audience effects are another important dimension to consider here. Twitter, like other digital platforms, renders participation roles difficult to identify. Bell’s (1984) audience design framework is of use here, but naturally was not originally theorized with respect to digital media. Intended addressees are not always made explicit by name, and in a public forum anyone’s presence may or not be ratified according to the speaker. There is evidence for audience design influencing style shifting on Twitter (Nguyen et al., 2015), but this is demonstrated mainly by use of at-replies, in which a particular account is named in a tweet. When the author names an addressee directly, it is easier to examine the effect of addressee on style. In this dataset, women occasionally mention one another specifically, making the addressee explicit:

(17) @TwitterUser723 I’m so sorry to hear that. Take solace that you are not alone. That’s why this hashtag is so important. #YesAllWomen

23 Anonymized
However, in most of the corpus, one can only speculate about the imagined addressees (Marwick & Boyd, 2011), though there are clues in the content of the tweets to suggest who the authors imagine is reading their messages. Consider the following examples:

(18) upside to #YesAllWomen reaffirming we’re not alone, and meeting awesome new Internet People. Thank you.

(19) #YesAllWomen We stand together and not alone [link]

(20) #YesAllWomen because if anything is clear tonight, it’s that we women are not alone in this. We’re here together.

(21) Trolls are engaging in harassment on the #YesAllWomen hash. Let’s band together and affirm those being harassed that their voice DOES matter

(22) Damn I love you guys on this trend. I feel so much better seeing I’m not alone #YesAllWomen

(23) Because #YesAllWomen reminds me that I’m not alone, but that doesn’t make me feel any better.

(24) If even one woman feels relieved to find she’s not alone, #YesAllWomen was worth it.

The frequent use of first-person plural pronouns as in (18)—(21) establishes positions the speakers in unity with the ratified audience members, a stance that suggests that other women are the imagined addressees for many participants in the #YesAllWomen conversation and constructs a collective group identity of women who have shared similar experiences and are expressing a communal feeling of outrage and determination to alter the status-quo. Even when the author acknowledges that non-allies read and use the hashtag, as in (21), she refers to these “trolls” in the third person, while using the first-person plural form “let’s” to align herself with her imagined addressees. Even in at-replies, like (17) above, the addressee is named by her Twitter handle, but this does not exclude the other imagined participants in the thread. The other women monitoring the thread remain imagined (ratified) auditors at the very least and perhaps even imagined addressees, sharing that role with the named account.
Though tweets like (18)-(24) suggest an imagined female, like-minded addressee, the stated objective of the #YesAllWomen movement was to raise awareness about the ubiquity of misogyny (Grinberg, 2014). This would require men to be at least exposed to the discourse, even if they are not the direct addressee. In fact, example (22) above acknowledges that men participate in the #YesAllWomen discourse productively, and commends those that know their place. In doing so, the author ratifies the men’s presence in the imagined space. To at least this speaker, men fill the role of “auditors” in Bell’s schema: their presence is both known and ratified. However, their presence is only ratified under condition of their respect for the metaphorical space as belonging to women. This tweet suggests that the presence or men who offer commentary in addition to the hashtag is unrated, and even unwelcome. The male participants are therefore constructed as (unrated) overhearers if they actively participate, and auditors if they read but do not compose any #YesAllWomen tweets. Men’s participation is not categorically unrated according to the examples cited here. It is assumed that at least some men will see the thread. In fact, it is hoped they will, lest the movement bring about no measurable change in behavior toward women. If women were simply sharing their experiences with the intention of keeping men in the dark about the Discourse, they could not hope to effect any change in the way women are treated. Some men contributed to the #YesAllWomen thread to express their genuine shock and to admit they did not realize the extent to which women lived in fear of assault, which is surely one desirable outcome of the discussion. Furthermore, given the subject of the discourse, men in general are a highly relevant population in the Discourse, including those who do not participate in or even read the thread. The movement seeks to draw attention to gender inequality and everyday sexism perpetrated by men. The broad category of “men” may thus constitute a “referee” in Bell’s framework. Referee design, in which a speaker does not adjust style according to a present interlocutor or audience member but to a relevant absent party, has been suggested by previous quantitative studies as motivation for style shifting (Hay et al., 1999; Rickford & McNair-Knox, 1994).
The understanding of “audience” in this study can therefore be broad or narrow. In the narrowest sense, the women appear to address other women more frequently than they address men. However, Twitter is a public medium and the Discourse seeks to draw attention from a broader, diverse audience. Hashtags in particular draw a broader audience to a person’s tweet, which may contribute to a tendency to use more standard language in hashtagged tweets (Pavalanathan & Eisenstein, 2015). Within the metaphorical space defined by the #YesAllWomen hashtag, the women are offering solidarity and community for victims of gender discrimination and are simultaneously broadcasting this discussion to men or other people who are unaware of the frequency of sexist behavior.

6 Conclusions and implications

The results presented here suggest that an increase in the use of profanity and a decrease in the use of hedging, politeness markers, and nonstandard abbreviations are partially constitutive of a style that emerges among women participating in feminist Discourse on Twitter. Significant intra-speaker variation suggests that these variables have indexical value in the process of feminist identity performance for speakers participating in the #YesAllWomen conversation on Twitter. This is the case regardless of whether each variable is explicitly tied to gender or power in itself or whether the variation is a response to the imagined audience, referee, or the subject matter. Interestingly, the decrease in hedging and overt politeness markers mirror the prescriptive instructions in many metalinguistic opinion pieces about how women can empower themselves linguistically, but the linguistic variation observed here should not be attributed to such prescriptions. Women do not need to have read any Internet thinkpieces about how to use empowering language to internalize ideas about the relationship between linguistic choices and displays of authority or aggression. The language ideologies reflected in popular essays may very well have influenced the variation measured in this study, but not by way of online advice columns. This is discussed in greater detail in Chapter 5.
It should be noted that this study is subject to the same critique of Lakoff’s early work that any findings apply to a narrow demographic of speakers and must not be extrapolated into any broader generalizations. In this case, the speakers are likely young, white, educated adults, Twitter data are notoriously metadata-poor, leaving the demographics of the population almost entirely up to speculation, but Chapter 5 provides some insight into the issue of diversity of the population sample. That said, social critics have targeted online feminism in general (Gleeson, 2017; Loza, 2014) and the #YesAllWomen movement in particular (Heideman, 2014; Rodino-Colocino, 2014) for excluding women of color, women with disabilities, and trans women, thereby erasing the intersectionality of their experiences and promoting only a narrow subset of women. Such criticisms suggest that the makeup of the speaker population here is likely predominantly white and cis-gendered. Any analysis of the #YesAllWomen movement, especially of only a small subset of participants, should not be treated as representative of all feminists, let alone all women.

With these caveats in mind, the results here are compelling despite the constraints of the dataset. The findings reported here hint at how speakers use linguistic resources to perform a gendered feminist identity online. The features studied in this chapter, which have been studied for decades with respect to gender, are not assumed to index gender or power structures directly. However, these results suggest that these particular linguistic resources have possible indexical meanings in the production of feminist identities. These indexical meanings are tied to the context in which they are used, and this includes the perceived (and simultaneously negotiated) identity of the speaker. People who do not identify as women may not use the same features the same way as reported in this chapter, and the features’ indexical values are therefore subject to reinterpretation. It is not sufficient to show these results and assume fixed, established indexical relationships between these linguistic resources and feminist identity. The following chapter examines the same features as used by men in the same context with the goal
of understanding how production of gender identity and production of feminist identity interact and inform each other, and how linguistic practice functions as a site for those processes.
Chapter 4. What about the men?!: Men’s accommodation and performance of feminist allyship

1 Introduction

Feminist literature has explored the various ways in which men can and ought to participate in feminist activism as allies (Almassi, 2015; hooks, 1984; 2000), but the expectations of men’s linguistic practice when expressing their allyship have not specifically been explored. This chapter examines the linguistic practices of self-identified male contributors to the #YesAllWomen Discourse and the process by which they become socially meaningful.

The results reported in the previous chapter suggest a style shift among women triggered by their participation in the #YesAllWomen Discourse. The implications of these results include a prediction that linguistic style is subject to effects of expression of political or social ideology. In this case, expression of feminist solidarity appears to influence linguistic choices of the speakers. However, as discussed in the previous chapters, meaning of linguistic features shifts according to the context in which the features are used, and the identity presented by the previous chapter’s population—self-identified women—is an integral component of whatever social meaning is indexically linked to studied features. The variants used by these women may not point to the same parts of a larger indexical field (Eckert, 2008) when employed by people who do not identify as women.

Several studies suggest that linguistic convergence can underscore attitudinal alignment with an interlocutor, including in short-term interaction (Auer & Hinskens, 2005; Giles et al., 1991b). However, linguistic accommodation can be socially fraught. In fact, there is abundant evidence that people, when expressing solidarity or sympathy with a group or cause that does not obviously include them, are mistrusted when their language approximates the people central to the issue. For example, young white people who identity with hip-hop culture or black social
justice issues may be accused of inauthenticity or insensitive appropriation by their peers when they use features of African-American English or other minoritized language varieties (Cutler, 2010; 1999; Rampton, 2017; Sweetland, 2002); northern politicians are ridiculed for adopting a “fake southern accent” when campaigning in the American South (Cole & Pellicer, 2012).

Considering the stigma of appearing to try too hard to imitate a group with which one is hoping to align oneself, it is not safe to assume that, when expressing solidarity with a feminist movement, men would approximate linguistic styles indexing women or femininity.

However, the linguistic style exhibited by women in the previous study does not appear to index women or femininity. As explained in previous chapters, language ideologies have persistently linked women’s linguistic styles with deference, politeness, prescriptive standards, hedging, and lack of profanity. When participating in the #YesAllWomen Discourse, the women decreased the frequency of markers of politeness and hedging and increased their frequency of taboo language. These shifts suggest that 1) features of stereotypically feminine language are meaningful in the performance of feminist stances, and 2) that women perform these stances in opposition to stereotypical femininity.

One natural follow-up question to the results presented in the previous chapter is whether men expressing solidarity with the #YesAllWomen message exhibit the same style shift observed among women, thereby approximating the stylistic norms established by the group at the center of the Discourse. A large body of literature on conversation accommodation theory predicts that stylistic convergence occurs as a symbol of social solidarity and divergence as a symbol of distance between conversation participants (Auer & Hinskens, 2005; Giles et al., 1991b; Giles & Ogay, 2007). Studies of spoken communication show correlation between entrainment of both lexical (Friedberg, Litman, & Paletz, 2012) and prosodic (Levitan et al., 2012) features as a mark of interpersonal alignment. Linguistic style accommodation has also been demonstrated to occur in Twitter conversations (Danescu-Niculescu-Mizil et al., 2011). If men are displaying their alignment with women as allies eager to speak on their behalf and
support their efforts to raise awareness and criticize patriarchal gender norms, it may be expected that they will converge to the linguistic practices established as an identifiable style by the female participants in the #YesAllWomen Discourse.

Prior work on gender in computer-mediated communication leads to the hypothesis that men converge to women’s discursive styles in this space. Previous by Herring and collaborators (S. Herring, 1994; S. C. Herring, 1992; S. C. Herring, Johnson, & DiBenedetto, 1998) has suggested that men approximate women’s discursive styles in female-dominated discourses, which is consistent with tenets of accommodation theory posited by Giles, Coupland, and Coupland (1991). We can assume the #YesAllWomen discourse is a female-dominated space, although Twitter metadata are insufficient to confirm that assumption completely. This chapter investigates this prediction. Does either the relatively overwhelming representation of women in this very gender-focused Discourse or a desire to express political solidarity with them a trigger for stylistic convergence among self-professed male allies?

2 Replicating an intra-speaker variation study with a new population

To test this hypothesis, I attempted to replicate the study described in Chapter 3 among a population of men who contributed actively to the 2014 #YesAllWomen corpus. In order to collect a sample comparable to the sample of tweets by women, I followed analogous selection criteria for self-identified gender (masculine-marked kindship terms or identifiers like “man”, “guy” or “dude” in biography sections or posted tweets). I then selected the self-identified men who had tweeted at least as many times as the woman with the smallest tweet contribution in the female #YesAllWomen corpus. The number of #YesAllWomen tweets authored by a single person in the previous chapter was 29, so only men who tweeted at least 29 times in the master #YesAllWomen corpus were eligible. This method yielded only six men who met the criteria, compared to 140 women in the first study. After considering collecting a corpus of male-authored tweets on a collection of several other explicitly feminist hashtags
(#StandWithWomen, #HeForShe, etc), it was determined that the soundest comparison would be between subsets of the #YesAllWomen corpus, and it is preferable to reduce the average sample size from each author than to include different hashtags from different periods of time.

3 Data

A population of 68 men who tweeted with the #YesAllWomen hashtag at least six times became the focus of this follow-up intra-speaker variation study. It is worth noting that the 68 men were those who remained after apparent trolls were manually removed from the sample. Over 130 men initially met the sample size criterion of six #YesAllWomen tweets, but roughly 25% of them were no long active on twitter, and an additional 25% were only contributing antagonistic posts. The resulting #YesAllWomen dataset contained a total of 1179 hashtagged tweets authored by 68 men, each contributing between six and 125 tweets to the corpus (M=17). The Twitter profiles of the same 68 men were scraped to build a control corpus, following the same methods as those used for the previous chapter (see chapter 3 for a detailed explanation of data collection and research methods). The control corpus contained 133,215 tweets after retweets and duplicates were removed, with each author having written 186 and 3194 tweets on a number of various topics that were not screened for hashtag or topic, published any time before the time of data collection (July 2017).

4 Results

4.1 Summary

In the first experiment, a direct analog to the experiment presented in the previous chapter, no significant intra-speaker variation was observed among the population of 68 men in the dimensions of profanity, politeness markers, hedging, or apostrophe deletion. Only frequencies of stable nonstandard abbreviations (wanna, gonna, etc.) showed significant differences between the #YesAllWomen corpus and the control. Notably, while results were far
less robust, the direction of change in feature frequency mirrors that of the women in each dimension. Section 5 of this chapter contains a more complete discussion of cross-gender comparison. Figure 4 below presents the mean frequencies in each corpus authored by the men. A detailed account of results follows.

Figure 4. Mean frequency of feature per tweet, men

4.2 Politeness

Paired-sample t-tests showed that politeness frequencies per tweet were lower in the \#YesAllWomen corpus (M=.025) than in the control corpus (M=.044), but the differences do not account for a statistically significant change (p > .1).

4.3 Hedging
Men’s frequency of hedging decreased when participating in the #YesAllWomen discourse (M=.016) than in the control corpus (M=.030), but this difference also does not constitute a statistically significant one (p > .1).

4.4 Profanity

The men appeared to use more profanity in the #YesAllWomen corpus (M=.051) than in the control corpus (M=.045), but only very by a very slight margin and the results do not approach significance (p > .1).

4.5 Nonstandard abbreviations

The frequency of stable nonstandard variants was the only feature to show statistically significant differences between the two corpora, with frequencies in the #YesAllWomen corpus (M=.004, SD=.017) significantly lower than in the control dataset (M=.018, SD=.022), t(67) = -2.37, p < .05.

4.6 Apostrophe deletion

Frequency of apostrophe deletion in negative contractions was, like other forms of nonstandard abbreviation, also lower in the men’s #YesAllWomen corpus (M=.060) than in the control corpus (M=.161), but the difference is not statistically significant (p > .1).

5 Resampling: comparing with a reduced women’s corpus

While these results show slight or insignificant differences in the dimensions studied, the direction of change mirrors the women’s in each case. The insignificance of the results may suggest that the features are not evaluated equally when uttered by men as when uttered by women, leading men not to conform to the stylistic norms established by women. However, the
fact that the direction of change is the same in each case, but to a lesser degree among men, raises the question of whether men are in fact converging to women’s discursive practices, but in a way that escapes identification in the given sample.

Sample size presents a challenge: it is possible that the results observed among the female participants would not have been as robust (or significant at all) if the women only numbered 68. To investigate this possibility, resampling techniques were employed in the women’s corpora to replicate exactly the men’s corpora. 68 of the 140 women were randomly selected and a number of their tweets that corresponded to the number of tweets authored by each man were also selected at random. For example, if Man 1 contributed 30 tweets, a set of 30 random tweets were sampled from Woman 1, and so on. This yielded two corpora of women’s tweets that match the men’s corpora in number of authors and number of tweets per author. The feature-counting methods were then re-employed in these reduced datasets. The entire procedure of collecting a reduced dataset and running the analysis was conducted 10 times. The results are shown in table 1.

Table 1. Results of comparison between reduced corpora by women

<table>
<thead>
<tr>
<th>Feature</th>
<th>Mean frequency in #YesAllWomen</th>
<th>Mean of mean frequencies in control</th>
<th>Mean t(67)</th>
<th>Mean p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profanity</td>
<td>0.060</td>
<td>0.052</td>
<td>0.498</td>
<td>0.477</td>
</tr>
<tr>
<td>Politeness</td>
<td>0.019</td>
<td>0.039</td>
<td>-2.060</td>
<td>0.128</td>
</tr>
<tr>
<td>Hedges</td>
<td>0.013</td>
<td>0.032</td>
<td>-2.258</td>
<td>0.085</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>0.006</td>
<td>0.119</td>
<td>-2.04</td>
<td>0.103</td>
</tr>
<tr>
<td>Apostrophe=0</td>
<td>0.064</td>
<td>0.098</td>
<td>-1.145</td>
<td>0.330</td>
</tr>
</tbody>
</table>

For one of the five features studied, hedging, the mean p-value of the observed differences between the #YesAllWomen corpus and the control corpus approaches significance
These results are not anywhere near as robust as the results found in the dataset authored by the full population of 140 women, though slightly more robust than the findings from the 68 men reported in the previous section. This suggests the possibility that the differences observed in the female population are larger or more consistent than those in the male population, but that conclusions are better drawn from larger sample sizes, and the 68 men studied here do not provide a large enough dataset to allow any statistically supported inferences.

6 Interspeaker comparisons

As another follow-up study to the intra-speaker variation analysis, each possible pair of corpora were compared using comparable methods as those described above. Sample sizes were equalized across number of authors and across tweets per author. This yielded four sets of results in addition to the two intra-speaker results already reported: women’s #YesAllWomen vs. men’s #YesAllWomen, women’s control vs. men’s control, women’s #YesAllWomen vs men’s control, and women’s control vs. men’s #YesAllWomen. In each case, no significant differences were identified in any of the feature categories analyzed. However, use of profanity was higher in the women’s #YesAllWomen corpus (M = .065) than in the men’s #YesAllWomen corpus (M = .028) to an extent approaching significance (.05 < p < .1). Furthermore, apostrophe deletion was also more frequent in the women’s control corpus (M = .015) than in the men’s control corpus (M = .007) to a nearly significant degree (.05 < p < .1). Considering the other measure of nonstandardness, common nonstandard abbreviations, showed no such difference in the same pair of corpora, the implications of this finding are unclear.

Figure 4.2 below shows the mean frequencies across all corpora and all feature categories. Note the similar distributions, particularly in politeness, hedges, and nonstandard contractions, where differences in sample sizes resulted in very different t-statistics. Additionally, the striking difference in apostrophe absence is an interesting case. Only the
difference in women’s frequencies was significant, but the more puzzling finding is that women overall had far higher rates of apostrophe deletion. The reason for this is unclear. It was considered that women were more likely to use hashtags containing contractions (punctuation is not permitted in hashtags), but that did not turn out to be the case upon closer inspection. This is a particularly unexpected difference when the rates of other nonstandard abbreviations were so similar across populations.

Figure 5. Average feature frequencies per tweet, all corpora

7 Qualitative analysis

The small sample size of men renders robust statistical results elusive, but a relatively small data set is not an intrinsically uninteresting one. A quantitative natural language processing research method can easily abstract over phenomena that a closer qualitative analytical method may reveal. The following sections explore some of these phenomena.
7.1 The role of silence

The difficulty encountered in collecting comparably-sized corpora authored by men and women is worth discussing in itself. While the #YesAllWomen hashtag was coined and popularized by women, it sought to draw attention to harassment at the hands of men and to educate men about the hazards of navigating the world as a woman. It was not necessarily intended to exclude men from the conversation. It is not a forgone conclusion that men would not participate in the discussion, and yet few of them appear to have. In this case, it is useful to consider the role of silence as an ideological entity and its position in this particular Discourse. Basso’s (1970) analysis of silence as a sociolinguistic variable suggests that silence has specific social meaning among the Western Apache. Of course, while it is unreasonable to extrapolate discursive norms of the Western Apache in the 1970s to contemporary participation frameworks in 21st-Century Twitter, Basso’s exploration of what motivates silence is of value here. Men have a reputation for contributing to online discourses when they are not invited to do so (Bridges, 2017), so why are they not in this case?

Gal (1989) points out the emphasis of silence in feminist theory, where silence is interpreted as a symbol of passivity or powerlessness. Several social scientists, including Basso, have theorized silence as indexically linked to ideas beyond meekness or powerlessness. Sattel (1983), for example argues that silence can be symbolic of masculine authority or intimidation. Silence in this study is further complicated by the digitally mediated space. One cannot be visibly present in the imagined #YesAllWomen space and silent at the same time. To make your presence known is to participate.

Researchers from the Pew Research Center investigated the “spiral of silence” in social media to examine more closely when people do or do not prefer to participate in political or other sensitive discussions on and offline (Hampton et al., 2014). They found that people were less likely overall to share their opinions on social media platforms than in person, and that for
those unwilling to discuss controversial topics in person (the article used Edward Snowden’s leaks of classified information as the example used in the survey), online platforms did not provide an alternate venue where they were more comfortable. In other words, some people were willing to discuss politics in person but not online, but far fewer were willing to discuss politics online but not in person. The survey also found that people were less likely to post about politics online if they felt they did not share the opinions of the majority of their followers.

This offers one possible explanation for why men seemed so scarce in the #YesAllWomen data set: that few men agreed with the message of the #YesAllWomen hashtag and were thus unlikely to express their opinions, assuming the majority of people reading the thread felt differently from them. If this were the case, this study would suggest that men who agree with the #YesAllWomen agenda were rare, and those captured in the data collection process for this research are a substantial, representative sample. However, Twitter is rife with disagreement, antagonism, and even harassment. The subject of the Pew study, Edward Snowden and privacy rights of American Citizens, is less charged by identity politics and perhaps therefore less attractive to Internet trolls than gender equality-focused discourses. Furthermore, Herring (1994) demonstrated that men are far more likely to deliberately disagree or engage in argument online than women. The Pew study made no analysis along gender lines, so it is unclear if hesitating to express an unpopular opinion is more common among women than men—Herring’s findings would predict so. My own data collection methods support Herring’s generalization. Of 141 women who used the #YesAllWomen hashtag several times, only one individual was doing so to provoke and antagonize other people posting with the tag. Among men, over 25 percent of men who used the hashtag repeatedly did so in a trolling fashion.

The silence observed among male allies in this data is unlikely to be due to a fear that their opinions will be unpopular among their Twitter audience. Twitter users regularly post unpopular opinions, often simply because of—not despite—their unpopularity. I propose an alternative explanation for this particular spiral of silence among men: that they believe silence
is the preferred method of expressing solidarity. This ideology is expressed in several tweets published by women, including (1) and (2) below:

(1) Thanks to the men who only said "#YesAllWomen" with no commentary & had the ability to realize this is not their space. #TakeBackSpace.

(2) Hey, fellas. If you don't think "male privilege," is a thing, please go read #YesAllWomen. Don't interrupt. Reflect - feel bad - do better.

These examples suggest that the women on this thread are not interested in hearing men’s professions of allyship and support. The first offers a method of making their supportive presence known without offering uninvited commentary: use the hashtag but do not write any accompanying content in the tweet. This is a potential solution to the challenge of signaling one’s presence in the Discourse without invading the imagined women’s space. Example (2) tells instructs men not to “interrupt”. This is an interesting directive, given that turn-taking on Twitter is nothing like that of copresent talk, and “interruption” in its technical sense, is not possible. One may receive replies to a tweet when they are in the process of writing follow-up tweets, but a series of tweets by a single author can be “threaded” together, indicating that the thread constitutes a single turn, and replies to individual tweets within that turn are not visible as interruptions. The author of example (2) frames any utterance by men in the space as an “interruption”, suggesting the #YesAllWomen Discourse has its own participation and turn-taking rules. This construction mens’ utterances as interruption also echoes an assumption that men interrupt women more than vice-versa, an often investigated claim but one which is not strictly factual, based on empirical studies (James & Clarke, 1993). The apparent belief that all of men’s contributions constitute interruption contributes to the language and gender ideologies at work in the Discourse, which are further examined in the next chapter.

The imagined audience of these examples (1) and (2) obviously includes men—men as a group are explicitly picked out an addressee in both cases—but their tweets are not intended as
an invitation to men to participate in the discourse. Men’s actions are required for any significant change in the norms that the women object to, but their active participation is more valuable offline (or “in real life”) than on Twitter. Men’s objection to sexist behavior as they witness it is believed to be more effective than objections by women (Cihangir, Barreto, & Ellemers, 2014; Drury & Kaiser, 2014), and increased frequency of such behavior as a desired outcome is impossible without men’s exposure to the #YesAllWomen discussion. Twitter provided a venue for women to share their stories and express support and solidarity, but also to educate men who may have been ignorant to the realities of the female experience. However, it is apparent that while men were intended as an audience, their participation is not necessarily ratified by the discussion. In other words, men should be listening but not talking. Interestingly (and somewhat paradoxically), this sentiment is echoed by some of the men:

(3) Dear fellow men. Please take some time to read the #YesAllWomen Tweets. We have much to learn. Don't comment; just read.

(4) Because, if as a man, you feel compelled to attempt to debate or defend the points being made by #YesAllWomen. Read. Think. Learn. Don't Speak.

These men take a stance toward other male audience members that is similar to that taken by the women in examples (1) and (2) above. The ideology of a good ally’s silence appears to have transcended gender in the studied population, as both men and women instruct well-meaning men to keep quiet, as in (3) and (4) above, in which the male authors identify other men as the intended addressee of the messages. This provides a possible motivation for the scarcity of men’s tweets in support of #YesAllWomen, but some men who espouse the opinion that the best male feminist is a silent male feminist display inconsistencies between self-professed language ideology and language practices. For example, the man who tweeted (3) above contributed to the #YesAllWomen corpus at least 26 times. While it is entirely plausible that the belief in silence as a virtue among allies prevented more men from contributing to the
Discourse, in some cases it was in direct contradiction with the practices those men engaged in. For those who did participate in the #YesAllWomen Discourse, there is no evidence that their utterance length tended to be any shorter than women’s; t-tests revealed no significant difference between the length of the men’s #YesAllWomen tweets and the women’s, measure either in character or word count.

The expectation of minimal participation is complicated by a competing ideology of allyship, best represented in example (5), taken from the men’s corpus:

(5) There aren’t nearly enough men contributing to this hashtag and that’s why it needs to exist. #YesAllWomen

Example (5) illustrates the tension inherent in men’s participation. Silence is both held up as preferred ally behavior and is also interpreted as a sign of apathy. This duality places men in a kind of double-bind similar to that which has been described as an impossible set of paradoxical expectations that women face regarding their language use (Lakoff, 1973, 2003; Tannen, 2008). As Lakoff (1973) described it, a woman “is damned if she does, damned if she doesn’t. If she refuses to talk like a lady, she is ridiculed and subjected to criticism as unfeminine; if she does learn, she is ridiculed as unable to think clearly, unable to take part in a serious discussion: in some sense, as less than fully human.” The double-bind that men find themselves in here is not equivalent but is defined by similarly mutually-exclusive expectations. Well-meaning men, in this context, are expected simultaneously to keep quiet and absorb information without offering an opinion, and to voice their support, without which any measurable change would be unrealistic to expect. How do men attempt to strike this delicate balance? The evidence suggests they often do so by targeting one another as implied or explicit addressees, rather than addressing women directly. An analysis of addressees in a 300-tweet subset of the men’s corpus reveals that while the majority of men’s #YesAllWomen tweets (205
of 300) had no apparent addressee at all, those that did were far more likely to address other men, as in (6) (original emphasis):

(6) MEN, instead of getting angry because women are talking about rape, get PISSED because WE still haven't learned to STOP. #YesAllWomen

Tweets like (6) accounted for 75 tweets in the 300-tweet sample, while tweets that addressed women were only 20. It is notable that 13 of those 20 were explicitly marked with overt politeness markers that were studied for the quantitative analysis, for example:

(7) #YesAllWomen ~ because I am learning so much from wise teachers, thank you to all those sharing and teaching. Some of us are listening

(8) Ladies please hang in there as best as you can and forgive these boys who call themselves men. There ARE a few of us who care. #YesAllWomen

In addressing other men and tagging their messages with #YesAllWomen, these men are performing their allyship while drawing attention from a female-dominated imagined audience, but by explicitly not addressing women, they avoid accusation of intruding on a female-controlled Discourse. This may serve to establish a deferential stance vis-a-vis the female auditors in the conversation. The presence of women in this conversation is almost certainly ratified to the male contributors, but the context is such that addressing the women directly may be interpreted as a face-threatening act. When they do address women, as in (7) or (8), the overt politeness strategies mitigate potential face threats.

Aside from the issue of silence vs. participation, additional cases of ideological mismatch emerged in the data in which linguistic practice undermined self-professed language ideology. For example, consider (9) and (10) from the same man:
This author expresses disapproval of sexist or aggressive language in response to women who have done nothing to warrant it. His description of women being called “bitches” when they get angry suggests that he believes women are entitled their anger and a sexist slur dismisses their legitimate experience, thereby locally authenticating and authorizing their identities as empowered women. His second tweet reinforces this ideology, implying that men are unduly aggressive or dismissive of women when they are doing nothing that would be perceived as marked if committed by a man. Consider this position against the following tweets, authored by the same person in the control group:

(11) my grandma really is a bitter bitch
(12) i fucking came into the orientation room to eat lunch alone and some fucking girl came and sat in here fuck you

Example (11) directly contradicts this man’s expressed disapproval of “bitch” as a slur directed at women, while (12) describes a scenario in which a women existed in the same public space he was occupying, prompting him to use sexist taboo language about her (presumably instead of directly to her, though one cannot know what transpired in the event itself). These examples are themselves performative, and constitute a different genre from the #YesAllWomen Discourse, where social meaning is constructed and interpreted differently. I do not cite these examples to draw attention to this Twitter user’s hypocrisy or to call the authenticity of his allyship into question. Rather, I use this to illustrate the stark differences in social meanings attached to certain linguistic variables and the broad indexical field they constitute. As this author uses quotation marks and double-voicing to define his stance toward the word “bitch”
because of what it indexes within the #YesAllWomen Discourse, his stance toward the same lexical items outside of the Discourse contributes to the construction of a different linguistic style that contains different systems of indexical meaning.

7.2 Taboo language and participation roles

Although the overall frequency of taboo language use among men was not significantly greater in the #YesAllWomen corpus, a closer analysis of the tweets themselves reveals a significant shift in how profanity and vulgarity were used and in the meanings attached to them. The previous chapter included a brief overview of voicing or participation roles in profanity use. Both the women’s and men’s #YesAllWomen corpora were rich in examples of taboo language used in a double-voiced context (Bakhtin, 1981), in which the person who posted the tweet was not necessarily the author or animator (Goffman, 1981) of the vulgar expression. While profanity use was overall significantly more frequent in the women’s #YesAllWomen corpus, much of the profanity was actually in embedded quotes, explicitly, as in (13) or implied, as in (14), both from the women’s corpus:

(13) To the guy in 8th grade who trapped me on the bus me and said, "Someday, whether you want to or not, we're going to fuck." #YesAllWomen

(14) Because if I don't smile at you, I'm a bitch. If I smile at someone else, I'm a slut. #YesAllWomen

Use of double-voicing of profanity was also prevalent among the men’s contributions to the #YesAllWomen corpus. Of the 56 profanities used by men in the #YesAllWomen corpus, 35 were single-voiced, in which the Twitter user was author, animator, and principal of the utterance containing the taboo terms, as in the following examples:
Because it's not like women are naturally better at being "nurturing". It's that most men don't even fucking try. #YesAllWomen

I also think it's bullshit we have to point out that men have a wife/daughter/sister/mother to care about sexual assault. #YesAllWomen

#YesAllWomen Because when *I* reply to douchebags trolling this hashtag, they don't say jack shit to *me*. Hmmm...

Compare these with the tweets in which the principal (person whose viewpoint is represented) or author (person whose words they are) of the taboo word is not the man who posted the tweet, as in (18) and (19), two of the 21 tweets fitting this description:

because if you have sex and are a woman, you're a slut, but if you have sex and are a man you're a stud #YesAllWomen

because I can't have a bunch of friends that are girls unless "I'm gay" or "we're fucking"

The male animator of (18) implied double-voicing in the same way the female animator of (14) did. The person whose words and viewpoint are being represented (i.e., the author and principle of the offensive words) is a nonspecific sexist person in both cases, and the utterance conveys that without use of punctuation or explicitly introducing another participant. The man who posted (19) used quotation marks to differentiate himself from the author and principle of the vulgarity. Of the profanities in the male-authored #YesAllWomen tweets, 62.5% were single voiced and 37.5% were double-voiced. Of the double-voiced profanities, 57% (12 out of 21) were explicitly embedded in quotative punctuation.

A comparison with a sample from the control corpus suggests that double-voicing is far more common in #YesAllWomen tweets than in tweets on other topics. I created a reduced control corpus composed of up to 10 random profanity-containing tweets taken from each of the
27 men who tweeted profanity in the #YesAllWomen corpus (N = 259 tweets). These 259 tweets contained 301 separate use of profanity, 283 of which were single-voiced (94%), and 18 of which were double-voiced (6%). Of the 18 double-voiced profanities, only 6 (33.3%) were explicitly marked with punctuation indicating an embedded quote.

This difference in profanity use suggests that double-voicing is a valuable discursive resource for men expressing their allyship in the #YesAllWomen conversation. When a Twitter user includes taboo language in his #YesAllWomen tweet, he is most closely associated with it if he is the author, principal, and animator of the expression. Placing the principal role onto another party, named or unnamed, distances himself from the potentially offensive phrase:

(20) Dear dude commenters: saying you want equality for all and then calling me a pussy traitor is sort of contradictory. Just FYI. #YesAllWomen

Quotation marks expand this distance and remove any possible ambiguity regarding authorship:

(21) because men like @TwitterUser1 use pejoratives like "dumb cunt" to try to get their point across when defending patriarchy #YesAllWomen.

8. Men-women profanity comparison

Compared to men, the women were more likely to swear when using the #YesAllWomen tag, both by measure of overall frequency of taboo lexical items (women’s mean frequency = .069 instances per tweet, men’s = .051), and by speaker. Of the 140 women whose tweets were analyzed, 120 used profanity or obscenity in their #YesAllWomen tweets (86%). Of the 68 men, 27 used taboo language at least once when tweeting with #YesAllWomen (40%). However, in an analysis of the taboo language that was used in each population, there were not stark differences in voicing or participation role tendencies. Women’s obscenities were single-voiced 56% of the
time and double-voiced 44% of the time, while men’s were single-voiced 62.5% and double-voiced 37.5% of the time. Of the double-voiced profanities in the women’s #YesAllWomen corpus, 45% of them were embedded in quotation marks, compared to 57% of the men’s profanities whose principal differed from the animator. Men, therefore, are more likely to take stances that mark the distance between themselves and the principals of offensive language by punctuation.

9 Discussion

The quantitative results discussed here suggest that conclusions are not easily drawn from a small sample, but that the style-shifting observed in the self-identified female population is also possibly observed among self-proclaimed male feminist allies, though perhaps to a lesser degree. The peripheral role of men in this particular discourse may be related to the underwhelming indication of intra-speaker variation. It is possible that central figures in the discussion are more prone to marked style-shifting while more peripheral participants are less likely to use linguistic markers of ingroup status. Another possible factor is the level of enregisterment of the linguistic style that male feminist allies exhibit. As men are less active contributors to this and similar discourses, there is a lower volume of text to which participants may be entrained. This leaves the male feminist register as less well defined or recognizable than those observed in the female contributors’ tweets. If stances constitute styles when used repeatedly over time (Bucholtz, 2009; Johnstone, 2009; Kiesling, 2009), the stances need to be deployed enough times over a long enough period, with relatively consistent indexical values, in order for a style to emerge. Men may simply not have achieved that in the time the #YesAllWomen Discourse transpired.

Though this analysis does not provide mathematically compelling evidence of style-shifting on the part of the male contributors to the corpus, there are apparent audience design- or accommodation-influenced linguistic practices, as evidenced by the mixed-methods analysis.
It appears that men’s ideological relationship to the particular linguistic features studied here is not necessarily identical to women’s, due in part to the power and gender politics at play in the discourse.

Men in this Discourse find themselves bound by a mutually-exclusive set of expectations: to speak up as allies and to sit down and let women speak. As they perform their own feminist stances and construct their own feminist identities, they face expectations that differ from those that women face. While women distance themselves from stereotypical femininity through their use of vulgarity or more direct language, men are not positioning themselves vis-à-vis feminine stereotypes, even if they are converging to the women’s linguistic style. Stylistic convergence in this case would constitute an adequation tactic (Bucholtz & Hall, 2004), marking solidarity or similarity with the women who dominate the Discourse. Similarly, their use of vulgarity in the Discourse differs from their use of vulgarity in other Discourses with respect to voicing (Bakhtin, 1981) and participation roles (Goffman, 1981). The use of double-voiced vulgarity in their performance of a feminist identity differentiates them from a stereotypically aggressive male and aligns them more readily with their female audience.

The construction of the audience also serves as a mechanism to maneuver around the delicate constraints these men encounter. The presence of women is certainly ratified (Bell, 1984), but they are frequently not addressed directly in the men’s utterances. Instead, they serve as relevant referees while men explicitly indicate that their imagined addressees are other men—to be more precise, men who are less engaged in feminism than they are.

The performance of feminist masculinity in this digitally mediated space challenges reified notions of power and gender identity. The power negotiations between men and women in the #YesAllWomen Discourse do not conform to assumptions that men dominate women here, or that the reverse is strictly true. Rather, gender and power dynamics are negotiated through linguistic practice, which is filtered through the language and gender ideologies of the
participants. The connections between these ideologies and the linguistic practices of the Discourse participants, conscious or unconscious, are the focus of the following chapter.
Chapter 5. “I’m not here to police people’s language”: Ideologies of gender and language among Twitter feminists

1. Introduction

The previous chapters report on quantitative studies of Twitter data produced by a particularly active subset of people who participated in the #YesAllWomen discourse on Twitter in 2014. The quantitative results suggest that self-identified women style-shift when participating in such a Discourse, and there is also some evidence that men do not necessarily converge to the established style that characterizes women’s performance of a feminist identity, but that they do show evidence of linguistic accommodation (Giles et al., 1991b; Giles & Ogay, 2007), likely informed by imagined addressee or other elements of audience design (Bell, 1984). A few follow-up research questions emerge from these findings. First, the data were very impoverished of metadata, and any known biographical information about the authors was limited to what was written in the short biography section on each person’s Twitter profile. The metadata-poor nature of the corpus invited an investigation into the authors’ social identities and backgrounds and the interactions between social variables and the authors’ linguistic practices.

Furthermore, the observed variation is productively analyzed with respect to language attitudes and ideologies, which offer insights into the motivations and social meaning of language variation. For the purposes of this study, I treat “language attitude” as a phenomenon on the level of the individual, and “language ideology” as a shared system of meaning that people use to make sense of linguistic difference, following Irvine and Gal (Irvine & Gal, 2000; Kroskrity, 2000). Ideologies and attitudes inform each other, and both interact with linguistic production and perception. In the case of this study, it is not expected that the observed variation is a deliberate or conscious decision, but the author’s attitudes about language, gender, and feminism may inform their behavior both within and outside of online feminist Discourse.
The motivating questions for this chapter include: What do these people imagine linguistic features of feminist Discourse to be? Do they have prescriptive opinions about language use in performance of a feminist identity, or in general? How do these attitudes coalesce into language ideologies that inform production and perception of linguistic variation?

To shed light on these issues as well as the demographic properties of the population studied, a survey was devised and distributed to as many participants in the #YesAllWomen corpus as possible. Distribution began with those represented in the analyzed corpora, and then expanded to those in the larger corpus who contributed actively but not as prolifically as the 206 whose tweets were computationally analyzed. Therefore, while the sample population analyzed in this chapter does not constitute a perfect one-to-one representation of the population who produced the analyzed Twitter data, the selection criteria (explained in detail in the “Recruitment” section below) increase the likelihood that conclusions drawn from one population can be productively applied to the other.

The results reveal a set of consistent themes illustrating coherent ideologies of language, gender, and socio-political advocacy. Certain language ideologies (most notably, features of Lakoff’s “women’s language” paradigm) emerge as salient, even in anti-establishment populations like this one, but are accompanied by distinct subversive attitudes about linguistic prescriptivism. These attitudes coalesce into a language ideology that has interesting implications for the observed intra-speaker variation reported in this project.

Linguistic accommodation and audience design is well documented on Twitter (Danescu-Niculescu-Mizil et al., 2011; Tamburrini, Cinnirella, Jansen, & Bryden, 2015) and social media is a rich site for research on identity practices (Bamman et al., 2014; Marwick & Boyd, 2011; Miller, 2011; Vaisman, 2014). These previous studies analyze linguistic practice, either quantitatively or qualitatively, but many lack testimony from the online language users themselves in such a way that elucidates how they view the process of performing gender. Recent trends in language, gender, and sexuality research have focused on frameworks that
interpret language as the site of the production of identity (Cameron, 1997, 2005, 2008; Talbot, 2010), similar to emergentist frameworks, which are concerned with how speakers recruit possible indexical values of linguistic variants to adopt meaningful stances in context (Levon & Mendes, 2015). For this study, I subscribe to these approaches while also treating identity production as relational act, in which people are both agents and patients of social processes that create meaningful categories (Bucholtz & Hall, 2003, 2004). Building on these theoretical bases, this study shows some of the ways people perform stances and styles on Twitter that are indexically linked to understandings of gender identity and sociopolitical activism.

For this chapter, the processes by which speakers connect social meaning to linguistic form in production of socially meaningful identities are revealed through self-reported biographical information and conscious attitudes that provide the speakers’ context in which they adopt their stances. The previous two chapters of this dissertation have provided evidence that participants in the #YesAllwomen Discourse recruit variables to adopt a feminist stance, but their motivations and the ideological processes by which they do so remain opaque. This chapter attempts to clarify some of those processes using data provided voluntarily and directly by #YesAllWomen contributors about their own conceptualizations of language, gender, and online feminism.

2 Recruitment and demographics

This section presents the process of identifying and recruiting survey participants and their self-reported demographic metadata.

2.1 Recruitment

The online survey was developed using the survey platform Qualtrics and published via an anonymous hyperlink. Participant recruitment was initially limited to Twitter users represented
in the experimental corpus analyzed for the two previous chapters, as the study sought to examine the language attitudes of a population alongside their linguistic practices. To that end, I extracted all Twitter handles (i.e., usernames) from the raw JSON data and composed a new Twitter list from them. Twitter lists are collections of Twitter users that one can curate to organize their social network thematically or by nature of their relationship. For example, a Twitter user may choose to label a group of other users as “linguists”, or “journalists”, or “comedians”, navigate to that list to see only what the people on that list have posted recently. Crucially, one can populate Twitter lists with people one does not follow, as long as they have a public account. I was therefore able to put everyone in the #YesAllWomen corpora (both men and women) into a Twitter list without personally subscribing to their Twitter feeds. Creation of the list was automated through the “add_list_member” method of the Twitter API.24

Since the time at which the Twitter corpora was collected from 206 people, 25 of them have privatized their profiles or left Twitter entirely, leaving 181 eligible for membership of the newly created list. Once the list was created, I used an application by a company called Followers_DM that was developed as a marketing tool that automates direct messaging to all members of a Twitter list.25 The application sent an identical message containing a brief explanation of my research goals and a link to the survey to everyone on the list. However, the message was only able to reach those whose direct messaging feature is open to everyone. Of the 181 on the list, 41 had their messages open to everyone. The 41 delivered messages yielded 16 completed survey responses, submitted between October 16 and November 5, 2018.

In an effort to collect more data, recruitment efforts were extended to a larger group of #YesAllWomen participants, beyond the group of 206 that were analyzed quantitatively for the previous chapters. The threshold for inclusion was lowered to 20 tweets contributed to the larger 500,000-tweet #YesAllWomen corpus (whereas the lowest number of original tweets

---

24 See full documentation at http://docs.tweepy.org/en/v3.5.0/api.html
25 https://www.followersdm.com/send-dm-twitter-list/
authored by a single female contributor to the experimental corpus was 29). This yielded a total of 1,243 Twitter handles, but many of those were no longer active on Twitter or had public profiles. The list of approximately 800 public Twitter profiles was reduced to 655 eligible survey participants. Eligibility was determined by similar criteria as were used for inclusion in the main study. That is, a Twitter account needed to be unambiguously an individual person rather than an organization, participating in good faith and not as a recreational antagonist (troll) or parody account, active on twitter in the 18 months preceding survey participant recruitment, and tweeting primarily in English. Trolls and parody accounts were easily identified. Examples (1) and (2) below, respectively the profile biography and a sample tweet by one account, offer an illustrative example of what a troll account looks like:

(1) gay trans-gendered trans-racial women, deal with it

(2) How many feminists does it take to mow the law? I don’t know, they neither go outside or do work #YesAllWomen

Of the 655 eligible list members, 124 had their direct messaging open to me. Another 16 of those 124 completed and submitted surveys between November 2, 2018 and January 25, 2019. An additional response was submitted through recruitment in a Facebook group in which a link to the survey was posted along with an appeal to people who remember tweeting with #YesAllWomen or similar hashtags, bringing the total number of competed survey responses to 33. The one respondent recruited through Facebook self-identified as a #YesAllWomen contributor on Twitter.

2.2 Demographics of population

The demographics section of the survey contained questions on gender, race, age, nationality, education level, and language background. Questions asking for gender, race, and language background were open text boxes, while education background and age group were
elicited with drop-down menus. The results of the demographic questionnaire section of the survey provide important context when considering the scope and generalizability of the quantitative results.

2.2.1 Gender

Of the 33 respondents, 28 provided their gender identity, which was elicited via unstructured response; the participants were given the prompt “What is your gender identity?” and a blank text box. 18 self-reported as a variant of “female” or “woman”. Three of these specified cisgender status, but the others did not mention cis- or transgender identity in that field. Of the remaining ten who disclosed their gender identity, five identified as male. None specified trans or cis-gender identity, but one self-reported as a “straight white married male”, despite the prompt only mentioning gender. Five of the respondents identified as nonbinary in some capacity, two providing only that term and two elaborating with additional details, seen in (3) and (4) below. Another reported (5), which was coded as Nonbinary for the purposes of the analysis:

(3) (NB, 25-34) I’m nonbinary (gender) and non-binary (how I relate to gender as a concept). Not gender-nonbinary, not gender-nonconforming. Nonbinary. Demigirl, specifically. I also accept genderfluid/genderqueer.

(4) (NB, 18-24) demiboy/demigirl and agender

(5) (NB, 45-54) My gender is oppressive, my identity is irrelevant, my sex is female.

There are a few elements of these results worth discussion in relation to the other results chapters. First, the surveyed population sample is overwhelmingly female, which is likely fairly representative of the #YesAllWomen corpus overall. My analyzed corpus was composed of

---

26 Each example response is in this chapter annotated with the self-reported age and gender identity (coded as Male, Female, or Nonbinary) of the participant who provided that response.
approximately two-thirds self-identified women, and the small surveyed population has a similar composition. The most notable difference is the that non-female-identifying third of the experimental corpus was all self-identified men, which were specifically selected for, whereas the survey respondents included equal numbers of self-identified male and non-binary participants. In the full #YesAllWomen corpus, only 396 of the 209,700 unique authors identified as nonbinary, genderqueer, agender, or with the personal pronouns “they/them” in their Twitter bios. This is very few relative to the 1,994 self-identified male Twitter bios. Nonbinary people are therefore overrepresented in the surveyed population, relative to male- or female-identifying participants.

2.2.2 Age

Participant age was elicited with a drop-down menu containing options of age ranges. See Figure 5.1 for distribution of self-reported ages and gender identities.

![Gender ID by age group](image)

*Figure 6. Survey respondents’ gender identity by age group*
2.2.3 Ethnicity

27 of the survey participants provided a self-description of their race or ethnicity. This question, like the question on gender identity, was in the form of an unstructured text box rather than a multiple-choice question. Of the 27 answers, 19 were simply “white” or “Caucasian”\(^{27}\) with one other naming specific European nations of ancestral origin, foregrounding region or nation over broad racial category. Another said “Arab”, exhibiting a similar choice. One identified as African-American, one as Afro-Latino, one as southeast Asian, and three as some kind of “mix” (“Eurasian mix”, “Mexican/white”, and “white/native American”).

People who identify as white or partly white dominate the sample, which is unsurprising and also likely fairly representative of the broader #YesAllWomen discourse. This is consistent with criticisms of the discourse as underrepresenting intersectional perspectives of women and non-binary people (Heideman, 2014).

2.2.4 Nationality

28 people reported their nationality (prompt: “What is your nationality (e.g., Canadian)?”). 25 self-identified as American, two as Canadian and one is reportedly “Sri Lankan/Australian”. This is unsurprising given that tweeting primarily in English was a selection criterion for recruitment.

The survey also asked for current location of residence. 27 report living somewhere in the United States or Canada, and one resides in Australia.

2.2.5 Language use

\(^{27}\)“Caucasian” here is interpreted as the American idea of “white”, although as one respondent points out, using “Caucasian”, which at one time only referred to people from the Caucasus region of western Asia, to refer to all people who fit the American category of “white” is problematic.
When asked what languages they grew up speaking at home, 22 respondents said only English (one of whom specified American English). Six reported English and at least one other code. Additional codes listed were Sinhala, French, German, Portuguese, Spanish, African-American Vernacular English, and Greek. 5 declined to respond.

The 22 who reported speaking only English at home growing up also report currently using only English in informal speaking and digital contexts in their daily lives. One reported using English and AAVE, one reported using English and Greek, and one reported using English and Spanish. These findings suggest a relatively homogeneous group with regard to language use and native or near-native English use among participants in the #YesAllWomen corpus.

2.2.6 Education

The highest level of education completed by the survey respondents was distributed across categories of post-secondary schooling. See Table 2 for the results of the drop-down survey.

<table>
<thead>
<tr>
<th>Highest level of education completed</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Some high school, no diploma</td>
<td>0</td>
</tr>
<tr>
<td>High school graduate, diploma or the equivalent (for example: GED)</td>
<td>0</td>
</tr>
<tr>
<td>Some college credit, no degree</td>
<td>6</td>
</tr>
<tr>
<td>Trade/technical/vocational training</td>
<td>0</td>
</tr>
<tr>
<td>Associate degree</td>
<td>5</td>
</tr>
<tr>
<td>Bachelor's degree</td>
<td>8</td>
</tr>
<tr>
<td>Master's degree</td>
<td>8</td>
</tr>
<tr>
<td>Professional degree</td>
<td>0</td>
</tr>
<tr>
<td>Doctorate degree</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
</tr>
</tbody>
</table>

This small sample suggests that while participants in the #YesAllWomen Discourse are likely to have received at least some post-secondary education, one should not make broad generalizations about the education level of Twitter feminists, considering this small group
displays considerable diversity in educational background. However, it is notable that the sample contained no one with only a high school diploma or less. Figure 5.2 shows a representation of education level by age group, indicating that some college credit without a degree and a bachelor’s degree only are concentrated among the younger participants. Associates degrees, however, are distributed across the older age categories, and overall education level appears not appear to be strictly a function of age in this population.

Figure 5.2: Education level by age group

![Education level by age group](image)

*Figure 7. Survey respondents’ education level by age group*

### 2.2.7 Political leanings

Of the 28 participants who disclosed a description of their political leanings, most listed some variation of “left”, “progressive”, or “democrat”. Three said some variant of “independent”. One each reported each of the following: “center”, “constitutional republican”, fiscal
conservative, social liberal”, “way left of any damn party”, and “girl, everything is a mess.

Leftish?”

Given the results of the demographic questionnaire, one may be tempted to project a sort of ideal persona onto the participants in the discourse I am studying: a young, American, white, cisgender, monolingual, left-leaning, educated woman. While this may be accurate on average, I do not wish to use these survey results as justification for abstracting over the diversity in the population. The internal variation in lived experiences should not be erased in the name of neat generalizations or trends.

3 Survey design

The survey covered topics of feminist identity, gendered language styles, reactions to particular linguistic stimuli informed by this study's corpus analysis, and demographic information. Response types were a combination of unstructured text boxes, multiple-choice questions, and Likert scale-style questions. A complete copy of the survey appears in the appendix.

The first set of questions addressed conscious attitudes toward feminism and gender-based language variation, as well as thoughts on use of #YesAllWomen and related feminist hashtags. The second section contained 12 minimal pairs of tweets, each differing only in one of the features tracked in the previous two chapters—overt markers of politeness, nonstandard abbreviated variants, hedges, or taboo language. The participants were asked to compare the two tweets in each pair to one another along a number of axes (see the examples in section 6 below). They were then prompted with brief passages and asked to rate them on a numbered scale with respect to various attributes. The following sections explore the emergent themes of the survey responses.
4 Language and gender attitudes

The first section of the survey addressed conscious language and gender attitudes, asking about feminist identity, beliefs about feminist hashtags, and impressions of gendered language styles.

4.1 Gender and feminist hashtag use

The participants were asked to rate their agreement with each of the following prompts:

A) It is helpful to the feminist cause when women use feminist hashtags like #YesAllWomen or #MeToo to draw attention to the issue and share experiences

B) It is helpful to the feminist cause when men use feminist hashtags like #YesAllWomen or #MeToo to draw attention to the issue and share experiences

Their responses illuminate some of the findings reported in the previous two chapters. The average rating for prompt A was 4.45, and the average rating for prompt B was 3.58, a significant difference (t=3.29, p < .01). The difference in support between men’s and women’s use of feminist hashtags contextualizes the paucity of data provided by self-identified male feminists discussed in chapter 4. Even the male participants in the survey were less enthusiastic in their expressed support for men to use such hashtags. The population’s attitudes toward hashtag use are conditioned by gender of participant, underlying the importance of gender in the way political engagement is performed in the experimental corpus. This implies that participants in the Discourse have different relationships to the Discourse itself, as well as to other participants, by virtue of their gender, with women being more readily ratified as participants than men.

4.2 Ideology of gendered language
Results of the Likert scale-style language attitude questions suggest ideologies consistent with the same public narratives about language and gender that are manifested in prescriptivist essays about how women speak. At the same time, elements of the results are at odds with the observed linguistic practices in the #YesAllWomen corpus. These questions are a mix of open-ended questions and 6-point Likert-scales with prompts asking for an agreement rating for a given statement, for example:

![Likert-scale-style survey question](image)

*Figure 8. Likert-scale-style survey question*

The survey respondents overall believed that gender-based differences language use exist, and they generally converged on similar assumptions about them. The mean score on the prompt “There are gendered styles of using language” was 4.06 (where 0 = “strongly disagree” and 5 = “strongly agree”). Follow-up questions asked about the existence of “stereotypical” men’s or women’s styles of language use. The agreement ratings with the statement “There is a stereotypical way that men use language” (M = 3.84) was not significantly different from the response to the analogous question about stereotypical women’s language use (M = 3.68) when compared using a paired two-sample t-test. Each respondent was relatively consistent in their attitude toward men’s vs. women’s language stereotypes. That is, if someone believes there is a language use stereotype for women they are likely to believe the same for men, and vice-versa.
The average responses suggest an overall acknowledgement that stereotypes exist, albeit an unenthusiastic one.

After quantitatively rating how much they agree with the existence of stereotypically gendered language, respondents were asked to provide characterizations or examples of stereotypical men’s or women’s linguistic styles. As will be made clear, the responses contain clear echoes of Lakoff’s 1975 women’s language paradigm as well as later investigations into her assertions. These include interpretations of gender difference as a reflection of male dominance (O’Barr & Atkins, 1980; Spender, 1990) and as socially learned difference (Cameron, 1992a; Talbot, 2010; Tannen, 1990). They also exhibit parallels with Kiesling’s (2007) posited main cultural discourses of masculinity in the United States: gender difference, heterosexism, dominance, and male solidarity.

These assumptions have been reproduced recently in the guise of popular nonscientific essays (Crosley, 2015; Leanse, 2015). The contemporary essays echo Lakoff’s work in their capacity as assuming a “deficit’ model of language and gender, according to which women’s linguistic practice reinforces and perpetuates sexist social structures (Kiesling, 2007; Talbot, 2010). They assume strong correlation between gender and women’s language features including hedging and apology and urge women to self-police their linguistic practice with the goal of leveling the power disparities that gave rise to the imagined differences in the first place. The participants in this study overwhelmingly reproduced some of these narratives of gender differences in language use (e.g., belief that women hedge or apologize more frequently than men), but they did not express preference for stereotypical men’s language over women’s. The main themes of the results are explored in the following sections.

4.3 Certainty, confidence, report

A prevalent theme in the responses was an emphasis on confidence, aggression, or certainty in men’s utterances. Some framed this characterization in contrast to their ideas about
how women speak, explicitly or implicitly. Consider the use of comparative language in the following examples (emphasis added), where “than women” is seemingly implied:

(6) (M, 25-34) Men, conditioned by patriarchy, tend to talk more loudly than woman; they communicate their ideas with more confidence and arrogance; men are less likely to talk over other people.

(7) (unknown) There's a lot more "you should" and "you need to" instead "have you thought about" or "you could try"

(8) (NB, 45-54) I live in more than one culture, and there is variation here, but men tend to omit “I think”, ”maybe”, “don’t you agree”, and other consensus-seeking phrases and to state their opinion as fact. They speak more directly, more often, and more boldly.

These examples speak to an assumption that women are more likely to hedge statements, whether to weaken epistemic force (as Lakoff argued) or as interpersonal connection or warmth (Holmes, 1987, 1995). They reproduce Kiesling’s (2007) understanding of gender difference (the assumption that gender is binary and that men and women are naturally and categorically different) and dominance (the expectation that men are more aggressive or in control, particularly when compared to women). These are two central tenants of hegemonic masculinity, which Kiesling argues is the primary or most valued or expected incarnation of masculinity as ideologized in the United States today. Hegemonic masculinity emerges as a salient theme in the survey results, an interesting pattern in a group of self-selected vocal feminists.

In addition to the assumption that men tend to speak with more certainty or confidence relative to women, the theme of tension between rapport and report (Michel, 1994; Tannen, 1994) arose frequently, though not using that exact framework. The assumption that men prioritize delivery of information (so-called “report”) over interpersonal or face-preservation
functions of language ("rapport"), while women do the opposite, is evident in responses like the following:

(9) (F, 55-64) [Men] use results-oriented focus

(10) (F, 25-34) Men speak without thinking about how their words impact others, because they have never had to fear that impact.

(11)(F, 25-34) Men are typically encouraged--and often forced--into aggressive and clipped styles of communication that neglect vulnerability and compassion.

(12) (M, 35-44) I think men get stereotyped using language that is more blunt and to the point, maybe even slightly aggressive.

(13) (F, 25-34) The way men speak, stereotypically, is devoid of emotion and helpful adjectives. My experience has been that they believe they’re more logical by speaking without pertinent details. Man: "Cathy is upset." Woman: "Cathy seemed a bit distant and was distraught over a loss."

Examples (9) through (13) illustrate the tendency to characterize men’s language as more direct, less hedged, with greater epistemic force, and with shorter utterances\(^\text{28}\), at the expense of politeness, which aligns with previous theories on gender difference in communication styles, including work by Tannen (1994) and Michel (1994). The rapport/report framework, though problematized in language and gender literature, emerges as a salient assumption in these results. Critics of the framework argue that a gender-as-cultural-difference model, of which rapport-report tension is a central component, both ignores structural power and social value inequalities and perpetuates a deterministic framework of gender as a predictor of linguistic practice (Cameron, 2005; Talbot, 2010). However, these answers contain clues that their authors are aware of complications of power and socialization. Answer (11) alludes to

\(^{28}\) The related belief that men’s utterances are briefer because they contain fewer affective discourse markers is, as previously discussed, not borne out in this study’s Twitter data, nor is it consistently borne out in other CMD studies; men have been shown to produce longer utterances than women in text-based online communication (Androutsopoulos, 2006; Panyametheekul & Herring, 2007)
socialization practices that condition men to speak a certain way, which does not suggest a particularly deterministic idea of the relationship between language and gender. Furthermore, example (10) explicitly includes a proposed motivation for men’s linguistic practice—that they “don’t need to fear the impact” of speaking without reflecting on how one’s words “affect others”. This sentiment speaks directly to Cameron’s and Talbot’s critiques that modeling gender differences as merely cultural differences in communication styles ignores the sociopolitical realities of those differences—namely, that styles that index masculinity are valorized and that styles that index femininity are not.

The generalizations about men’s communication styles also included two references to homosociality (Cameron, 2014; Kiesling, 2007; 2005) among men:

(14) (F, 35-44) Maybe the general comradery that men share

(15) (F, 65-74) Also, many men will avoid conflict with other men and not disagree or stop a friend who makes sexist jokes. The ‘pack loyalty’ seems more important than loyalty to their ethics.

These responses are folk generalizations about male homosociality and the tenet of male solidarity, the expectation that men as an ideological entity prefer to bond and spend time with other men rather than women. This belief coexists and stands somewhat in contrast with assumptions that men categorically prioritize information transfer over the face needs of interlocutors (Holmes, 1995; Mills, 2002), but is qualified in both cases by the assumption of a male audience. In these comments, men are assumed to prioritize their relationships with other men over ethics regarding sexism.

4.4 Apology, hesitation, rapport

The generalizations about women’s language overwhelmingly pointed to a tendency toward prioritizing the audience’s face needs and foregrounding rapport over report, in
Tannen’s parlance. Many highlighted apologizing, and many focused on the downplaying of one’s certainty, expertise, or conviction:

(16) (NB, 25-34) Women have a habit of apologizing more, speaking in more uncertain terms even when they are sure of things, and taking on emotional labor because it is expected of them.

(17) (F, 35-44) Women sound more questioning when we speak. You may hear it referred to as "up talking." It makes statements sound like questions. There is a lot of apologizing.

(18) (F, 55-64) Self deprecating: "I don't know if we'd want to do this, but I think we should...." Hesitant to come out and make a point, for fear others will disagree.

(19) (NB, 45-54) They tend to speak more inclusively and often less directly, with more invitations for other people’s opinions, more indications they could be wrong.

Overall, generalizations about forcefulness, hedging, and politeness were by far the most salient attitudes when analyzed with respect to the linguistic features that are the focus of this study. Ideas about apologizing and other explicit politeness strategies were limited to discussion specifically about women, suggesting that men’s politeness and apology practices are the unmarked case and women’s practices are divergent from a baseline, an inference that one participant explicitly articulated, though couched in a discussion of stereotypes rather than personal perception:

(20) (F, 25-34) I think stereotypically male ways of speaking are considered the "default" more or less.

It is notable that only one respondent framed stereotypical men’s language in this way. Linguistic styles perceived as masculine are often interpreted as the unmarked norm in mainstream Discourses (Cameron & Kulick, 2005; Kiesling, 2007), but that does not emerge here as a common thread in the responses. While the survey elicited generalizations about
men’s stereotypical language use as well as women’s, likely leading participants away from treating supposedly masculine variants as the unmarked case, there were no participants who offered generalizations about women’s linguistic practice but not men’s. With the exception of example (20), the expressed attitudes in this survey notably diverged from the persistent treatment of men’s language as a bundle of unmarked default linguistic features.

Moreover, men were not only the subject of stereotyping, but frequently negatively targeted for their linguistic practices, including by male respondents. This constitutes a notable divergence from patriarchal evaluations of linguistic norms that emerge in prescriptivist deficit-based understandings of language and gender. Among these responses there was a distinct sense of disdain for stereotypically masculine linguistic styling, contrary to public narratives of stereotypically feminine language being negatively evaluated. This is most pronounced in the many answers about men’s stereotypical language use that focused on sexist content of men’s speech, as in examples (21) and (22):

(21)  (F, 45-54) Calling women emotional or hysterical when having discussions or debates, as well as calling them cunt, bitch, slut, etc. Men speak in masculine terms, “be a man”, “man up”, “stop acting like a pussy”

(22)  (M, 35-44) In private, men making sexualize jokes about women as objects. In public, interrupting and mansplaining concepts to their female coworkers.

Another important site of divergence from the canonical women’s language paradigm is the total absence of any mention of formality, slang, or grammatical “correctness”. This is particularly striking considering that element of women’s language is among the most documented principles of language and gender variation  (Cameron, 2008; Cheshire, 2004; Labov, 1990) However, no participant raised the issue, while generalizations about politeness, hedging, and prioritizing face or informal transfer were nearly ubiquitous in the results.
4.5 Observation or stereotype?

It is important to note here that the question was about impressions of *stereotypes* of
gendered language, not their impressions of observed language practices. That said, the framing
of the answers rarely differentiated between those ideas, and more often generalizations were
treated as observed reality rather than spurious stereotypical assumptions. Compare the
following two examples:

(23) (F, 25-34) The closest stereotype I can think of for women is saying omg or
apologizing too often

(24) (NB, 25-34) Women have a habit of apologizing more, speaking in more
uncertain terms even when they are sure of things, and taking on emotional labor
because it is expected of them.

Example (23) is one of only three answers that mention stereotypes or perception without
editorializing on whether those stereotypes were accurate. All other responses referred to
generalizations or personal observations. Though they were asked about stereotypes, a
remarkable number of participants’ answers included evidentiality markers like “in my
experience”, as in (13) above or expletive constructions like “there is”, as in (7) or (17), all
repeated here for reference:

(7) (unknown) **There's** a lot more "you should" and "you need to" instead "have you
thought about" or "you could try"

(13) (F, 25-34) The way men speak, stereotypically, is devoid of emotion and helpful
adjectives. **My experience has been** that they believe they're more logical by
speaking without pertinent details. Man: "Cathy is upset." Woman: "Cathy seemed a
bit distant and was distraught over a loss."

(17) (F, 35-44) **Women sound more** questioning when we speak. You may hear it
referred to as "up talking." It makes statements sound like questions. **There is** a lot
of apologizing.
4.6 “They” or “we”?

It is also revealing to examine the stance a participant takes vis-à-vis the generalizations they discuss. Use of a first person plural serves to construct a group identity while use of a third-person plural may distance the speaker from a group whose membership they may be eligible for (De Cillia, Reisigl, & Wodak, 1999). The responses contain various strategies serving different interpersonal metafunctions (Halliday & Matthiessen, 2014), i.e., how the authors position themselves relative to the group they are referring to. In the following two cases, the authors invoke their own linguistic practices (emphasis added):

(25) I think women are taught to behave "young and innocent" when they are not. I remember the first time I saw the word "infantilize". I slowly started to notice myself speak with a false high pitch to my voice and put a fake innocent look on my face.

(26) A lot of the way women talk is coded to be gentle and apologetic. I sometimes find myself falling into roles that are more sympathetic or apologetic. Despite my own temper, I often find myself feeling like I am obligated or expected to be the peace maker. I often wonder if I am coming across as too masculine and consciously make the effort to be more feminine.

Note the construction “I notice myself” or “I find myself”, suggesting an involuntary linguistic practice that is reaches the level of conscious awareness after production. This relates to the idea of agency that emerges again in section 4.7 below.

Three other women alternated between first and third person when talking about how women speak (emphasis added):

(27) (F, 45-54) I think women tend to speak in a way that’s more submissive and non-confrontational. We’re more likely to apologize, to ask rather than order. We speak in a way that’s less aggressive, more assertive.

(28) (F, 25-34) We women are socialized to think carefully through every word, because there is the constant threat of what could happen to us should a listener interpret our words a certain way. i.e. When attempting to be friendly, women have to take great measures to minimize the possibility that a male listener will be "lead on" or read romantic/sexual intention into our words, because if a man feels that we "promised" him affection via our interaction (and then "deny" him when we only
wanted to be polite or nice), then there is the ever-present threat of physical force and violence.

(29)  (F, 35-44) **Women** sound more questioning when **we** speak. You may hear it referred to as "up talking." It makes statements sound like questions. There is a lot of apologizing.

Only one man used any first-person pronoun when discussing stereotypical men’s language. All other answers were strictly in the third person when discussing generalizations about gendered language use, establishing a stance that distances the participant from what they describe as broad generalizations. This was true of all of the nonbinary participants, but frequent even among those who self-identified elsewhere as male or female:

(30)  (F, 25-34) Women are usually more careful when **they** speak, especially in a social sense.

(31)  (M, 35-44) In private, men making sexualize jokes about women as objects. In public, interrupting and mansplaining concepts to **their** female coworkers.

These examples are authored by people who self-identify as part of the category they are describing, but they employ only third-person pronouns when providing their descriptions, distancing themselves from the generalizations they make. Example (31), in addition to exhibiting only third-person assertions about men, reflects assumptions about the frequency of interruptions by men, the reality of which is far more nuanced and subject to numerous discursive variables (Grob, Meyers, & Schuh, 1997; James & Clarke, 1993).

4.7 Perceptions of motivation for gendered language use

A salient difference in perceptions of men’s and women’s language is the projection of underlying cause for a particular language style. When discussing women’s linguistic practices, the respondents were far more likely to frame communicative choices as a response to external pressures. Consider the following examples:
(32)  (F, 35-34) **We women are socialized** to think carefully through every word, because there is the constant threat of what could happen to us should a listener interpret our words a certain way. i.e. When attempting to be friendly, women have to take great measures to minimize the possibility that a male listener will be "lead on" or read romantic/sexual intention into our words, because if a man feels that we "promised" him affection via our interaction (and then "deny" him when we only wanted to be polite or nice), then there is the ever-present threat of physical force and violence.

(33)  (Unknown) **Women have been conditioned** to compete with other women and to be unquestionably loyal to their assigned males, and thusly use gender-based language and ideals to appear more feminine to appease their males or to demean other women they have **been conditioned** to despise.

(34)  (NB, 25-34) I think as a general rule, **women are taught** to be kinder, more compassionate, more accommodating, and less, overall, dehumanizing.

(35)  (M, 25-34) **Women, conditioned by patriarchy**, tend not to speak up as much as men; women are more likely to apologize when talking with others.

The above examples emphasize women’s language practices as a reflex of societal conditions, whether women are “taught”, “conditioned”, or “socialized”, while men more often purportedly act with no mention of practice as a reflex of external pressures. The perpetrators of this supposed conditioning are left undefined; “the patriarchy” is the most specific external force mentioned. While men were also depicted in two responses as acting in response to “conditioning”, more frequently their linguistic habits were discussed as a result of internal motivation on the speaker’s part, as in (36) below, where men are posited to “like to” speak with more certainty. Most often, men’s language practices were presented as fact without any causality or motivation, unlike women’s, whose communicative tendencies were often attributed to societal pressures. The recurring stance toward women’s practice as less agentive than men’s recalls Butler’s (2004) treatment of agency as paradoxically determined by an uncontrolled social world that constitutes the self. The paradox lies in the idea that ideological constructs define agency and its use. Carter (2007) similarly argues that agency does not exist outside of ideology or social limitations but is rather circumscribed by and exercised within them. Here, women are ideologically constructed as speaking a certain way in response to external factors,
thereby removing their agentive motivations for their choices. Men are conversely attributed agency where women are not. (36) below offers an example of generalization about men’s language use as motivated by the speaker’s desires (although the author also incorporates biological determinism into her answer):

(36) (F, 25-34) They like to use more certainty in their language. It would seem to be a function of testosterone, based on the research.29

The respondents, therefore, are fully aware of the stereotypes of gendered language, and overwhelmingly express belief that the stereotypes are grounded in truth, but that they are motivated and maintained by sexist social structures. The ideology of women’s language as a result of an oppressive patriarchal system is aligned with Lakoff’s earliest theories of women’s language but also with reframing of the language-gender relationship as a reflection of male-dominant social structures (Spender, 1990). The construction of agency in contemporary feminist Discourse is a clearly a rich site for future study, though thorough analysis falls outside the scope of this project.

4.8 Taboo language and feminist style

Regarding other features studied in this dissertation, only two people mentioned profanity, vulgarity, or other types of taboo language. One only mentioned that swearing was considered “unladylike” but professed no personal opinion on the matter. The other made a passing reference to profanity but also expressed discomfort at the premise of the question:

(37) (NB, 35-44) I think they use more deferential language and don’t always want to swear. Am uncomfortable with the gender descriptions here.

29 Example (5) was the only comment alluding to biological determinism of any kind, and introduces questions about influences of nature and nurture on linguistic practices. This study assumes foundationally that language practices are social and therefore socially conditioned, and that understanding of biological sex is equally mediated by social and ideological processes (Butler, 1990, 1993a, 2004).
Two answers mentioned high rising terminal (as “uptalk”) (Warren, 2016), and one mentioned creaky voice (as “vocal fry”) (Henton, 1986; Wolk, Abdelli-Beruh, & Slavin, 2012; Yuasa, 2010), suggesting there is some awareness of acoustic stereotypes but they are may be less salient than lexical, syntactic, or discursive variables.

Only one answer mentioned feminist discourse, making different generalizations based on feminist identity and context:

(38) (F., 35-44) In certain situations, feminists may be more likely to be confrontational; in certain situations, women who are not feminists may be more likely to be confrontational with other women.

This suggests that, depending on context, “feminists” employ different communication styles as compared to non-feminists. Non-feminists are posited to be more confrontational vis-à-vis other women, suggesting that the generalization of feminists as more confrontational assumes a male (or non-female) audience. This construction of “feminism” as an immutable quality that women either possess or do not is in contradiction with this study’s treatment of feminism as a stance that is performed and encoded through linguistic and other semiotic processes.

5 Features

Four questions in the survey asked for explicit attitudes about what linguistic features, of the ones studied here, ought to be deployed when discussing gender equality in a public sphere. The questions were elicited, like the questions about whether gendered language styles exist, using a 6-point Likert scale, where 0 = “strongly disagree” and 5 = “strongly agree”. The prompts and their mean responses are as follows:

(39) When speaking out about gender inequality, people should avoid taboo or inappropriate language (swearing, vulgarity, etc.) (M = 1.52)
When speaking out about gender inequality, people should avoid words or phrases that suggest uncertainty, like "maybe" or "it seems like". (M = 1.81)

When speaking out about gender inequality, people should use proper or grammatically correct language. (M = 1.94)

When speaking out about gender inequality, people should avoid being explicitly nice, polite, or deferential. (M = 2.0)

The results of this section reveal that in general, the respondents were not interested in prescribing linguistic practices or even predicting what is more helpful to a sociopolitical agenda. The mean responses for each prompt were all between 1.5 and 2, meaning they leaned toward “disagree” but were not strong opinions. A score of “disagree” does not necessarily entail that the respondent believes in prescribing the opposite of the feature in question; i.e., disagreeing with “one should avoid vulgarity” does not automatically convey the opinion that “one should use vulgarity”, but may imply disapproval of prescriptive language attitudes in general. This latter interpretation appears more likely considering the opinions expressed in response to “Do you have anything else to add?":

(F, 45-54) I’m not here to police people’s language. All of these modes of expression can be useful, and they’re up to the person typing, not me.

(F, 25-34) I don’t really think there’s a style of speech that people "should" use when speaking out about gender inequality. People express themselves in different ways. Polite, impolite, it doesn’t matter to me. What matters is the content of the speech, not the style.

(M, 25-34) To have these ideas about how people should speak and broach a topic is very much tone policing to me and this makes me uncomfortable. One cannot tell someone else how to speak to their experience.

(F, 45-54) While niceness, grammatical correctness, and avoiding vulgarity can be nice, we have to understand that all women are different and have varying experiences that will contribute to the ways in which they discuss gender inequality. It is tone policing and classist to require the discourse always be polite, grammatically correct, and unoffensive.
Examples (43)-(46) constitute metalinguistic commentary that establishes the authors’ stance vis-à-vis the discourse itself. Their stancetaking is distinctly anti-elitist (Jaworski & Thurlow, 2009), implying that only classist or otherwise prejudicial people place value on linguistic style in the context of feminist discourse, or perhaps at all. The concept of “tone policing” emerged in activist circles in response to criticisms of how activist stances are expressed. It refers to the idea that oppressed populations are criticized not for the content of their activism but their “tone”. Many activists argue that tone policing is a tool of the privileged to decenter oppression in political conversation by distracting participants with questions about the style in which a message is delivered, and that it establishes an expectation that oppressed populations must express themselves in a way their oppressors find palatable in order to be heard[^30].

6 Response to Twitter stimuli

To gauge conscious attitudes toward the particular linguistic variables studied in the previous chapters, the survey invited participants to rate two similar tweets that differed only by one of the four features studied quantitatively. The participants were asked to compare the two tweets on multiple axes. An example stimulus is below, showing a pair of tweets with one member containing a hedge and the other omitting it:

Figure 9. Pair of tweets
### Figure 10. Likert scale accompanying pair of tweets

<table>
<thead>
<tr>
<th></th>
<th>Definitely tweet A</th>
<th>Neither A nor B</th>
<th>Definitely tweet B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convincing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggressive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weak</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sarcastic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nice or friendly</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rude</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assertive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polite</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effective</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Results of this section reveal a striking lack of consensus, and possibly of any conscious opinion about any of the variables. Mean ratings were between 2 and 4 on every axis and for each pair and no tweet in a pair was rated significantly different from the other in any dimension. In general, a few patterns emerged, but none carrying any statistical significance. Still, a few inferences may be drawn.

First, hedged variants were consistently rated as weaker and also nicer or friendlier than their hedgeless counterparts, while the non-hedged stimuli variants were rated as more aggressive, assertive, and convincing than the hedged ones. This finding is consistent with previous work showing indexical link between powerlessness and hedges (Fiona Farr, 2013; Jalilifar & Alavi, 2011; O’Barr & Atkins, 1980).

Tweets with overt politeness features were judged, unsurprisingly, as higher on the scales of “polite” and “nice/friendly”, while those without were judged as more aggressive, assertive, and convincing. Tweets with politeness markers that were overtly directed at men (Figures 11 and 12) were judged as more sarcastic, while the stimulus with an apparently female audience was judged as no more or less sarcastic than its comparison stimulus (fig. 13):

![Figure 11. Tweet](image1)

![Figure 12. Tweet](image2)

![Figure 13. Tweet](image3)
Vulgar or profane language in the stimuli predicted higher ratings on scales of assertiveness, convincingness, and aggressiveness. Tweets without vulgarity were judged as nicer, friendlier, and more polite. Somewhat unexpectedly, judgements of rudeness did not negatively correlate with judgements of politeness; tweets without profanity were consistently judged as more polite than their profanity-laden equivalents, but the latter were not judged as ruder. This implies that abstaining from vulgar language may be seen as facework, but that using it is not necessarily a face-threatening act, and that “rude” and “polite” are not, in fact, endpoints of a single continuum in the minds of these speakers.

Judgements of tweets containing marked nonstandard variants showed no pattern whatsoever, with each pair showing different ordering of properties, often with opposing properties with identical or near-identical scores. This is consistent with the total absence of any mention of standardness as a salient feature of gendered language ideology, but is a surprising result considering the consistency of production results in this dimension as reported in the previous two chapters.

7 “Women’s language” in workplace correspondence

Considering the focused scrutiny of women’s language use in professional contexts, the survey also included five stimuli of example phrases commonly cited as self-damaging when used by women in the workplace. This section included stimuli drawn from the kinds of non-academic writings on contemporary ideas of women’s language in professional settings (Crosley, 2015; Leanse, 2015). The purpose of this segment was to determine if people’s impressions of language in feminist discourse was consistent with the type of gendered language variation frequently discussed in popular media (e.g., women hedge or apologize too much, which works against their professional standing or ambitions). The stimuli were presented as excerpts from a
professional email and the respondents were asked to rate each on a 10-point scale along several axes:

"If it's not too much trouble..."

<table>
<thead>
<tr>
<th>Not at all</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>Very 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impolite</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polite</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commanding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weak</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authoritative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subordinate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Masculine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feminine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 14. Language in the workplace prompt and Likert scale
The mean ratings (on a scale of 10) mostly hovered closely around the center position, which a few exceptions. In each case, “polite” received the highest rating (M ranged from 7.41 to 8.33) and “impolite” the lowest (M ranged from 1.7 to 2.37). These were highly significant in each case (p < .001). Each of the stimuli was rated more “feminine” (M between 5.0 and 5.74) than “masculine” (M between 3.43 and 4.30), but this difference, though consistent across the stimuli, was not significant.

8 Discussion

The results discussed here reveal four key findings about social meaning of linguistic variants among the survey participants:

A. Some features of women’s language emerge as salient and indexically linked to gender.
B. Hedges and politeness are the most frequently mentioned features of stereotypically feminine language
C. Profanity is interpreted as a dimension of politeness
D. The participants do not subscribe to standard language ideology and are vocally anti-prescriptivist.

These findings supplement the quantitative results in surprising ways. Patterns emerge between the language ideologies of the studied population and the observed linguistic practices in the Twitter corpus, suggesting a link, however inadvertent or subconscious, between salient attitudes about gendered language practices and style-shifting patterns.

8.1 Women’s language as language ideology

Though Lakoff’s women’s language paradigm is problematic as a scientific framework, many of its central tenets persist even in the minds of active crusaders for gender equality,
thereby making it a practical point of reference for discussing whether speakers conform to the expectations that accompany this paradigm. Like in Hall’s findings among phone sex operators (1995), women’s language emerges here as a meaningful bundle of features in construction of gender in this mediated space.

The participants’ understandings of gendered language stereotypes are consistent with narratives of hesitant, self-deprecating, deferential female speakers in contrast with assertive male speakers. Those who offered generalizations about gendered language stereotypes, whether framing them as such or, more often, presenting them as fact, gravitated toward the same assertions. Women are thought to apologize more, sound more hesitant or questioning, are more focused on the face needs of their interlocutors than the efficiency of information transfer, relative to men. This is not explicitly negatively evaluated; in fact, the complementary impressions of how men speak were evaluated with more derision than those of women.

These results suggest that women’s language constitutes a coherent language ideology as defined by Irvine and Gal (2000) as the process by which people frame their understanding of language differences and map those understandings onto people, places, or events. Furthermore, the gender categories of women and men are themselves ideological constructs, and the answers reported here characterize “ideal” masculinities and femininities (Kiesling, 2007).

These ideal constructions of gendered language are consistent not only with a women’s language paradigm, but also with later theories of gender difference in language use as comparable to cultural differences (Michel, 1994; Tannen, 1990). The respondents report men and women having different communication goals and constraints and cite socialization and cultural learning processes as motivations for the differences they observe. These are also part of the women’s language ideology in addition to the particular linguistic features it includes.

The respondents also frame women’s linguistic practices as a response to patriarchal expectations, which is compatible with how Lakoff framed them, as well proponents of a
dominance model (Spender, 1990). The features of language use that index construction of
gender are treated as conditioned rather than biologically determined, but gender identity is still
treated as a static, predictive variable. The women’s language ideology as it emerges here is
consistent with both dominance and difference approaches to language and gender theories
(Talbot, 2010). However, the survey participants did not express ideological preference for
stereotypical men’s language over women’s; this diverges from Lakoff’s deficit-based view of
women’s language as reinforcing of patriarchal structures, as well as from frequent messages
blaming women’s language use for their professional marginalization. No one in this sample
suggested that the way women speak has a negative impact on their social standing, or that
men’s language is more valuable in their estimation.

8.2 Salience of hedges and facework in language attitudes

This is not to say that the entirety of Lakoff’s canonical women’s language paradigm was
reproduced in these survey results. The respondents focused on two features in particular:
hedging and politeness. Overwhelmingly, the survey results highlight a perceived tendency for
women to temper the epistemic force of their statements, or their own certainty or epistemic
rights, delegitimizing themselves. This generalization includes hedges like those examined in
the corpus studies reported in the previous chapters.

The other recurring focus of gendered stereotypes reported here is the practice of
apologizing, prioritizing face needs of conversation partners, and other overt markers of
politeness. This was expressed sometimes as a generalization about women with men serving as
the implied unmarked case, and sometimes as the reverse, but the theme was consistent through
all of the results.

8.2.1 Profanity as a politeness feature
While profanity was not as frequently mentioned in the discussion of gendered language practices, it did come up in one open-ended prompt (example (32) above). More telling was the response to example Twitter stimuli, in which perceptions of taboo language use emerged as tied to perceived politeness of a tweet more than any other attribute on the provided Likert scales. Tweets with swearing, profanity, or vulgarity were consistently judged as less polite, but not “ruder”, than tweets that omitted or censored those lexical items. Taboo language was thereby revealed to be a dimension of politeness for this population, which is consistent with politeness studies that treat taboo profanity and vulgarity as face-threatening acts (Brown & Levinson, 1987; Jay & Janschewitz, 2008).

8.3 Anti-prescriptivism and the absence of standard language ideology

The final main finding from this survey that I wish to highlight is the lack of grammatical prescriptivism among these participants. This emerges in two ways. First, in the elicitation of gendered language stereotypes, no one at all mentioned use of slang, nonstandard variants, overtly prestigious or academic English varieties, or varieties with working-class connotations. This is especially noteworthy considering the of the tendency of women to use more prestigious linguistic variants is among the more frequently documented principles in gendered language variation research (Cheshire, 2004; Schilling-Estes, 2002). The dimension of standard language as a gendered feature is perhaps more empirically demonstrated than any of the linguistic features cited by the participants in their discussion of gendered linguistic practice. The dimensions they do mention consistently (hedges, apologies, politeness, questions, etc.) are by contrast apparently very salient but far less empirically demonstrated.

Not only was standard language ideology absent from elicited attitudes about gendered tendencies in language practice, but when pressed about whether marked nonstandard variants were helpful in feminist discourse, the participants further declined to express any prescriptivist attitudes at all. They had no clear opinion on that particular question, and in the open-ended
follow up questions revealed that they do not condone linguistic prescriptivism of any kind (framed by multiple respondents as “tone policing”). In doing so, they establish stances opposed to standard language ideology (the idea of “correctness”) and prescriptive attitudes toward appropriateness or acceptability in the context of online feminist discourse.

8.4 Relationship with production findings

These findings become more revelatory in conjunction with the intra-speaker variation results presented in chapters 3 and 4. Recall that when tweeting in the #YesAllWomen corpus the female population shifted into a style marked by decreased frequency of hedges, overt markers of politeness, and marked nonstandard variants, and increased frequency of profanity. Under a women’s language paradigm, then, their practices became less “feminine” along the dimensions of politeness, hedging, and vulgarity, and more “feminine” in the dimension of overt linguistic prestige. The nature of style-shifting therefore mirrors the ideological consensus revealed in these survey results. The salient features of gendered language in the minds of the Twitter users are also those that women deploy in more masculine ways according to their own ideas about gendered language practice. Profanity, politeness, and hedges all emerged as salient components of a gendered language ideology in the survey data, and also showed a shift away from ideal (in Kiesling’s sense) women’s language in the #YesAllWomen’s corpus.

On the other hand, marked nonstandard features of American English were not mentioned as a feature of men’s language in the survey results. That category of features also showed significant differences in frequency in the #YesAllWomen corpus relative to the control data, but in a way that is closer to canonical women’s language in Lakoff’s sense. The speakers therefore became more “feminine” in a women’s language framework, but not in their own understanding of language and gender, as this group does not show conscious indexical links between nonstandard variants and stereotypical gendered language. However, the features of a linguistic style considered masculine by this population are the same ones that the participants
gravitate toward when performing feminist activism. I reiterate that these features do not intrinsically or exclusively index gender, as most linguistic variants do not (Ochs, 1992). However, the variants that index feminism through their patterns of production also index masculinity in the authors’ ideological framework.

This is not to say that the Twitter users studied here voluntarily shift their linguistic style in ways that they consider more masculine when they discuss feminism in public. Linguistic variants index social acts or activities more directly than they index gender, but those acts may also be associated with one gender or another. In this case, the linguistic resources used in the #YesAllWomen discourse are indexical markers of empowerment, authority, and legitimacy, and the speakers’ understanding of ideal men’s language also connects to these features. The women are not tweeting “more like men” when they perform feminist stances, but they are tweeting in way that signals stances and styles that also appear indexically linked to an ideal masculinity. Meanwhile, though use of nonstandard English variants may be more demonstrably linked to gender in observed language use, it is not saliently gendered to this population. When this population is observed to use fewer nonstandard variants in the #YesAllWomen discourse, it may be interpreted as increasingly feminine according to a women’s language theory or other sociolinguistic principles but would never be interpreted as such by the speakers. Participation in the #YesAllWomen discourse, perhaps, liberates the speakers from the patriarchal forces they cite as influencing women’s stereotypical language practices, licensing more stereotypically masculine language use according to their own language ideology. This does not include an increase in nonstandard variants, which is not interpreted as a feature of masculine language.

9 Implications and limitations

These results suggest this population possesses an unmistakable awareness of stereotypical men’s and women’s language and distinct ideologies that reproduce the central
tenets of those stereotypes. The framing of these ideologies, however, suggests an understanding of linguistic practice as socially mediated rather than innate or immutable. The participants in the Twitter discourse in question apparently conceptualize language as a social practice that exists in dialog with other social structures and expectations. The way people speak, as this population is acutely aware, is inseparable from broader systems of indexicality and ideas about what linguistic forms are valued or stigmatized by social norms. This population’s linguistic practice is itself inseparable from its language and gender ideology. Their style-shift mirrors their ideology in a way that implies that ideas about power and authority remain tightly bound with ideas of gender, and that performing a feminist stance online is not isolated from performance and perceptions of gender identity. This does not imply a causal relationship between the participants’ perceptions of gendered speech and their production of feminist language, but it does provide evidence that neither can be isolated and studied in a vacuum. These speakers perform a gender identity making use of linguistic resources that have indexical links to gender by way of ideological processes. They simultaneously construct a gendered self and a feminist self with linguistic resources that are given meaning through their repeated use in certain stances.

However, these implications do not emerge without qualification. The small sample size and the difficulties inherent in discussing abstract concepts in an impersonal and constrained medium are the primary limitations of these findings. While the surveyed population expressed clear patterns in how they think of language, gender, and feminism in relation to one another, it is presumptuous to extrapolate over the thousands of people who authored this study’s experimental dataset from only a few dozen self-selected Twitter users.

Furthermore, the lines between perceived linguistic stereotypes and perceived linguistic variation were not always sharply drawn in the responses, and interrogating the ideological processes involved in constructing those boundaries fell outside the scope and capabilities of an online survey that was already a significant time commitment for the participant. Similarly, the
survey did not include very much personal introspection, leaving only speculation as to what the authors of the data set believe their own linguistic practices to be in the genre of feminist tweeting. There was no opportunity for the participants to express self-perception regarding linguistic variation online. Due to their reluctance to declare what linguistic variants are more or less helpful in feminist discourse, it is a reasonable assumption that they would not identify their own style-shifting tendencies, but one cannot reasonably speculate on anyone’s perception of their own language use.

Despite these concerns, the results provide a rich context for the quantitative results reported in the previous two chapters. The linguistic style-shifting apparently triggered by participation in the #YesAllWomen discourse does not appear to be the result of conscious opinions of how anyone “should” use language on Twitter, but is undoubtedly informed by the perceived indexical relationship between linguistic practice and gender identity.
Chapter 6. Conclusions, limitations, and future directions

Three primary research questions have guided the trajectory of this project:

1. What linguistic features characterize participation in an overtly feminist digitally-mediated Discourse?

2. Are features of a classic conceptualization of “women’s language” salient for people when discussing contemporary gender and power politics?

3. In what ways are linguistic practices in online feminist Discourses informed by or in dialogue with language and gender ideologies?

To address these questions, this dissertation combines lines of inquiry from variationist approaches to sociolinguistics with intellectual traditions from linguistic anthropology, gender and feminist studies, and digital media studies. It explores the performance of gender and political stances in digitally mediated text by examining variation of a set of linguistic features on Twitter and the social meaning that speakers connect with those features. The relationships between linguistic forms and social meanings are not fixed (Eckert, 2008) and this study assumes dynamic indexical relationships between the variables and the stances, people, styles, and ideologies that surround them. Language in this study is treated as one locus of identity production rather than an effect of static social identities (Bucholtz & Hall, 2003). In non-co-present text-based interaction (e.g., Twitter), linguistic form is an even more crucial resource for processes of identity production and negotiation because many other potential resources (physical appearance, voice, prosody, gesture, etc.) are less readily available, if available at all.

This study examined written linguistic variants that have been linked to gender identity in linguistics literature and which serve as potential resources for performance of gender and feminism on social media. Gender is not assumed as a natural binary system that is projected

---

31 “Big D Discourse” here is taken to refer to linguistic communication, the participants in the communication, and the systems of meaning they bring to it, after Gee (2015). This contrasts with a “small d discourse”, which is constituted by a series of linguistic utterances in a given situation.
onto the speakers in the study, but rather a locally meaningful construct among the population that produced the dataset.

The features in question (hedges, politeness markers, nonstandard variants, and taboo language or profanity), are analyzed not as directly indexing gender (Ochs, 1992), but as resources employed repeatedly to signal similar stances, which give rise over time to coherent linguistic styles (Bucholtz, 2009; Johnstone, 2009; Kiesling, 2009). The social meanings of these styles are mediated by language ideologies (Garrett, 2010; Irvine & Gal, 2000; Schieffelin, Woolard, & Kroskrity, 1998) that reinforce indexical relationships between language use and social categories and entities. The possible ideas and people that these features index vary according to the medium, the speaker, the audience, and other components of a Discourse. This dissertation illuminates some of the ways in which these relationships between form and meaning are produced.

1 Summary of results

This study combines quantitative and qualitative methods to investigate the relationships between linguistic practice, gender identity, and feminism on Twitter. Four Twitter corpora served as the dataset for the first two results chapters. One corpus contained tweets authored by 140 self-identified women and tagged with the hashtag #YesAllWomen, a popular feminist hashtag from 2014. A control corpus authored by the same group of women but not screened for hashtags or topics served as a point of comparison. Two other corpora authored by a group of 68 men allowed for comparison in style-shifting practices across men and women performing feminist stances on Twitter. Additionally, a survey focusing on biographical information and attitudes toward language, gender, and feminist activism provided context in which to situate the variation study results. The survey data help complete a picture of the construction of gender and feminist identity on Twitter and the process of building indexical relationships between linguistic form and ideological constructs.
The results of two intra-speaker analyses indicate that men and women appear to take distinctly different stances, as signaled by their linguistic variation, and that these stancetaking strategies give rise to a style-shift when participating in a Discourse on Twitter that is made explicitly and publicly feminist through use of the viral hashtag #YesAllWomen. The intra-speaker variation studies suggest that women systematically use linguistic resources differently when performing a feminist self. The style shift is characterized by an increased frequency in vulgar language, a decreased frequency of hedges and politeness markers, and a decreased frequency of nonstandard contractions or abbreviation. Under Lakoff’s (1973; 1975) theory of women’s language, the female Twitter users use language in a less canonically feminine way in the dimensions of hedging, politeness, and profanity when they participate in the #YesAllWomen Discourse, and in a more canonically feminine way in the dimension of prescriptive standardness. The robustness of these patterns suggest that these features are associated with a feminist stance as produced by women.

When men perform feminist allyship—that is, their feminist stance and their identity as men simultaneously—they mirrored the shift but not in a statistically significant way, a result which is likely due in part to the smaller population size of men constructively participating in the Discourse. Closer examination of the linguistic practices of male feminist allies in the #YesAllWomen Discourse suggest that linguistic features index different things when different people deploy them, and men’s variation is evaluated differently from women’s in connection with their differing performances of gender identity. The linguistic practices in the #YesAllWomen Discourse suggest a set of linguistic features that are coded as anti-patriarchal or subversive, but deeply intertwined with the identity politics of the participants.

The findings underline the complicated political and social constraints on the Discourse. Gender identity and equality is fraught terrain. People participating in the Discourse are doing so because they are emotionally charged, hurt, or angry, and those expressing sympathy must navigate such sensitive issues. The previous chapters suggest that the
expectations of male allies in particular are exceptionally difficult to navigate. Men are expected to be silent and vocal supporters simultaneously, to stop talking and let women speak, and to speak up when they see gender inequality because they are granted more respect and authority in social interaction. Male allies appear to negotiate this by addressing one another, or no one in particular, while women freely address both other women and men.

Though audience design (Bell 1984) influences communicative practices between men and women, yielding different stances, the subset of participants who provided survey responses appeared to converge on a coherent language ideology concerning the features that have historically been tied to constructions of women’s language. The survey data reveal that politeness, tentativeness, and profanity (particularly as a dimension of politeness) are all highly salient features of gendered language in the minds of the population in question. These conceptualizations of gendered language are consistent with Lakoff’s canonical view of women’s language and with narratives that are manifested in contemporary metalinguistic essays and opinion pieces that argue that women hedge, apologize, or self-efface too often, to their own detriment.

The survey respondents, however, displayed one significant departure from a canonical women’s language paradigm, both in production and perception; this is the finding related to the social meaning of standardness or prestige variants. Lakoff and several linguists before and since have asserted that women overall use linguistic varieties that carry mainstream symbolic capital or index upward economic mobility (Cheshire, 2004; Labov, 1990; Schilling-Estes, 2002; Wolfram, 1969). In the #YesAllWomen corpus, both men and women appeared to avoid nonstandard or markedly non-prestigious variants in their feminist-hashtagged tweets. According to a women’s language paradigm, such a shift translates to an increase in canonical femininity, while the other observed features of the apparent style-shift are more consistent with stereotypically masculine styling. The survey results, meanwhile, reveal that standardness,
correctness, or other ideological constructs rooted in economic or institutional power structures are not saliently gendered for this population at all.

Not only is standardness not ideologically tied to gender the way the other studied features are, a distinct anti-prescriptivist language ideology is evident in the survey data. The #YesAllWomen Discourse is unequivocally subversive with explicit anti-establishment overtones, providing a natural context for an anti-hegemonic language ideology to emerge, so a subversive language ideology is fitting. However, the observed practices are surprising in this context. While, when asked, the survey participants say that mainstream American English standardness is irrelevant to the value of contributions to #YesAllWomen or similar forums, the participants’ language appears to become more formal or marked for prestige when they participate in the #YesAllWomen Discourse. A decrease in nonstandard language may establish a stance that serves to legitimize or enhance the gravity of an utterance, stance by endowing it with an authority granted to those with access to prestige varieties of English and thereby “authorizing” the speaker (Bucholtz & Hall, 2004). Furthermore, the shift in nonstandard variant use may be linked to an increase in the size of the imagined audience, which a possible result of hashtag use (Pavalanathan & Eisenstein, 2015). A hashtagged tweet is more public than would be otherwise, and the effect on linguistic style may approximate one triggered by public speaking or another genre with a large or unfamiliar audience.

It is worth pointing out here that a possible limitation of this finding is the education level of the population. In the survey results, those who reported all had some form of post-secondary education, with none reporting an unfinished high school education, only a high school diploma or GED, or trade school degree. It is possible that a less educated population would be more consciously prescriptivist or hyper aware of their practices with respect to mainstream prestige and the expectations associated with privilege.

2 Challenges of online sociolinguistics
The possible interaction between education level or socioeconomic status and linguistic practice and ideology leads to another research goal of this dissertation: to identify and address some of the challenges inherent to conducting sociolinguistic studies of online media. While social media indisputably constitute a gold mine of searchable, archived linguistic data from more diverse populations than many researchers previously had ready access to, the typical poverty of online metadata is highlighted by this study’s methods and results. The corpus results were statistically significant but lacked much information about the authors of the corpus and benefitted immensely from a supplemental study employing different types of data and analytical methods.

The survey results contextualize the quantitative corpus study results in a way that potentially affects their interpretation. Before an investigation into the language attitudes of the population, it was unclear what meaning, if any, was attached to the features tracked in the study. The quantitative results indicated a shift away from conventionally feminine language according to some of the dominant trends in language and gender research, with the exception of practices regarding overt linguistic prestige or standardness. The language attitude survey revealed that the latter variant is not as linked to expectations of gender performance in the same way the other three feature categories are, at least according to the survey respondents. Certain components of the construction of women’s language—those that the speakers associate with women—are those that are used less when speakers perform a feminist persona online, while the aspect of women’s language that is not salient to them is enhanced under the same circumstances.

Furthermore, the self-reported demographic information illuminates the limitations of this type of study. Given how overwhelmingly white, educated, and monolingual the survey respondents appear to be, generalizing from the observed variation patterns is a problematic venture. Online communication research frequently abstracts over questions of demographics and intent of the people who produced the data.
3 Implications

3.1 CMC research and the value of mixed-method studies

None of this is to say that study of Internet communication is not worthwhile or only valid when ethnographic or survey methods are available. These conclusions only serve to demonstrate the value in combining quantitative and qualitative methods in contexts where each only contributes part of the picture. Variation studies can benefit from an understanding of the speakers’ constructions of the social categories that are the focus of study. Here, the Twitter user’s conceptualizations of language, gender, and power help complete an image of possible motivation and consequence of style-shifting. Certain features are perceived by participants in the #YesAllWomen Discourse as gendered or markers of empowerment, liberation, or control. These perceptions help reveal how meaning is produced through linguistic practice in a given genre, and only emerged in the follow-up survey research following the corpus study. This dissertation serves not only to show how social and political stances are linguistically performed and how linguistic variation is tied to ideological frameworks, but to demonstrate the value of supplementing quantitative linguistic research with more anthropological approaches.

Androultosopoulos (2006) has advocated for the value of “online ethnography” as a component of both quantitative and qualitative CMC research. The methods here are not ethnographic in the sense that they did not include prolonged offline interaction with the studied population, which is not always an available option to the researcher. The anonymity and physical displacement of online communication can simultaneously entice and stymie social scientists; construction of identities, social networks, interaction, and communities that transcend physical proximity invite investigation into how people build and maintain systems of social meaning across space and time. At the same time, traditional ethnographic methods face new practical challenges. This study provides support for methods that go beyond variationist approaches but have a lower barrier to entry than extended hybrid online/offline ethnographic forays.
3.2 Gender and feminist research

The findings presented here build on previous work treating gender as mutable, socially mediated, and performed. I have contributed evidence that digital media are worthy research sites where gender and feminism are negotiated in new and complicated ways. A growing body of research from the past few years has emphasized the role of digital media in studies of contemporary feminism (Baer, 2016; Daniels, 2016; Jarrett, 2015; Loza, 2014), a necessary step for feminist theory. Gender and feminist studies may benefit from a coherent theoretical construct—an update on “cyberfeminism”—that captures the centrality of digital communication in feminist practice. The Internet has become a primary marketplace for feminist Discourses. The way people theorize and engage with feminism is now frequently worked out on social media platforms and feminist theories should account for the shift in modality that social activism and feminist identities have recently undergone.

3.3 Sociolinguistic research

This research contributes to the body of work treating social categories not as static properties that predict or linguistic variation, but as ideologically mediated constructions that negotiated through interaction. It also serves as an example of the ways in which social categories intersect and inform one another. Traditional macro-categories like gender or race are not only overly simplistic and essentializing, but they are not inherently more meaningful than categories like “feminist” or “ally” or “activist” among members of a given population sample. Furthermore, identity-building processes are not isolated within a person. For example, this study has shown that performing a female feminist identity is not the sum of a performance of a female identity and a performance of a feminist identity. Likewise, doing male feminism online is not equal to doing maleness online and doing feminism online. The intersecting social
categories that are meaningful to the population cannot be segmented and studied in isolation without acknowledging that process.

4 Future directions

4.1 Feminist Twitter since #YesAllWomen

This study invites several possible avenues for future research that may address some of the questions that emerged from the findings reported in this dissertation. First, the Twitter data that are the focus of this research date as far back as 2014. Since then, Twitter has seen the rise of the hashtag #MeToo, a hashtag with a similar message to #YesAllWomen’s, but which has proven to be farther-reaching and more enduring. The #MeToo movement provides a resource far richer in some ways than the #YesAllWomen movement, inviting additional research into the role of linguistic variation in performance of a socio-political stance on gender inequality. #MeToo began as a hashtag used by those who had been sexually abused or assaulted, particularly at work, and became popular following revelations about the Hollywood producer Harvey Weinsten’s record of sexual harassment and abuse in the American film industry. The inclusion of a first-person pronoun introduces an interesting difference from #YesAllAomen, which refers to women as a group but carries no entailment of the speaker identifying as a member. #MeToo, on the other hand, serves as a cry from people of any gender identity about their experiences with sexual victimization, and its proliferation in social media illustrated the ubiquity of such experiences. The subsequent online cacophony of sexual misconduct allegations led to reckonings across varied industries, the most public occurring in the film industry, but with prominent figures from several other professions promising to investigate, punish, and prevent sexual harassment among their ranks. Eventually, #MeToo became an idiomatic transitive verb, frequently used in passive voice to foreground the phenomenon of successful men being publicly shamed or exiled from their professional circles following accusations of sexual misconduct (e.g., “Did you hear about Ryan Adams? He got Me
Too’ed”\textsuperscript{32}. This usage de-centers the experiences of the victims and refocuses the conversation on the effects felt by perpetrators. Overall, #YesAllWomen and #MeToo are very different Discourses with different implications for the participants, taking place in different historical contexts—a lot has happened on and off Twitter in the past five years. Language of #MeToo merits its own investigation to determine how these differences are related to linguistic practice.

4.2 Understanding of gender

A binary gender framework was practical for this dissertation because it was a salient feature of the dataset and the context in which it was produced and was thus used to organize and analyze the population and their linguistic data. Tweets tagged with the #YesAllWomen hashtag often assumed two mutually-exclusive genders. The driving research questions of this document were built on the premise of a male/female dichotomy, but this trend has been examined and criticized as reductive according to some scholars, particularly in feminist linguistics (Lorber, 2000; Rodino, 1997). As treatment of gender in linguistic research is interrogated, problematized, and reworked, conversations about gender need to be re-evaluated to reflect evolving understanding of it. The representation of nonbinary people in the survey results suggest that some of the most engaged participants in gender equality Discourses do not relate to gender as a concept in the way the majority of language researchers do. If we consider speaker ideology in analysis of variation research, as this dissertation suggests is valuable, it is crucial that scholars be reflexive when conducting their research. Projecting a binary gender framework and analyzing language use with respect to it when the speakers themselves do not conceptualize gender in such a way undermines the value of the research. Linguistic research that relies on social categories of its research subjects should reflect or at least consider current

\textsuperscript{32} See Gaillot (2018) for a discussion of this linguistic innovation.
understandings of those social categories, particularly those expressed by the research population.

4.3 Additional questions and final thoughts

This dissertation has demonstrated that linguistic features are indexically linked to performance of a feminist identity, and that this indexical relationship interacts with production of gender identity and language ideology. Women participating in a feminist Discourse on Twitter recruit linguistic variants that index an idealized masculinity according to their own expressed ideas about language and gender. These variants, combined with a more standard register, contribute to meaningful stances that index authority and confidence. Several questions follow from this dissertation’s findings. For example, what other linguistic features emerge as markers of an anti-patriarchal linguistic style? Is the linguistic variation seen here similarly evident on other online social media platforms, or in spoken genres? Are speakers aware of their style-shifting? Do other hashtags trigger similar variation? All of these questions fall beyond the scope of this dissertation but offer important insights into how social and political positions are expressed and interpreted. They invite further research using hybrid methodologies to examine the nature of linguistic variation and related systems of meaning, particularly with respect to political and social activism online. As methods of political activism evolve with digital media technologies, linguists and other social scientists have the opportunity use the diverse tools at their disposal to uncover relationships between mediated language use, politics, and identity production. These issues become increasingly consequential as the Internet becomes accessible to more people worldwide and political and social movements continue to be generated and their messages circulated online and beyond.
### Appendices

#### Appendix A: Search terms by feature category

<table>
<thead>
<tr>
<th>Category</th>
<th>Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Politeness markers</strong></td>
<td>thank you, thanks, please, sorry</td>
</tr>
<tr>
<td><strong>Profanity/taboo language</strong></td>
<td>fuck, shit, damn, bitch, crap, piss, cock, fag, bastard, dick, slut,</td>
</tr>
<tr>
<td></td>
<td>douche, cunt, tits, pussy, asshole</td>
</tr>
<tr>
<td><strong>Hedges</strong></td>
<td>i guess, sort of, sorta, kind of, kinda, maybe, i think, seems like</td>
</tr>
<tr>
<td><strong>Nonstandard abbreviations</strong></td>
<td>gonna, gunna, wanna, imma, ima, aint, ain't, wtf</td>
</tr>
<tr>
<td><strong>Null apostrophe</strong></td>
<td>cant, wont, isnt, dont, wouldn't, couldn't, shouldn't, didn't, wasnt,</td>
</tr>
<tr>
<td></td>
<td>arent, hasnt, havent, hadnt, doesn't</td>
</tr>
</tbody>
</table>
Appendix B: Blank survey

Informed consent

THE CITY UNIVERSITY OF NEW YORK
Lehman College

INTERNET BASED INFORMED CONSENT

Title of Research Study: Investigating linguistic variation and language ideology in feminist discourse on Twitter

Principal Investigator: Professor Cecelia Cutler, PhD

My name is Nora Morikawa and I am a doctoral candidate in the Linguistics department at the City University of New York. Under the supervision of Dr. Cecelia Cutler, I am conducting my dissertation research on feminist discourse on Twitter.

You are being asked to participate in this research study because you have tweeted with the hashtag #yesallwomen. The purpose of this research study is to get a better idea of who contributes to feminist discourse on Twitter and how they use and perceive language. If you agree to participate, we will ask you to respond to a written survey containing questions about yourself and your opinions on language and gender equality. The survey will likely take you a maximum of 20 to 30 minutes. Aside from possible discomfort when discussing controversial issues, there are no known risks associated with this study. Your responses will be entirely anonymous and not linked to your identity or Twitter handle in any way. All responses will be stored on a password-protected drive and will only be accessed by the researchers. Your participation in this research is
voluntary. If you have any questions, you can contact Nora Morikawa at egoldman@gradcenter.cuny.edu. If you have any questions about your rights as a research participant or if you would like to talk to someone other than the researchers, you can contact CUNY Research Compliance Administrator at 646-664-8918 or HRPP@cuny.edu.

Please click the arrow to go to the next page and start the survey.

***NOTE: This survey works best on a computer. You can complete it on a mobile device, but some text might be difficult to read on a small screen.***

**General language/gender**

You have been contacted for this survey because some of your tweets or retweets were included in a publicly available collection of tweets from 2014 containing the hashtag #yesallwomen. This survey asks for your opinions on language use, gender issues, and some information about yourself. If you wish, you may leave any of these questions blank, but more information really helps our research project.

How much do you agree with the following statement?

I consider myself a feminist.
How do you define feminism?

What does the #YesAllWomen hashtag mean to you? Why did you tweet with it?

How much do you agree with the following statement?

**There are stereotypically gendered styles of using language.**

How much do you agree with the following statement?
There is a stereotypical way that men use language.

| Strongly disagree | 0 | 1 | 2 | 3 | Strongly agree | 4 | 5 |

If you think there is a stereotypical way men speak, how would you describe it? Give examples if you can think of any.

How much do you agree with the following statement?

There is a stereotypical way that women use language.

| Strongly disagree | 0 | 1 | 2 | 3 | Strongly agree | 4 | 5 |

If you think there is a stereotypical way women speak, how would you describe it? Give examples if you can think of any.
How much do you agree with the following statement?

It is helpful to the feminist cause when **women** use feminist hashtags like #YesAllWomen or #MeToo to draw attention to the issue and share their experiences.

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1  2  3  4  5</td>
</tr>
</tbody>
</table>

How much do you agree with the following statement?

It is helpful to the feminist cause when **men** use feminist hashtags like #YesAllWomen or #MeToo to draw attention to the issue and share their experiences.

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1  2  3  4  5</td>
</tr>
</tbody>
</table>

How much do you agree with the following statement?

When speaking out about gender inequality, people should **avoid** taboo or inappropriate language (swearing, vulgarity, etc.)

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1  2  3  4  5</td>
</tr>
</tbody>
</table>
How much do you agree with the following statement?

When speaking out about gender inequality, people should avoid words or phrases that suggest uncertainty, like “maybe” or “it seems like”.

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th></th>
<th></th>
<th></th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Circle]</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

How much do you agree with the following statement?

When speaking out about gender inequality, people should use proper or grammatically correct language.

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th></th>
<th></th>
<th></th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Circle]</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

How much do you agree with the following statement?

When speaking out about gender inequality, people should avoid being explicitly nice, polite, or deferential.

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th></th>
<th></th>
<th></th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Circle]</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
Do you have any additional comments about anything on this page?

Tweet pair 1

In this section, you'll see 12 pairs of tweets. Below each tweet, you will see a series of attributes next to a scale. For each attribute, indicate which tweet in the pair is better described by it, with a value of 0 meaning "Definitely tweet A", and value of 6 meaning "Definitely tweet B".

A.

Maybe it's just me but I have a problem with songs talking about "blurred lines" & lyrics like "you know you want it." #YesAllWomen

B.
I have a problem with songs talking about "blurred lines" & lyrics like "you know you want it." #YesAllWomen

Which tweet of the above pair is more...

<table>
<thead>
<tr>
<th>Definitely tweet A</th>
<th>Neither A nor B</th>
<th>Definitely tweet B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tweet pair 2

A.
Because #YesAllWomen is even necessary to talk about the stuff that ALL women deal with just for being women

B.

Because #YesAllWomen is even necessary to talk about the bullshit that ALL women deal with just for being women

Which tweet of the above pair is more...

<table>
<thead>
<tr>
<th>Definitely tweet A</th>
<th>Neither A nor B</th>
<th>Definitely tweet B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Convincing
Aggressive
Weak
Sarcastic
Nice or friendly
Rude
Assertive
### Tweet pair 3

**A.**

A lot of men spend all their time trolling the #YesAllWomen HT. It's Sickening.

**B.**

A lot of men seem to spend all their time trolling the #YesAllWomen HT. Kind of sickening.

Which tweet of the above pair is more...

<table>
<thead>
<tr>
<th>Definitely tweet A</th>
<th>Neither A nor B</th>
<th>Definitely tweet B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

- Convincing
- Aggressive
Tweet pair 4

A.

If you respond to #yesallwomen with "but not all men!", you don't really get it.

B.

If you respond to #yesallwomen with "but not all men!", I'm sorry, you don't really get it.

Which tweet of the above pair is more...
<table>
<thead>
<tr>
<th>Definitely tweet A</th>
<th>Neither A nor B</th>
<th>Definitely tweet B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Convincing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggressive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weak</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sarcastic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nice or friendly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rude</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assertive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polite</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effective</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tweet pair 5

A.

#YesAllWomen Please listen to your children, your friends, your sisters, your mothers when they tell you what happened. 1 in 3 women

B.

Which tweet of the above pair is more...

<table>
<thead>
<tr>
<th>Definitely tweet A</th>
<th>Neither A nor B</th>
<th>Definitely tweet B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

- Convincing
- Aggressive
- Weak
- Sarcastic
- Nice or friendly
- Rude
- Assertive
- Polite
- Effective

Tweet pair 6

A.
Because talking about how a man shot up women because they wouldn’t f**k him is met with, literally, cries of “what about the men?” #YesAllWomen

B.

Because talking about how a man shot up women because they wouldn’t f**k him is met with, literally, cries of “what about the men?” #YesAllWomen

Which tweet of the above pair is more...

<table>
<thead>
<tr>
<th>Definitely tweet A</th>
<th>Neither A nor B</th>
<th>Definitely tweet B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

- Convincing
- Aggressive
- Weak
- Sarcastic
- Nice or friendly
- Rude
Tweet pair 7

A.
Even after being told #YesAllWomen is a space to discuss women’s issues, men wanna cry about not being included like petulant toddlers.

B.
Even after being told #YesAllWomen is a space to discuss women’s issues, men want to cry about not being included like petulant toddlers.

Which tweet of the above pair is more...

Convincing
Tweet pair 8

A. 

I didn't use #NotAllMen because if u need that reassurance then u aren't my ally & I ain't here to teach ya #YesAllWomen

B. 

I didn't use #NotAllMen because if you need that reassurance then you are not my ally & I am not here to teach you #YesAllWomen
Which tweet of the above pair is more...

<table>
<thead>
<tr>
<th>Definitely tweet A</th>
<th>Neither A nor B</th>
<th>Definitely tweet B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

Convincing
Aggressive
Weak
Sarcastic
Nice or friendly
Rude
Assertive
Polite
Effective

Tweet pair 9

A.

Men, read the #YesAllWomen tweets, believe our stories, listen to our experiences. We don't need you to "play devil's advocate" right now.
B.

Dear Men, read the #YesAllWomen tweets, believe our stories, listen to our experiences. We don’t need you to “play devil’s advocate” right now. Thank you.

Which tweet of the above pair is more...

<table>
<thead>
<tr>
<th>Definitely tweet A</th>
<th>Neither A nor B</th>
<th>Definitely tweet B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

- Convincing
- Aggressive
- Weak
- Sarcastic
- Nice or friendly
- Rude
- Assertive
- Polite
- Effective
A. 

#YesAllWomen because it's clear that if marriage/children is NOT my goal then I am not a "Real Woman."

B. 

#YesAllWomen because sometimes it seems like if marriage/children is NOT my goal then I am not a "Real Woman."

Which tweet of the above pair is more...

<table>
<thead>
<tr>
<th>Definitely tweet A</th>
<th>Neither A nor B</th>
<th>Definitely tweet B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Convincing
- Aggressive
- Weak
- Sarcastic
- Nice or friendly
- Rude
Tweet pair 11

A. 

because I'm gonna make a name for myself, not gonna become just someone's "mrs." #YesAllWomen

B. 

Because I'm going to make a name for myself, not just going to become someone's "Mrs." #YesAllWomen

Which tweet of the above pair is more...
<table>
<thead>
<tr>
<th></th>
<th>Convincing</th>
<th>Definitely tweet A</th>
<th>Neither A nor B</th>
<th>Definitely tweet B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

- Aggressive
- Weak
- Sarcastic
- Nice or friendly
- Rude
- Assertive
- Polite
- Effective

Tweet pair 12

A.

Because I'm called a radical feminist \_\_\_\_\_\_ for wanting to feel safe outside of my own home. #YesAllWomen

B.
Which tweet of the above pair is more...

<table>
<thead>
<tr>
<th>Definitely tweet A</th>
<th>Neither A nor B</th>
<th>Definitely tweet B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Convincing
Aggressive
Weak
Sarcastic
Nice or friendly
Rude
Assertive
Polite
Effective

feminism means

"Feminism" means to me (check all that apply)...

- Equal rights and equal pay for men and women
Women's workplace language

For each of the following scales, indicate how would rate each of the phrases if they were in a work-related email from a colleague.

"Just following up..."

<table>
<thead>
<tr>
<th>Not at all</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Very</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impolite</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polite</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commanding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weak</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authoritative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Subordinate

<table>
<thead>
<tr>
<th>Not at all</th>
<th>Very</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2 3 4 5 6</td>
<td>7 8 9 10</td>
</tr>
</tbody>
</table>

Masculine | Feminine

## "Could you please..."

<table>
<thead>
<tr>
<th>Not at all</th>
<th>Very</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2 3 4 5 6</td>
<td>7 8 9 10</td>
</tr>
</tbody>
</table>

Impolite | Polite | Commanding | Weak | Authoritative | Subordinate | Masculine | Feminine

## "I'm so sorry for the delay"

<table>
<thead>
<tr>
<th>Not at all</th>
<th>Very</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2 3 4 5 6</td>
<td>7 8 9 10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Impolite</th>
<th>Not at all</th>
<th>Very</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 1 2 3 4 5 6 7 8 9 10</td>
<td></td>
</tr>
<tr>
<td>Polite</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commanding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weak</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authoritative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subordinate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Masculine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feminine</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

"Thanks so much for your help."
"If it's not too much trouble..."

Demo questionnaire
The questions in this section are completely optional, but the more information you’re willing to share, the more helpful it would be to our study.

What is your gender identity?

What is your age group?

Please select

How do you identify racially or ethnically?

What is your nationality (e.g. Canadian)?

What is your current City/State/Country of residence?
What language(s) did you grow up speaking?

What other language(s) do you speak besides those listed above? Leave blank if none.

What language(s) do you use when texting or messaging friends, family, or colleagues?

What is your highest completed level of education?

Please select

How would you describe your political beliefs/party/position?
Would you be willing to be contacted with follow-up questions? If so, please enter your email address below.

Thank you very much for your participation! If you have any questions or comments about the study, please e-mail Nora Morikawa at egoldman@gradcenter.cuny.edu.
References


Chenoweth, E. (2017, February 7). This is what we learned by counting the women’s marches. *The Washington Post.*


Fiona Farr, A. O. (2013). Would as a hedging device in an Irish context: an intra-varietal comparison of institutionalised spoken interaction. *Journal of Chemical Information and
Modeling, 53(9), 1689–1699.
Gerbaudo, P., & Treré, E. (2017). In search of the ‘we’ of social media activism: introduction to the special issue on social media and protest identities.
Grieser, J. A. (2019). When the White Knight is a Woman: An Intertextual Analysis of Antagonism in a Female-Oriented Fandom. Language @ Internet, 15, 1–20.


Lane, L. (2015). Feminist Rhetoric in the Digital Sphere: Digital Interventions &the


social movements. *Feminist Media Studies, 14*(6), 1094–1095.


Poland, B. (2016). *Haters: harassment, abuse, and violence online*.


