NORTHWESTERN UNIVERSITY

Transgressing the binary: Gendered language practices on Twitter

A THESIS

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

MASTER OF ARTS

Field of Linguistics

By

C. Michael Senko

EVANSTON, ILLINOIS

June 2022

ABSTRACT

Cisnormativity is in part perpetuated through harmful language practices such as misgendering and deadnaming. I extend previous work on this topic by collecting tweets discussing four trans celebrities before and after their coming-out events (COEs) to quantify the distribution and content of these practices. Data consist of 7m tweets mentioning any of four 'target' celebrities with publicly documented COEs (two nonbinary, one trans woman, one nonbinary transmasculine), as well as three 'comparison' celebrities without such events (one trans woman, one cis woman, one cis man). This study constitutes the first computational social media analysis of nonbinary *they*, pronominal misgendering, and deadnaming.

Distributional analysis reveals that the target celebrities who use binary pronouns (trans-binary) have their pronouns affirmed at a slightly lower rate than cisgender celebrities; however, we observe that the celebrities who use nonbinary *they* (trans-nonbinary) are affirmed at a much lower rate. Furthermore, I find that the trans-binary celebrities, Caitlyn Jenner and Elliot Page, are deadnamed persistently in around 17.5% of tweets following their COEs. Following other work in computational sociolinguistics, content analysis reveals that misgendering and deadnaming tweets are significantly correlated with lexical items indicative of hate speech, biological essentialism, and binary gender terms. I show that while trans-binary celebrities' identities are pronominally ratified at higher rates than trans-nonbinary celebrities, tweets that misgender them also exhibit stronger signs of dehumanization. I ultimately argue that these patterns are driven by cisnormativity, an ideology that maintains a rigid relationship between individual identity and binary gender in a coherent and immutable manner. The patterns unearthed in this study demonstrate the power of cisnormativity as an organizing principle.

ACKNOWLEDGEMENTS

In short, it's like learning to swim again. From where I was before, I caught sight of a faraway beach; the distance was great, the waves were daunting, and I wondered if I'd stay afloat. But I wanted to go the distance, and I sort of had to, so I dove in. It wasn't too bad at first. I built speed slowly, and felt like I was getting somewhere, but it was never easy. I told myself to breath, but only succeeded sometimes; often, I needed prompting from others. At times I felt near my destination. But, Murphy's Law proving itself yet again, I would look up from the waves just to realize how far the shore remained. And so I kept kicking.

Along the way, I learned about mettle, vulnerability, and the tricking balancing act between the two. I learned, begrudgingly, when to flip onto my back and rest. I learned how to surround myself with good tides, how to paddle hard, how to take stock of things. I learned how to keep my head above water. In short, I learned how to swim.

Most of this learning was not independent, of course, but completed with the gracious help of others. Thank you to Annette and Rob, who started the fire and kept it going, respectively, collectively. You both inspire me so much, and I feel beyond lucky that I get to learn from and alongside you. Thank you to the rest of the Linguistics Department, especially Prof. Masaya Yoshida, Prof. Jennifer Cole, Prof. Gregory Ward, SocioGroup, CoLLab, and the many other linguists I met along the way. It takes a village!

Thank you also to my friends, who maybe never really understood what I was studying, but were nonetheless there to celebrate with me at every milestone along the way. I love you all. Finally, thank you to my family, to whom everything is owed, always. If I missed a call, a text, or a card, it's because I was too busy counting my blessings.

This work is dedicated to anyone ever cast down by society. Thank you for changing it.

Glossary

Biological essentialism A cisnormative principle that centers anatomical features in the notion of sex and social expression in the notion of gender, polarizing the two and reducing the former to a biological imperative that precedes any sort of sociocultural influence.

Biosocial gender The aspect of gender experienced by an individual internally based on socialization, phenotype, cultural norms, and other factors.

Cisgender A self-identification with the sex/gender category assigned to one at birth.

Cisnormativity An ideology and organizing principle that maintains a strict relationship between individual identity and an immutable gender binary. Cisnormativity maintains only cisgender individuals as natural, dominant, and good. Related terms include *cisgenderism* and *transphobia*.

Coming-out event The moment of declaration by the trans celebrities in this study that their inner, self-realized experience of gender matches their social, identity-oriented expression of that gender.

Conceptual gender How people perceive, interpret, and actively construct the gender of others. Often, this dimension draws upon sociocultural norms surrounding gender, gender expression, and gender identity.

Deadnaming When a speaker uses the former name of a transgender individual (one often assigned at birth) rather than the individual's correct, gender-affirming name.

Epicene they The usage of singular *they* to refer to a generic antecedent when gender is unspecified, unknown, or irrelevant. Compare to epicene *he*, *he or she*, and *s(he)*.

Gender expression The aspect of gender that comprises the ways an individual chooses to appear and behave in relation to cultural expectations for gendered dress, speech, behavior, and embodied aspects.

Gender identity The aspect of gender that describes an individual's sense of self considering the alignment between biosocial gender, gender expression, and external conceptions of their gender.

Gender self-determination An individual's right to define their own identities and relationships to sex, sexuality, and gender.

Gender-inclusive language Language reform motivated by social activism that, at many different levels, makes possible the linguistic expression of trans and nonbinary gender identities. Related terms include *gender-neutral language*.

Listing pronouns The folk meaning of the phrase 'Michael uses *he/him* pronouns', which is intended to convey an individual's pronoun suite for desired uptake by others – present and future interlocutors – who might rely on this information in the process of reference.

Misgendering When a speaker selects a third-person pronoun in the process of reference that does not align with the self-asserted pronoun suite listed by their referent as part

of gender expression.

Nonbinary A self-identification with a sex/gender category that is not (exclusively) woman or man. Related terms include *genderqueer* and *genderfluid*.

Nonbinary they The usage of singular *they* to refer to a specific individual who identifies as nonbinary and lists *they* pronouns. Not all nonbinary people list *they* nor is *they* listed only by nonbinary people.

Notional gender The system of grammatical gender present in English, whereby language users draw upon temporally-bound cultural expectations, norms, and notions surrounding gender and sex when selecting the gendered forms of pronouns and some lexical items.

Transgender A self-identification with a sex/gender category different from that assigned to one at birth, including binary-identifying ('woman','man'), nonbinary, and other identities.

Using pronouns Instances where language users make reference to individuals through gendered third-person pronouns; when they use pronouns in this way, language users assert and assign the conceived gender of their referent.

Note: These definitions are merely those used consistently throughout this work, and do not necessarily denote the 'correct' meaning for any of these terms.

TABLE OF CONTENTS

List of Figures 9

List of Tables 10

Chapter 1 :· Introduction 12

Chapter 2 :· Language and Gender 16

2.1 Defining gender 16

2.2 English: A gendered grammar? 20

2.3 Expanding gender, neutralizing language: Cross-linguistic evidence 23

2.4 The creation and subversion of cisnormativity 26

Chapter 3 :· Pronouns and Gender 30

3.1 Pronouns in English 30

3.2 Pronouns enmeshed: An ideological perspective 35

3.3 The great pronoun shift: Changes in singular *they* 39

Chapter 4 :· Harmful Speech: Misgendering and Deadnaming 46

4.1 The regulation of transgender identities 46

Chapter 5 :· A Computational Sociolinguistic Approach to Gender 53

5.1 Computer-mediated language variation and change 53

5.2 Research questions 56

Chapter 6 :· Data 59
6.1 Seven celebrity corpora 59
6.2 Data collection 65

Chapter 7 :· Methods 66
7.1 Statement of positionality 66
7.2 Pre-processing 67
7.3 Filtering 69
7.4 Lexical processing 75

Chapter 8 :· Results 77
8.1 Distributional analysis 77
8.2 Content Analysis 86

Chapter 9 :· Discussion 102

Chapter 10 :· Conclusion 112
10.1 Summary of findings 112
10.2 Limitations 114
10.3 Future work 114
10.4 Ethical considerations 115

References 115

Appendix A :· Terms Used in ALTENT Filter 131

Appendix B :· Twitter bio field demographic data 132

Appendix C :· Additional fightin' words lexical analyses 133

LIST OF FIGURES

8.1 **Affirming name uptake for trans-binary celebrities across entire analysis period, with week 0 representing the coming-out event.** . . . 78

8.2 **Mean affirming pronoun rate by week for all celebrities. Data for comparison celebrities spans entire analysis period, while data for target celebrities is POST-COE.** 82

8.3 **Affirming pronoun uptake for target celebrities across entire analysis period, with week 0 representing the coming-out event.** 86

8.4 **Effect of binary measures on gendered pronominal usage for all celebrity groups. Positive coefficients represent tweets that were more likely to affirm, while negative coefficients represents those more likely to misgender. Only significant measures (p-value≤0.05) are displayed.** 100

LIST OF TABLES

7.1 Increasing selectivity throughout the tweet filtering process. PRONOUN analysis occurs at ALTHAND filter level; NAME analysis occurs at NAME filter level. 70

7.2 Percent change in mean affirming pronoun rate across most selective filters follows expected patterns. 71

8.1 NAME ADF tests for trans-binary celebrities in GROSS, PRE-COE, and POST-COE conditions. Test-statistic lower than $\tau 3$ rejects null hypothesis, suggesting stationarity. 79

8.2 Affirming name rate descriptive statistics for trans-binary celebrities in PRE-COE and POST-COE conditions. 79

8.3 Usage rates by pronoun form for comparison celebrities across entire analysis period. AFFIRM-RATE represents affirming pronoun usage, ALT-RATE represents misgendering usage, and THEY-RATE represents singular *they* and instances of plural *they* that slipped through the filtering process. 81

8.4 Affirming pronoun rate descriptive statistics for target celebrities in PRE-COE and POST-COE conditions. 83

8.5 Affirming pronoun rate descriptive statistics for comparison celebrities across entire analysis period. 83

8.6 PRONOUN ADF tests for target celebrities in GROSS, PRE-COE, and POST-COE conditions. Test-statistic lower than $\tau 3$ rejects null hypothesis, suggesting stationarity. 84

8.7 PRONOUN ADF tests for comparison celebrities across entire analysis period. Test-statistic lower than $\tau 3$ rejects null hypothesis, suggesting stationarity. . 85

8.8 Crossover NAME and PRONOUN results for trans-binary celebrities across Aff (rate=1.0), Mix (0.0<rate<1.0), and Mis/Dead (rate=0.0) conditions. . . 86

8.9 Trans-binary VADER sentiment scores across PRE-COE and POST-COE conditions in the NAME analysis. 87

8.10 NAME lexical associations for trans-binary celebrities by DEADNAME (name rate=0.0) and NAME-AFFIRM (rate=1.0) in the POST condition. 88

8.11 Target celebrity VADER sentiment scores by pronoun rate across PRE-COE and POST-COE conditions in the PRONOUN analysis. 90

8.12 Comparison celebrity VADER sentiment scores by pronoun rate across entire analysis period in the PRONOUN analysis. 90

8.13 NAME binary logistic regression in POST-COE condition. Data aggregated from equal number of tweet-level observations for trans-binary celebrities. . . 95

8.14 PRONOUN binary logistic regression for target celebrities in POST-COE condition. Separate data aggregated from equal number of tweet-level observations for trans-binary and trans-nonbinary celebrities, respectively. 97

8.15 PRONOUN binary logistic regression for comparison celebrities across entire analysis period. Cisgender data aggregated from equal number of tweet-level observations for Doja and Holland. 98

B.1 User demographic data for target celebrities, ordered chronologically by time of COE. 132

C.1 PRONOUN lexical associations for Elliot Page post-COE. 134

C.2 PRONOUN lexical associations for Demi Lovato post-COE. 135

C.3 PRONOUN lexical associations for Sam Smith post-COE. 136

C.4 PRONOUN lexical associations for Caitlyn Jenner post-COE. 137

CHAPTER 1
INTRODUCTION

As societal notions of gender change, so too do the aspects of language that encode it. In many ways, these changes have been sparked by the need to linguistically account for lived identities that challenge both traditional understandings of gender and the ways in which gender becomes represented in language. Zimman (2020) identifies the emergence, recognition, and study of these innovations as part of a global and ongoing 'transgender moment'. From the gender-unmarked pronoun 'TA' in written Mandarin (Sluchinski, 2019) to the use of _ as a gender-inclusive morpheme in Slovene (Popič & Gorjanc, 2018), these changes are occurring at many different levels in languages across the world.

In English, there are changes occurring in both pronoun and practice. On one hand, change in the scope of *they* represents the most recent in a long line of socially-motivated shifts in the language's pronominal system (Bodine, 1975; Curzan, 2003; Silverstein, 1985). Aside from *they*'s newfound dominance over *he* and *he or she* as the most common epicene pronoun in English (Balhorn, 2004; Lascotte, 2016; Noll, Lowry, & Bryant, 2018), researchers have begun to explore the usages of and attitudes toward *they* as a singular form to represent nonbinary identities – termed here as nonbinary *they* (Ackerman, 2019; Camilliere, Izes, Leventhal, & Grodner, 2021; Conrod, 2019; Hekanaho, 2020).

On the other hand, specific gender-related language practices have sparked much metalinguistic discourse in English-speaking society as they entered the linguistic mainstream. Some of these practices, such as pronominal misgendering (Conrod, 2020) and deadnaming (Turton, 2021), function to reject self-determined gender identities (Koyama, 2003), par-

ticularly those of trans people. Here, misgendering occurs when a language user selects a third-person pronoun in the process of reference that does not align with the self-asserted pronoun suite listed by their referent as part of gender expression. Deadnaming is a more contextualized practice, whereby in the process of reference a language user selects the former name of a transgender individual (one often assigned at birth) rather than their correct, gender-affirming name. Meanwhile, other language practices challenge the belief that gender identity can and should be assumed, such as the listing of third-person pronouns as part of introductions (Zimman, 2019) or on social media profiles (Jones, 2021). These developments raise new questions about the relationship between proper names, third-person pronouns, and gender identity.

The present study attempts to shed light on these questions through the utilization of computational linguistic methods to provide large-scale evidence identifying the distribution and lexical content of two gendered language practices. In a corpus analysis of 7m tweets, I demonstrate the extent to which Twitter users linguistically ratify the gender identities of seven celebrities of diverse genders through the use of affirming proper names and third-person pronouns. The 'target' data set comprises four celebrities, two of whom are nonbinary and list *they* pronouns (the trans-nonbinary group) and two who are transgender and list the binary pronouns *she* and *he* (the trans-binary group). The 'comparison' data set comprises both transgender and cisgender celebrities who list binary pronouns and do not have a publicly-documented coming-out event (COE). Following Zimman (2009), COE refers here to the moment of declaration by the trans celebrities in this study that their inner, self-realized experience of gender matches their social, identity-oriented expression of that gender.

I ultimately find that, in the time period immediately following these celebrities' COEs, usage of nonbinary *they* to refer to the trans-nonbinary group trails behind the usage of

binary pronouns *he* and *she* to refer to the trans-binary group. However, both of these groups are misgendered more than the comparison celebrities. Furthermore, I find that the two trans-binary celebrities, Caitlyn Jenner and Elliot Page, are deadnamed persistently in around 17.5% of tweets on average following their COEs. These quantitative findings reveal significant disparities in gender-affirming nominal and pronominal reference that align with the gender identity of the (celebrity) referent and the historical time in which they came out.

Following other work in computational sociolinguistics, I also show how the ideological aspects of misgendering and deadnaming are representative of dehumanization (Haslam, 2006; Mendelsohn, Tsvetkov, & Jurafsky, 2020), whereby trans individuals are denied the agency to determine their own gender identity. Misgendering/deadnaming tweets are significantly more likely than gender-affirming tweets to contain lexical items indicative of hate speech, biological essentialism, and a binary understanding of gender. I ultimately argue that these lexical patterns are driven by cisnormativity, an ideology that maintains a stable, coherent, and immutable relationship between individual identity and a rigid gender binary (Borba & Milani, 2017; Ericsson, 2018; McInroy & Craig, 2015).

Like other ideologies, cisnormativity is productive; it encapsulates how individuals interpret the gender of others and determines their subsequent social and linguistic behavior (Gal & Irvine, 2019). This project joins more recent scholarly work that has begun to inspect this ideology as the sum of its discursive parts (Ansara & Hegarty, 2014; Borba & Milani, 2017; Ericsson, 2021). Considering that third-person pronouns and proper names are among the first acts of linguistic self-determination that trans individuals make (Konnelly & Cowper, 2020), misgendering and deadnaming emerge as two loci where cisnormativity can regulate – through erasure, denial, and dehumanization – the expression of transgender identities (Irvine & Gal, 2000; Zimman, 2019). The different distributional and lexical patterns that

emerge in this study between transgender and cisgender celebrities inform us as to the common linguistic strategies that not only reflect but actively construct this ideology. Within the transgender celebrities, further differences between those who list nonbinary *they* and those who list binary pronouns (*she* or *he*) point us towards some of the ideology's more fine-grained mechanisms: as linguistic ratification of (trans)gender identity increases, the presence of dehumanizing speech in the tweets that misgender and deadname these celebrities increases in tandem.

This thesis represents an uncommon bridging of theory and methodology; I aim to contextualize its focus within the linguistic literature on (trans)gender and its methods within the burgeoning field of computational sociolinguistics (Nguyen, Doğruöz, Rosé, & de Jong, 2016). I provide in Chapter 2 a broad overview of how gender and language interface before turning specifically to the status of gender features in English's third-person pronouns in Chapter 3, where I also discuss the ongoing change in singular/nonbinary *they*. Then, in Chapter 4 I explore misgendering and deadnaming as language practices that reinforce cisnormativity. Following, I discuss throughout Chapter 5 the treatment of gender in computational sociolinguistic work and present the research questions guiding this thesis. Chapter 6 introduces the seven celebrities that comprise the corpus, the purpose-oriented tweet filtering process, and the computational methods used to analyze the data. Finally, I share the results of my distributional and content analyses in Chapter 7, after which I contextualize them theoretically in Chapter 8. Chapter 9 presents a brief conclusion, limitations of the study, and directions for future work.

CHAPTER 2
LANGUAGE AND GENDER

The two language practices discussed in this work – deadnaming and pronominal misgendering – gain symbolic power through their relationship to gendered power structures. As such, it is necessary to articulate my theoretical approach to gender. In this chapter, I define the dimensions and identities of gender relevant to the present study. Then, I explore how language and gender interface, with a special focus on English's notional system of gender. I subsequently turn to cross-linguistic evidence, particularly from Swedish, that explores the impact of gender-inclusive language reform from both a linguistic and an ideological perspective. Finally, I explore how negative stances towards gender-inclusive language reform are emblematic of cisnormativity, an organizing principle of gender, language, and identity – the mechanisms of which scholars are only recently beginning to dismantle and inspect.

2.1 Defining gender

Gender is a complex social phenomenon. In the West, the prevalent view of gender historically has been one that takes biological sex to be binary: male or female. (Morgenroth & Ryan, 2018). These two sexes, to which individuals are assigned at birth, subsequently form the basis for the gender binary: man or woman. In this way, it is possible to view gender as a social classification system. Starting with the sex assigned by medical establishments at birth, individuals are organized and accordingly characterized (by themselves, by others, and by institutions) through their position on either side of the gender binary.

Since at least the late 20th century, however, scholars have worked to dissect the many

dimensions that constitute gender, challenging the prevailing Western model that links sex, gender, and sexuality as essential, immutable qualities. Butler (1990) proposes in her seminal work that neither gender nor sex are essential, biologically determined classifications. Rather, gender is achieved through repeated performance. That is, we draw upon a rich tapestry of social behaviors, physical materials, and other, embodied aspects to both perform our own gender and interpret the gender of others. In turn, these acts create the gender categories of 'woman' and 'man' in society. These categories, along with the process by which individuals ascribe to them, are culturally variable and have the potential to change over time – often, as ideas about sex, gender, and sexuality shift as well.

The dominant classifications – to some, the only classifications – within the Western system of gender remain 'man' and 'woman.' However, there is an emerging recognition that many people do not fit so neatly within this binary and that the binary is in fact a spectrum. Some individuals reject this spectrum entirely, while others demonstrate that individual relationships to it can change throughout a lifetime or even setting to setting in response to contextual factors. While individuals who transgress the man-woman binary have always existed (Vincent & Manzano, 2017), they have gained increasing salience over the last two decades in Western society as part of an ongoing 'transgender moment' (Zimman, 2020). To account for these lived experiences, scholars have again begun to rethink the relationship between sex, gender, and sexuality.

There are many factors that influence how gender is determined, the most significant perhaps being sex. The relationship between sex and gender has long been operationalized in academic work through the 'coat rack' metaphor (Zimman, 2014). In this model, the rack is framed as a largely invariable object, like sex, that does not change in use, shape, or appearance over time. Onto this rack, a variable coat – a given culture's expectations for

'man' or 'woman' – is placed (Delphy, 1993). In the perspective of this metaphor, gender is sex dressed up.

While the theoretical decoupling of sex and gender has enjoyed much scholarly uptake in queer, feminist, and sociolinguistic work (e.g. Cameron & Kulick, 2003; Eckert, 1989; Hall, 1995), Zimman (2014) draws upon ethnographic fieldwork with American trans populations to problematize approaches that adopt a 'coat rack' metaphor. The most relevant of these critiques here is the disguising of sex as the 'inevitable truth of nature' which ultimately reinforces biological essentialism. Though it is in the rack's nature to have a culturally varying dimension – in this case, gender – a system that treats only the bodies of cisgender people as naturally and biologically male or female implies that, however a transgender person might experience or transform their body, their 'sex', it will never be natural. Thus, biological essentialism emerges here as a force that centers anatomical features in the notion of sex and social expression in the notion of gender, polarizing the two and reducing the former to a biological imperative that precedes any sort of sociocultural influence (Stryker, 2008; Zimman, 2014).

For the purposes of this paper, I adopt the definitions of gender offered by trans scholars such as Zimman (2020). **Cisgender** can describe individuals who self-identify with the sex/gender category assigned to them at birth. Meanwhile, **transgender** can describe individuals who do not self-identify with the sex/gender category assigned to them at birth. After decades of usage in-community, this term gained widespread use as an identity label in the 1990s (D. Valentine, 2007). **Nonbinary** individuals are in this way also transgender, as they reject specifically the entire rigid binary that was used to classify their sex/gender at birth.

Given the extreme diversity with which gender and sex have been conceptualized in

the social sciences, it is important to use clear and consistent terminology across empirical examinations of gender, sex, and sexuality (McElhinny, 2003). To this end, this paper adopts an approach to gender that centers transgender experiences, beliefs, and notions of gender. One central tenant of transgender activism and theory is the precept of self-determination. Self-determination in this context describes an individual's right to define their own identities and relationships to sex and gender (Zimman, 2014). This approach entails the belief that no individual should be pressured into or out of decisions regarding their gender identity, gender expression, or body in order to become a 'real' or 'natural' man or woman (Koyama, 2003).

In this view – one that challenges understandings of sex and gender that center cisgender experiences – the individual is the ultimate authority on their **biosocial gender**, which describes the aspect of gender *experienced* by a individual based on phenotype, socialization, cultural norms, gender expression, and gender identity. Biosocial gender is the first in a series of gender terms proposed by Ackerman (2019). This notion of gender contrasts with **conceptual gender**, which is the aspect of gender that is *inferred* by a perceiver and subsequently *assigned* to an entity in discourse. We further define **gender expression** as the way an individual appears and behaves in relation to cultural expectations for gendered dress, speech, behavior, etc. and **gender identity** as an individual's sense of self considering the alignment between biosocial gender, gender expression, and external conceptions of their gender.

This latter point is critical: gender in part emerges through interaction. An individual may determine the success of their performance of gender – in expression and in identity formation – through the linguistic, attitudinal, and behavioral responses of interlocutors. In this sense, gender identity is a 'dialogic construction' (Bucholtz & Hall, 2004; Zimman, 2017).

If individuals fail to adequately perform their gender in the mind of the perceiver, they can face negative consequences, stigma, and even violence (Butler, 1990). Though this paper promotes a self-determined view of biosocial gender, it nevertheless focuses on the linguistic mechanisms used to assert gender, the ideologies that mediate external reception of these assertions, and the linguistic practices that, in turn, ratify or reject certain expressions of gender.

2.2 English: A gendered grammar?

Language is arguably the most dynamic piece of the gender tapestry, as it plays a central role in both the performance and perception of gender. Whether hearing a voice or processing a third person pronoun, we use linguistic information to conceive gender (Gal, 1989). Since the feminist critique of linguistics began in the mid 1970s (Bodine, 1975; Kramarae, 1981; Lakoff, 1973), scholars have explored the myriad ways language can encode, represent, and perform gender. Within this great sea of work, linguistic notions of gender have been wielded towards diverse, often contrasting ends. As such, it is necessary to discuss how exactly gender interfaces with language.

When linguists refer to gender, they are often referring to formal syntactic and/or semantic features that are morphosyntactically defined (Ackerman, 2019). Languages that have gendered noun classes ('feminine','masculine', and 'neuter', for example) also have the morphosyntactic category 'gender' (Curzan, 2003). Corbett (1991) proposes that languages with gender feature systems generally fall into one of two categories. The first of these is strict semantic gender systems, such as Djirbal (Silverstein, 1985), in which the meaning of the noun determines its gender, and vice versa, for all or almost all lexical items. The second is grammatical gender systems, such as English, where most nouns do not follow semantic

assignment rules; the nouns that are gender-marked depend on morphological factors for their assignment.

Though English once had a more complete system of grammatical gender, it now only marks gender on third-person pronouns, names, and a handful of lexical items ('actress','cowboy', and 'congresswoman', to name a few). As such, English is commonly proposed in the linguistic literature to possess a 'natural' gender system, though this terminology reflects problematic assumptions made throughout decades of research on the topic. McConnell-Ginet (2014) argues that the depiction of English as a language with natural gender reinforces biological essentialism[1]: inherent in most discussions of 'natural gender' is the assumption that the biological sex of a noun's referent determines its gender agreement (c.f. Curzan, 2003). Assumed in less considerate approaches that adopt natural gender is the proposition that it is the natural 'sex' of a given human referent that a language user draws upon to determine which gendered pronominal or lexical form they will use ('she' vs. 'he' or 'chairman' vs. 'chairwoman', e.g.). The previous distinction between biosocial and conceptual gender makes immediately clear the problems with this approach. In reality, a language user's selection of gendered forms during reference hinges entirely on the *conceptual gender* that they assign to their referent, which may or may not align with the referent's self-determined biosocial gender. For example, the use of *he* reflects only a masculine conceptual gender feature that the language user *assigns* to their referent. These assignments are dependent on beliefs surrounding sex, gender, and sexuality at a given time. McConnell-Ginet (2014) thus proposes that instead of natural gender English possesses a **notional gender** system, corresponding to the cultural expectations, norms, and *notions* that a language user relies upon in conceiving and assigning the gender of their referent. In this way, pronominal usage

[1] And, as we will see, cisnormativity as well.

cannot be understood without considering the mutability of conceptual gender. Whereas natural gender implies that the expression of gender in language is dictated by biological imperatives, notional gender captures the social and dialogic dimensions of gender that are culturally variable, constantly shifting, and produced through interaction.

To this end, Curzan (2003) and N. S. Baron (1971) offer a thorough documentation of the many changes grammatical gender has undergone in English. Their work demonstrates how the minimal marking of gender in English leads to complexity when identifying concrete changes in the linguistic structures that reflect new ideas and experiences surrounding gender. However, a notional gender approach allows us to consider the ways in which shifting cultural tides surrounding gender and sex can manifest in innovative uses of language.

One recent development in this regard is documented by the ongoing change in English's third person singular pronouns, a change that is largely driven by trans community members, activists, and linguists striving to make English more reflective of expansive experiences of gender (Calder, 2020; Zimman, 2020). As such, trans communities are largely responsible for the evolution currently taking place in English; in myriad ways, the language is evolving to account for and reflect lived transgender experiences. Though we return to this dimension of English in Chapter 3, it is important to note here that, for transgender individuals, gendered names and pronouns are some of the most significant sites at which transgender people stake their gender identity (Conrod, 2020; Zimman, 2017).

In this section, I clarified the ways in which gender features manifest in English as part of a system of notional gender (McConnell-Ginet, 2003). However, English is only one of many languages that in some way encodes gender. Though gender features are by no means a linguistic universal (Corbett, 1991), approaches to gender-inclusive language reform can vary widely depending on the type of gender system present in a given language. As we will

see in the next section, the ongoing 'transgender moment' (Zimman, 2020) is truly global – culturally and linguistically.

2.3 Expanding gender, neutralizing language: Cross-linguistic evidence

As attitudes towards and ideologies surrounding gender expand in many cultures, so too are their respective languages undergoing change. While some of this work comprises gender-neutral language reform – whose main purpose is to dismantle androcentrism, or the centering of (cisgender) men, in some arenas of linguistic gender (Bodine, 1975) – this paper takes a more expansive view. Any attempt to make language more gender-neutral is an attempt to make language gender-inclusive; if linguistic manifestations of the man-woman binary can be deconstructed, a fuller expression of nonbinary trans identities is made possible (Darwin, 2017; Hekanaho, 2020).

Perhaps the most complete literature on gender-inclusive language change comes from Swedish, where researchers have documented shifting ideologies surrounding the introduction, uptake, and usage of the gender neutral third person pronoun *hen* (Hord, 2016; Lindqvist, Gustafsson Sendén, & Bäck, 2016; Sendén, Bäck, & Lindqvist, 2015; Vergoossen, Renström, Lindqvist, & Sendén, 2020). Like English, Swedish is a notional gender language (Hord, 2016). There, *hen* emerged from trans and queer communities as a novel pronoun, or neopronoun. This development received much attention and uptake in Swedish society over the last decade. While neopronouns exist in English as well, they have not yet received widespread adoption (Hekanaho, 2020).

Hen began to enter the Swedish mainstream in 2011 as both a generic or 'epicene' pronoun and as a pronoun to represent nonbinary identities. In this way, the addition of *hen* rounded out *hon* and *han*, the language's masculine- and feminine-marked pronouns, respectively. In a

longitudinal study measuring attitudes toward *hen* using a combination of open-response and Likert-scale questionnaire items, Sendén et al. (2015) observed explicit negative sentiment towards the inclusion of the new pronoun in 2012. These attitudes changed dramatically and rapidly over time; by 2013, a majority of participants in their sample no longer expressed negative attitudes toward *hen*. By 2015, only a very small sample of Swedish participants retained negative views. However, the researchers also observed that uptake in the usage of *hen*, whether referring to nonbinary individuals or used as an epicene, consistently lagged behind the generally positive attitudes surrounding the neopronoun.

Many social factors significantly predicted attitudes towards and usages of *hen*. Positive attitudes and usage of *hen* were associated with younger participants, left-leaning participants, participants who expressed more interest in gender issues, and participants who espoused less sexist views (Sendén et al., 2015). Vergoossen et al. (2020) further contextualizes these findings by exploring the ideological dimensions that comprise criticisms of and negative attitudes towards *hen*. Participants were more likely to take issue with the nonbinary usage of *hen* rather than its more common epicene usage. Additionally, a qualitative coding analysis revealed that the vast majority of arguments against *hen* were comprised of those that defended the linguistic status quo and explicitly sexist or transphobic views, albeit to a lesser extent. These findings mirror previous attempts at gender-inclusive language reform in the United States surrounding epicene *they* (Parks & Roberton, 1998).

Other work on Indo-European languages such as French (Kosnick, 2019) Spanish (Bonnin & Coronel, 2021; Papadopoulos, 2018), Portuguese (Borba, 2019; Verguiero, 2016), and Slovene (Popič & Gorjanc, 2018) articulate the differential challenges faced by activists working within languages that exhibit far more complete systems of grammatical gender. Hord (2016) shows how these languages, in comparison to those that exhibit a domain-

scarcity in grammatical gender like Swedish and English, struggle to implement gender-inclusive language reforms, especially in the face of more conservative gender notions. For example, Borba (2019) traced how the usage of the letter X as a gender-inclusive morpheme in Brazilian Portuguese incited a 'linguistic guerilla war' that resulted in increased public animosity between some of the country's political constituencies. Relatedly, Kosnick (2019) tracked the highly militaristic response of the Académie Française, a prescriptive French-language institution, towards a package of pronouns and articles – known as *écriture inclusive* – that were designed to make French more gender-inclusive. Finally, Popič and Gorjanc (2018) conducted a corpus analysis of Slovenian newspapers to explore the implementation of the underscore as a gender-inclusive morphological innovation, similar to the X in Brazilian Portuguese. They show that, despite intense challenges to its validity and usability, use of the innovative underscore morpheme is increasing over time.

This cross-linguistic evidence exemplifies the typical attitudes and social patterns driving the adoption of gender-inclusive language use. Though these ideological patterns are discussed more fully in Chapter 3.2, some conclusions can be drawn now. First, Bonnin and Coronel (2021) note that attitudes surrounding gender-inclusive language are generally characterized in their relationship to two variables: prescriptive language (relative openness to linguistic innovation and change) and conservative gender notions (relative commitment to gender as a binary category). Second, adoption of gender-inclusive language strategies often lag behind positive attitudes (Sendén et al., 2015; Vergoossen et al., 2020). Third, there exists a dyadic relationship between expansive conceptions of gender and gender-inclusive language reform (Hord, 2016).

2.4 The creation and subversion of cisnormativity

Attempts at gender-inclusive language reform are almost always met with strong social, political, and cultural opposition. The driving force behind this opposition, I argue, is cisnormativity: an ideology and organizing principle that maintains cisgender identities as natural, dominant, and good (Ericsson, 2018). Cisnormativity represents a package of ideas, perspectives, and behaviors that maintains close proximity to other hierarchical regimes of social power, such as sexism and heteronormativity (Hall, Borba, & Hiramoto, 2021). Though scholars have been deconstructing these latter ideologies for decades (Calder, 2020), only recently have academics begun to explore the ways cisnormativity is produced and perpetuated (Turton, 2021). Challenging cisnormativity entails moving away from critiquing or justifying individual experiences of biosocial gender by focusing instead on understanding the ideological and discursive practices that marginalize transgender experiences (Ansara & Hegarty, 2014).

Drawing upon recent anthropological, linguistic, and medical work towards identifying the mechanisms of cisnormativity (Bauer et al., 2009; Borba, 2019; Borba & Milani, 2017; Ericsson, 2018, 2021; Hobaica, Schofield, & Kwon, 2019; Hornscheidt, 2015), I propose a taxonomy of its ideological principles. Under cisnormativity,

- **The gender of all individuals can be classified using the man-woman binary.** The only two gender classifications in this system are 'man' and 'woman'. By presupposing the gender binary, cisnormativity continuously affirms it (Hornscheidt, 2015). In order to fully participate in society and avoid being pathologized, individuals must correspond to one of these categories. Borba and Milani (2017), for example, draw upon fieldwork in a Brazilian gender clinic to demonstrate how cisnormativity mani-

fests itself in diagnostic medical assessments. To access gender-affirming health care, transgender patients must prove that their identities correspond directly to the man-woman binary. These institutional practices in turn frame the evaluations made by health professionals within the cisgender assumption that gender and sex align naturally and immutably. This tenet of cisnormativity is also an ethnocentric force: gender has been conceived of and classified in countless ways throughout human history, far beyond the two categories of 'woman' and 'man' (Vincent & Manzano, 2017).

- **The man-woman gender binary corresponds to the male-female sex binary, to which individuals are naturally assigned at birth.** In this system, the sex that medical institutions assign children at birth immediately forms the basis for their gender identity (Zimman, 2015). Diverse instantiations of phenotype or genitalia are made to fit within this one-to-one mapping (Stryker, 2008). From birth on, individuals are expected to perform gender in ways that accord with the sex assigned at birth. If an individual were to prevent novel expressions of gender, they would never become a 'real' or 'natural' member of the presented category if it was not the one assigned to them. In this way, the only true exemplars of a given gender are those who were assigned to its corresponding sex at birth (Hobaica et al., 2019).

- **An individual's gender identity, expression, and physical features must cohere intelligibly.** Cisnormativity enforces an understanding of sex and gender that is intelligible and immutable. Ericsson (2021) articulates this notion in terms of a coherency constraint. A cisnormative perspective assumes that the relationship between biosocial gender, gender identity, gender expression, and anatomical features must cohere in understandable, even predictable ways. This thinking also manifests

itself in conceptual gender, whereby these features are utilized in the attribution of others' gender: if the relationship between conceptual and biosocial gender is unclear or indeterminable, then the other is not the gender they self-proclaim to be. The idea that it is offensive to ask people what their gender identity is reflects this constraint – that an individual's gender expression must felicitously correspond to their gender identity (Zimman, 2017). To affirm that a person's gender identity is not obvious is to suggest that they have failed to properly express their gender. This precept is intimately related to sociocultural norms surrounding gendered speech, dress, and physical appearance in particular, as it is these features that are rapidly accessed in the determination of conceptual gender (Ackerman, 2019). In this way, cisnormativity also discards self-determination as an organizing principle in relationship to gender and sex.

- **An individual cannot change their gender identity, as gender is experienced in a consistently stable fashion throughout a lifetime.** Within a one-to-one mapping between sex and gender initiated at birth, no potential exists for changes in gender or sex, whether these are contextually dependent or more crystallized changes (Hornscheidt, 2015). To realize this assumption, cisnormativity must regulate, inhibit, and erase transgender identities, who provide evidence that experiences of gender can shift throughout one's life (Bauer et al., 2009).

As Cameron and Kulick (2003) emphasize in their approach to heteronormativity, just because an ideology goes largely unmarked does mean that it holds a natural place in the world. Rather, these ideologies are actively produced through interaction at the individual and institutional level. Recent research has thus begun to identify the ways in which cisnormativity is perpetuated in society, especially through language practices (Zimman, 2017). In

Chapter 4, I argue that deadnaming and misgendering are the most prevalent and harmful of these language practices. This is critical to the present study, as large-scale patterns of gendered language practices allow us to understand the extent to which Twitter users, in drawing upon conceptual gender to determine the gendered pronouns and names used to refer to trans celebrities, perpetuate cisnormativity.

CHAPTER 3
PRONOUNS AND GENDER

Pronouns are functional elements (Conrod, 2017). As such, they typically constitute a closed class, meaning that they only rarely admit new members. However, recent evidence from Swedish (Sendén et al., 2015; Vergoossen et al., 2020) and from throughout the history of English (Balhorn, 2004; Curzan, 2003; Silverstein, 1985; Wales & Katie, 1996) suggest that pronouns are not as resistant to change as grammarians might expect. Rather than a matter of historical chance, I argue that pronouns are especially susceptible to change (in membership and in scope) due to their implication in personal reference, a consistently social action. Any aspect of a given language is susceptible to change if there is enough impetus at the metalinguistic level to do so (Bodine, 1975); the relevance of the pronominal system's status as a closed class of functional elements only goes so far in explaining why change in this arena might be more difficult. In this chapter, I first review the status of pronouns in English before discussing the pronominal system as a site for ideological contest. Then, I present recent work from the psycholingusitic, sociolingusitic, and syntactic literature that captures systematic pronominal change in progress.

3.1 Pronouns in English

This study focuses on two separate linguistic practices related to third-person pronouns: their listing by individuals as part of gender expression and their subsequent use by language users in instances of pronominal reference. By **listing pronouns**, I mean to capture the folk meaning of the phrase 'Michael uses *he/him* pronouns', which is intended to convey an

individual's pronoun suite for desired uptake by others – present and future interlocutors – who might rely on this information in the process of reference. Listing pronouns, whether in a Twitter bio, Zoom handle, or as part of an introduction, is an act of assertion rather than of preference. By **using pronouns**, I describe the instances where language users make reference to individuals through gendered third-person pronouns; when they use pronouns in this way, language users assert their conceived gender interpretation and assign it to their referent.

In English, the most common third-person pronoun suites are *she/her/hers*, *he/him/his*, and *they/them/theirs*, though these are not the only ones. Some individuals instead list novel pronouns, or neopronouns, such as *ze/hir/hirs*, as part of their gender expression. However, research has demonstrated that neopronouns are currently both used and accepted at much lower rates than the singular *they* suite (Bradley, Salkind, Moore, & Teitsort, 2019; Hekanaho, 2020; Sheydaei, 2021). Importantly, the pronouns that an individual decides to list do not necessarily and/or directly map onto the gender identity of that individual (Zimman, 2017). I see this reflected in my data, as some Twitter users package together multiple pronoun suites – *she/they*, for example – when listing their pronouns in their account biographies. Indeed, pronoun listing as a practice of self-identification has rapidly spread throughout the American mainstream over the last decade. Jones (2021) observed that, out of all tokens, the pronouns *she*, *her*, *he*, and *him* exhibited the most positive increase in prevalence in a longitudinal analysis of Twitter users' bio fields from 2015 to 2020. Differences between the individual celebrity corpora in my data set seem to confirm this finding. Whereas only 5.37% of unique Twitter users in the Caitlyn Jenner corpus (comprising tweets written between late 2014-early 2016) list their pronouns in their bio, 35.8% of unique users in the Elliot Page corpus (comprising tweets written between mid 2020-late 2021) list their

pronouns. In fact, there is a increase in pronoun listing across historical time when comparing all target celebrity corpora (see Appendix B.1). Pronoun choice thus emerges as just one interfacing of language and gender that an individual can take advantage of in forming and performing their gender identity. Though cisnormativity might penalize individuals whose listed pronouns do not cohere with other components of their gender expression, an agentive perspective makes room for the fact that pronouns do not directly index gender identity and can change contextually and longitudinally.

The arrival of singular *they* into the English mainstream is in fact understood through two distinct changes – epicene *they* and nonbinary *they*, the latter of which is discussed in the final section of this chapter. Examples of the diverse uses of singular *they*, following analysis by Bjorkman (2017) and Conrod (2020) with slight modification, are laid out below.

1. **Generic, indefinite antecedent:** It's okay to tell someone that they cannot sing.

2. **Generic, definite antecedent:** The ideal singer has great breath control, even if they are also dancing.

3. **Specific, definite**[1] **antecedent:** The music director of my group is talented, but they often show up late to rehearsal.

4. **Specific, name antecedent:** Owen is great at opera, but they will never make it in theater.

5. **Specific, binary-gender antecedent:** I'll let my sister introduce themselves.

[1] In some situations, this may also be considered an ungendered or distal usage, as to signal that the referent is unimportant or uninteresting to the language user based on the implicature that if their gender were relevant, the language user would have selected a gendered pronoun (see Conrod, 2019, chap. 4)

Here, **epicene they** is the usage of singular *they* to refer to generic antecedents of all genders; when linguists discuss epicene pronouns, they are often talking about the form used when the language user does not know the gender of the referent or the gender of the referent is unspecified, as in (1) and (2) above. Meanwhile, **nonbinary they** describes the usage of singular *they* to refer to a specific individual as nonbinary, as in (4) and (5) above[2]. This is not to say that all nonbinary individuals list *they*, or that *they* can only be listed as a third-person pronoun by nonbinary individuals; rather, I use nonbinary *they* because this specific usage emerged from nonbinary communities (Conrod, 2019), is the most common pronoun listed by nonbinary people (Buch, 2017; Darwin, 2017; Hekanaho, 2020; Hord, 2016), and is listed by all of the nonbinary celebrities in this data set.

The emergence of epicene *they* to replace *he* is just one significant change that the pronominal system in English has already undergone throughout the history of the language (N. S. Baron, 1971). Indeed, social factors can explain how some pronouns expanded in scope while others were eliminated entirely. In an investigation of the second-person pronouns *thou* and *you*, Silverstein (1985) demonstrates how the collapse of the formality dimension between these pronouns can be explained by religious and political ideologies. Historically, *thou/you* differed grammatically along the dimension of formality, with the more formal *you* being used to index respect for the referent, consideration of the pragmatic speech context, or an elegant, classical style. In the lead-up to the 18th century, religious groups such as the Quakers found this formality distinction oppositional to their belief in civic equality for all beneath God. Accordingly, they began to collapse the *thou/you* distinction in favor of *thou*. Driven by distaste for the Quakers and these other deviant religious groups, the rest of English society settled on the *you* form instead. This finding contextualizes recent and ongoing

[2](3) could feasibly work as another example of nonbinary *they*, in the case that the music director uses *they*

changes to English's pronominal system as ideological in nature: individuals who espouse particular ideologies (Quakerism, gender-inclusivity, or cisnormativity, i.a.) can focus on a particular, structured area of language – such as the pronominal system – and mutate it into a vehicle of said ideology, charging all future usages with ideological valence (Silverstein, 1985). Conrod (2019) further discusses how, like the formality dimension still operationalized in many of the world's languages, conceptual gender is another such dimension along which pronouns can vary.

Another broad change at the intersection of social movements and pronouns began centuries ago: the rise of epicene *they*. While many cite the 1970s feminist critique of language – the backlash against epicene *he*, specifically – as the starting point of epicene *they*, it has in fact been in use since at least the 1400s (Balhorn, 2004; Curzan, 2003). Many feminist scholars, most notably Bodine (1975), argued that the use of *he* as an epicene pronoun was attributable to the misogynistic cultures in which English was spoken rather than as an accident of historical language change alone. For example, responding 'what is his specialty?' to the statement 'I love my doctor' with no premeditated knowledge of the doctor reflects biases regarding appropriate roles for specific genders in society. While some proposed *he or she* or *(s)he* as less sexist options for the epicene pronoun in English, these innovations now emerge, viewed under the light of cisnormativity, as more gender-exclusive than they gender-inclusive; by depending on the masculine-feminine binary encoded in these pronouns, they limit the expression of certain identities – especially nonbinary gender – in discourse. Nonetheless these activists successfully initiated a broad change in English's pronominal system challenging gendered power structures that still exist today through political work (D. E. Baron, 1981) and the lobbying of prescriptive style manuals (Schaefer, 2013). In the time since, however, epicene *they* (comprising usages (1) and (2) in the above list) has

emerged as the most common form. While psycholinguistic research conducted at the turn of the century found longer reading times for epicene *they* than for epicene *he* among university undergraduates, (Foertsch & Gernsbacher, 1997; Sanford & Filik, 2007), more recent work has demonstrated that this processing pattern regarding epicene *they* has reversed: *they* is now the most frequently used epicene pronoun (Hekanaho, 2020; Lascotte, 2016; Noll et al., 2018).

However, epicene *they* is not without its critics, outrage, and discourse. Bodine (1975), remarkably, dates the first prescriptive argument against singular *they* to the 1700s. Despite this, over 600 years of usage and 300 years of prescriptivism paled in comparison to a few decades' worth of targeted activism and rapidly shifting social tides in the case of epicene *they*. The broad uptake of epicene *they* demonstrates that the ongoing change in English's pronominal system is not new. Instead, it is simply the most recent in a long line of shifts motivated by social and ideological change (Bodine, 1975; Curzan, 2003; Silverstein, 1985).

3.2 Pronouns enmeshed: An ideological perspective

Over the last decade, third-person pronouns have caused quite the stir in American society. Though much of this change has trended towards inclusivity and expansion – Facebook's move to implement a pronoun field (Zimman, 2015) and the inclusion of personal pronouns in introductions, email handles, and social media profiles, among other sites – some commentators have folded these novel language practices in with other critiques of liberalism (Zimman, 2017). How and why does a culture imbue pronouns with such ideological gravitas? How, for example, did the discussion surrounding nonbinary *they* and neopronouns rise to such a level that enabled Fox news anchor Tucker Carlson to state that such pronoun usages make the English language 'dumber, less precise and embarrassing' (Ring, 2018)? Metalinguistic

comments such as this demonstrate how political, social, and linguistic attitudes towards nonbinary *they* have become enmeshed in a much larger ideology: cisnormativity.

Third-person pronouns are heavily implicated in personal reference and identity construction. On one hand, we encode our own conception of a person's gender into the pronominal form we select when referring to them in everyday discourse. This is an incredibly frequent act across many distinct natural language settings – in my own data, I observe that pronouns are used in just over 20% of tweets that also contain the names of celebrities. On the other hand, pronouns are one dimension of gender identity that, like all others, have the potential to shift at both the societal and individual level in response to changes in experience. Any discussion of pronouns 'entering the mainstream' is more accurately an observation that gender-affirming language practices involving pronouns that have been circulating in trans communities for decades (Zimman, 2014) are now receiving adoption in wider circles as matters of linguistic and social justice (Zimman, 2017). These attitudes towards changes in singular *they* have become inextricably linked with attitudes towards the trans and nonbinary communities in which these changes originated (Konnelly & Cowper, 2020).

More specifically, the listing of pronouns among trans populations as part of gender expression challenges many tenets of cisnormativity. This practice affirms that pronouns can change over a lifetime or context to context as a matter of safety or gender fluidity (Conrod, 2020). Like other forms of expression used in the performance of gender, cisnormativity edicts that these variable expressions must align in straightforward and predictable ways. However, the pronouns people choose to list may not cohere in ways intelligible to those holding cisnormative concepts of gender, either because these pronouns do not follow sociocultural gender norms or because they are packaged together with additional pronoun suites.

Related language practices, such as the normalization of listing and asking for self-determined third-person pronoun suite, further cements the trans notion of gender self-determination in mainstream society (Koyama, 2003). In this way, novel usages and practices surrounding third-person pronouns directly subvert cisnormative ideas, challenging the gender ideologies of those who believe sex and gender can and should be accurately inferred through anatomical, aesthetic, and behavioral cues (Zimman, 2017).

As stated earlier, attitudes towards gender-inclusive pronoun practices can be thought of in their relation to two variables: prescriptivism and conservative gender notions (Bonnin & Coronel, 2021). However, it is important to note that these dimensions are interrelated: often, the former serves as a guise for the latter. Hernandez (2020) found that participants' grammaticality judgements towards epicene *they* were significantly correlated with both prescriptive attitudes and attitudes towards transgender people, as participants with negative attitudes reported lower grammaticality scores. Meanwhile, attitudes towards transgender individuals were a much stronger predictor of ungrammatical judgements than prescriptive grammar attitudes regarding nonbinary *they* specifically.

This finding is emblematic of more large scale observations on the richly social nature of pronoun-related language practices. In 2019, a majority of Americans said that they had heard about the use of nonbinary *they* and/or neopronouns while around one in five personally knew someone who uses such pronouns (Geiger & Graf, 2019). Younger participants were more likely to fall into both camps than older participants. Furthermore, Democrats were more likely than Republicans to express comfort in using one of these gender-inclusive pronoun suites, though the overall metric was about 52%. As Borba (2019) demonstrates in Brazil, the adoption of gender-inclusive language practices can quickly become politicized and elicit reactions of fear, outrage, and misunderstanding because they target cisnormative

assumptions about the gender binary.

Accordingly, widespread discussion of gender-inclusive language reforms has taken place on social media, in the news media, and at many educational institutions. Expanding on this observation, Conrod (2018a) identifies the implementation and uptake of nonbinary *they* as a sociolingusitic change from above, whereby language users are aware of the innovation and metalinguistically refer to its implementation (Labov, 1966). However, nonbinary *they* did not manifest itself in a variable that carries high prestige nor did it originate in a high-status group in American society at large (Konnelly & Cowper, 2020). The discourse surrounding the innovation ranges from comments that deride its existence (as we saw with Tucker Carlson) to those that derive immense personal value from *they* due to its importance in nonbinary identity formation (as evidenced by comments made by Sam Smith in Chapter 6) (Darwin, 2017). More likely, nonbinary *they* and gender-affirming pronoun usage carry high prestige in specific communities of practice, such as LGBTQ+ communities or groups of younger language users (Conrod, 2018a). Within these communities – contemporary academia, for example – the failure to affirm someone's gender through the use of pronouns is often met with assertive correction.

Meanwhile, the possibility of gender-inclusive language reform (nonbinary *they*, chairperson, etc.) and practices (listing pronouns, using gender-affirming names, etc.) is perceived as unacceptable – or found injurious, even – to individuals who hold cisnormative ideologies. Because such reforms transgress the gender binary, people who strongly maintain cisnormative ideological principles assert that these reforms must be rejected, ignored, or erased. Changes in usage and practice surrounding third-person pronouns are just one such gender-inclusive reform. Any change to an aspect of language that encodes gender will be treated as an attack against the status quo by those who invest the most in the maintenance of

gendered power structures.

3.3 The great pronoun shift: Changes in singular *they*

Having situated the current change in the pronominal system of English within previous socially motivated shifts and demonstrated how innovative usages of third-person pronouns challenge the ideological principles of cisnormativity, I now characterize the shift in more detail through a review of relevant psycholinguistic and sociolinguistic research.

Rapidly changing ideas, experiences, and notions surrounding the gender binary in American society are pushing us toward a new organization of third-person pronouns (Konnelly & Cowper, 2020). As demonstrated by the introduction of Swedish *hen*, pronoun shift can occur rapidly (Sendén et al., 2015). Scholars of language and gender are uniquely equipped to track these changes in linguistic systems and especially so in those quite rigidly organized like the pronominal; in the process, we can learn about how highly saturated metalinguistic discourse and ideologies (Conrod, 2018a) might inform more discrete linguistic patterns.

The distinct usages of singular *they* discussed earlier in the chapter are here reprinted for ease:

1. **Generic, indefinite antecedent:** It's okay to tell someone that they cannot sing.
2. **Generic, definite antecedent:** The ideal singer has great breath control, even if they are also dancing.
3. **Specific, definite antecedent:** The music director of my group is talented, but they often show up late to rehearsal.
4. **Specific, proper name antecedent:** Owen is great at theater, but they would never make it in opera.

5. **Specific, binary-gendered antecedent:** I'll let my sister introduce themselves.

Contemporary research into usages (3) and (4) has stimulated a reconsideration of the ways gender information can factor into sentence processing and generated novel analyses of the distribution of pronouns in English. Singular *they* has been found to be more acceptable with a specific, definite antecedent than with a proper name antecedent overall (Bradley et al., 2019; Conrod, 2019). Doherty and Conklin (2017) manipulated the conceptual gender expectancy of specific antecedents (low–cyclist, high–mechanic, known–spokeswoman, for example) with singular pronouns (*him, her, them*). Here, conceptual gender expectancy refers to the degree to which an utterance exhibits (mis)alignment between a pronominal form and its referent for a given language user; thus, 'mechanic' elicits strong conceptual gender expectancy effects because it is a highly gender-normative profession in American society. The authors found that that naturalness ratings for *they* decreased as conceptual gender expectancy increased. The authors take this finding as evidence that experience with language – in this case, the type and token frequency with which an antecedent is referred to by a gendered pronoun – can be rapidly integrated during language processing. More specifically, Ackerman (2019) argues that this information enters processing as part of an exemplar tier, consisting of observations drawn from individual exposure to diversity in gender expression.

In an exploratory study expanding on previous work exploring epicene *they*, Bjorkman (2017) theoretically accounts for this evidence by arguing for a new distribution of singular *they*. Relying on the judgements of about twenty participants, Bjorkman argues that there is a growing class of language users for whom (3) is completely acceptable: innovative *they* users. For these language users, the gender feature on the subject 'music director' in (3) is optional; that is, when no conceptual gender information is introduced into processing by an

antecedent, innovative users find that *they* can refer back to it with no loss in acceptability. However, proper names and binary-gendered nouns in English either carry indirect or direct gender information in the form of distributional frequency or grammatical gender, respectively (Conrod, 2017). Bjorkman (2017) contends that, even for innovative users, *they* will always fail to match the features of any specific referent of known gender, as in cases (4-5). For nonbinary individuals who list *they*, then, language users must unlearn the generalization that certain proper names are uniformly specified with gender features.

Konnelly and Cowper (2020) locate these innovative *they* users as being intermediates in a three-stage process that more accurately characterizes the ongoing change:

Stage 1 All three gender features [masc], [fem], and [inanimate] are contrastive and obligatorily specified for nouns and proper names.

Stage 2 All three gender features remain contrastive but specification of these features is no longer obligatory, thus changing the distribution of nouns and proper names that carry a contrastive gender feature.

Stage 3 Contrastive gender features become optional modifiers. Thus, *they* becomes the default for singular, animate referents.

They report on the judgements of eight language users, some of whom are nonbinary and others who have at least one nonbinary family member. The participants in this sample are behaviorally representative of an even more innovate usage of *they* – Stage 3 – as they judge usages (4-5) as completely acceptable. This starkly differs from language users in Stage 1, for whom all three gender features [masc], [fem], and [inanimate] are contrastive and for whom proper names and binary-gendered nouns must bear one of these inserted gender

features[3]. For Stage 2 language users, who can accommodate usages of singular *they* like (3), gender features remain constrastive; what has changed, however, is that the insertion of these features becomes no longer obligatory. Nouns like 'music director' do not need to be assigned gender features even if the gender of the referent is known or conceived.

Returning to the evidence presented by Doherty and Conklin (2017) and Ackerman (2019), individual Stage 2 language users will vary in the proper names and nouns that lexically bear contrastive gender features according to their world experience. For a Stage 2 language user to accept nonbinary *they*, as in (4), they can either delete the contrastive feature for the name (*Owen* joins the class of nouns that are not obligatorily specified for gender) or they can add a new lexical entry for the name that has no marked gender feature (the new entry for *Owen* makes (4) grammatical, but as an exceptional use-case). In the case of a transgender person who begins to use a new, gender-affirming proper name and pronoun suite, language users at Stage 2 may similarly create a novel lexical entry to refer to them, rewriring the form-meaning connection and erasing the previous form. There is another option – misgendering – in which a language user consciously rejects or habitually avoids the antecedent's affirming pronoun suite, though this practice is more fully explored in the next chapter.

Contrastive gender features instead become optional modifiers for language users at Stage 3, enabling them to use *they* to refer to any non-inanimate singular antecedent regardless of morphological gender marking, distributional name frequency or conceptual gender expectancy. The reconfiguration finalized by Stage 3 is, in many ways, a grammatical manifestation of the understanding that gender is not a binary property. Though Konnelly and Cowper (2020) posit that most language users remain at Stage 2, the growing number and

[3]For complete discussion of a model of gender feature insertion that theoretically aligns with this paper, see Ackerman (2019).

visibility of individuals who use and/or list nonbinary *they* provides impetus for these language users to simplify their grammar. As previously argued in this work, Konnelly and Cowper (2020) provide more anecdotal evidence[4] that social factors – in this case, trans identity or experience with nonbinary *they* users – play a role in motivating language users along the three stages of this grammatical change.

Other studies have provided empirical evidence towards identifying the social factors that might be mediating this change and for characterizing it into three distinct stages. Sexist attitudes and negative attitudes towards feminist language reform (epicene *they*, for example) predict lower acceptance of nonbinary *they* through explicit survey items (Bradley, 2020; Hekanaho, 2020). Relatedly, cis women have been found to accept nonbinary *they* and have more positive attitudes towards gender-inclusive language than cis men (Hekanaho, 2020). Linguistic prescriptivism and 'grammarian' ideologies lead to higher resistance against epicene and nonbinary *they* (Bradley, 2020; Hernandez, 2020). An effect of age has also been observed, where younger language users are more likely to accept the innovation than older language users (Camilliere et al., 2021; Conrod, 2019; Hekanaho, 2020). Similarly, as suggested by the small participant sample in Konnelly and Cowper (2020), transgender (including nonbinary) individuals and those with greater experience with or awareness of transgender people[5] are more likely to accept nonbinary *they* (Ackerman, 2018; Camilliere et al., 2021; Conrod, 2019; Hekanaho, 2020). Similar patterns of age and transgender experience have been observed in studies of comparing *themselves* and *themself* (Ackerman, Riches, & Wallenberg, 2018; Davenport, 2020). Finally, Camilliere et al. (2021) offer empirical

[4]The authors themselves note that their participant sample does not comprise 'a comprehensive cross-section of the population who share the grammar we describe... these additional speakers augment our own judgements, accord with Conrod's (2018) results, and show that the system we present is not exclusively our own'.

[5]Some differences have been observed between experience levels with nonbinary, as opposed to binary, transgender people (see Camilliere et al., 2021).

evidence that language users align with one of three stages in the changing acceptability of nonbinary *they* in ways partially predictable by social factors. In an analysis of results from an acceptability task, the authors statistically determine three clusters of participants that mirror the three-stage distribution predicted by Konnelly and Cowper (2020). Participants in the Stage 3 cluster tended to be younger and also included all of the transgender participants in the study.

Furthermore, most of the acceptability and usage patterns observed with Swedish *hen* have now been replicated for nonbinary *they*. As in Sendén et al. (2015), liberals and cis women are more likely to accept nonbinary *they* than conservatives and cis men (Hekanaho, 2020). Turning towards usage, we find also that younger language users are more likely to use nonbinary *they* (Conrod, 2019; Sheydaei, 2021). Most relevant to the present study is the repeated observation that actual usage trails behind the positive attitudes surrounding usage. That is, the acceptability rate of nonbinary *they* at a metalinguistic level is often higher than its use as a functional element (Conrod, 2019; Sheydaei, 2021). Similarly, in a corpus analysis of news articles comparing those written about individuals who list binary pronouns and those about nonbinary individuals who list nonbinary *they*, Arnold, Marquez, Li, and Franck (2022) find that, controlling for discourse context, usage of *they* is significantly lower than binary pronouns *she* and *he*. This suggests that, like the timecourse of *hen*, the relative unfamiliarity of nonbinary *they* and nonbinary expressions of gender may inhibit its usage in production (Arnold et al., 2022).

Collectively, this research indicates that trans identity, as well as attitudes towards and experience with those who hold such identities, is a socially meaningful predictor of syntactic change in this arena. We also see rapid change in the gender expectancy, or the conceptual gender alignment between referent and pronominal form, of nonbinary *they* over

a roughly five year period (Camilliere et al., 2021; Doherty & Conklin, 2017), and that between-partipant differences are largely predictable by social considerations, such as age, gender, and ideology (both in relation to linguistic prescriptivism and cisnormativity). It is clear that gender-inclusive language practices involving pronouns challenge the ideological principles of cisnormativity at work in societies, institutions, and individuals. Furthermore, this research has begun to characterize this change into three stages driven by social patterns, mirroring previous shifts in English's pronominal system.

CHAPTER 4
HARMFUL SPEECH: MISGENDERING AND DEADNAMING

In opposition to gender-inclusive language reform, pronominal misgendering and deadnaming operate to enforce cisnormativity. Through these language practices, the agency of transgender individuals to self-determine their own gender is denied as their identities and expressions of gender are forced to align with an immutable gender binary. In this chapter, I discuss how the relationship between names, pronouns, and gender can be manipulated through harmful language practices that reject trans identities. I first discuss how cisnormativity is in part perpetuated through the regulation and erasure of transgender identities. I then discuss the two specific language practices analyzed in the present study: misgendering and deadnaming. Misgendering, here, is the use of third-person pronouns to refer to an antecedent in a way that does not align with their biosocial gender and/or gender identity. Deadnaming is a more contextualized language practice, whereby language users select the former name of transgender people rather than their correct, gender-affirming name. Ultimately, I argue that gathering large-scale data on these practices is one promising direction towards dismantling cisnormativity, which may be inhibiting discrete patterns of uptake in nonbinary *they* and novel pronoun listing more generally.

4.1 The regulation of transgender identities

For cisnormativity to be a totalizing and complete ideology, it must find a way to deal with transgressive practices and people (Gal & Irvine, 2019). Cisnormative agents – whether individuals or institutions – must police the boundary between man and woman against those

who in their perspective might seek to cross, subvert, or erase it. In the case of binary transgender individuals, it it the sense of 'crossing' that is problematized under cisnormativity. In the case of nonbinary transgender individuals, it is the erasure of the gender binary that cuts right at the core cisnormative notion that all individuals can be classified as either 'man' or 'woman'.

The most intense articulations of cisnormativity are often espoused by those who have the most to gain from the maintenance of a rigid gender binary. We see this manifest in language across multiple social hierarchies: more misogynistic, cisnormative, conservative, and grammarian beliefs all predicted lower acceptability of gender-inclusive language reforms (Bradley, 2020; Hekanaho, 2020; Hernandez, 2020; Sendén et al., 2015). Indeed, transgender people and gender-inclusive language practices directly contrast and contest the cisnormative assumptions outlined in Chapter 2. Thus, in the maintenance of a natural, stable, and coherent gender binary, trans identities must either be ignored or explained away. But the transgender moment is here (Zimman, 2020), and with it, an increase in visibility. As agents of cisnormativity explicitly work to erase trans expressions of gender, they must regulate them; they must make visible expressions of transness adhere to their concepts of gender and the man-woman binary.

Because neither cisnormativity nor the gender binary is natural, they both must be actively produced. This regulation is in part completed through two specific language practices: misgendering and deadnaming. Some research has defined misgendering as a discursive practice that contains under its umbrella both deadnaming – the use of a trans person's former name (often, one that was distributionally gender-normative and assigned to them at birth) – and 'mispronouning', or the use of third-person pronouns in an act of reference that does not correspond to the asserted gender expression of the referent (Conrod, 2017; Hekanaho,

2020; Zimman, 2019). In addition to deadnaming and mispronouning, there are of course many other linguistic practices that misgender trans individuals, including the use of lexical items with overt grammatical gender ('woman','sister','mother', etc.) (Zimman, 2019), through medical classifications (Borba & Milani, 2017), and the obligatory specification (and often conflation) of binary gender and/or sex in legal proceedings, on licenses and passports, and in some legislative contexts (Cannoot & Decoster, 2020). For analytical clarity, I take misgendering to refer specifically to the use of third-person pronouns in instances that do not align with those listed by a referent as part of their gender expression. Like other language practices, misgendering can be classified along axes of context and intent. Adapting Simpson and Dewaele (2019) for the more specific contexts discussed in this paper, I classify three types of misgendering, ordered from most to least intentional:

- **Intentional misgendering**, whereby language users – fully aware of the self-determined gender identity/expression of their referent – misgender them in order to deny their agency and gender validity, harass them, and/or to explicitly reinforce the immutability of gender/sex under cisnormativity. While the language user's relationship to gender-inclusive ideological principles may vary in understanding or acceptance, the act of misgendering is intentionally purposed to deny the gender identity of the referent.

- **Conceptual misgendering**, whereby language users unaware in advance of the gender identity of their referent produce a mismatch between the language user's conceptual gender they assign to the referent and the referent's asserted gender expression, resulting in misgendering. The language user's conceptual gender forms in response to material/physical indicators (appearance, speech, dress, etc.) or in accordance with other interlocutors.

- **Habitual misgendering**, whereby language users are aware of the gender identity of their referent, have no intention to reject their right to gender self-determination, but nonetheless fail to pronominally affirm the referent. The language user's intention to gender-affirm is blocked by either historical alignment, whereby the referent recently changed their pronouns and the language user relied on the old lexical entry (Ackerman, 2019), or perceptual alignment, whereby sociocultural gender norms relating to speech or physical appearance more strongly activated a certain conceptual gender toward the referent in the mind of the language user.

While this taxonomy clarifies how intentionality and contextual information can play a part in misgendering, it is important to recognize that all instances of misgendering and deadnaming can harmfully impact transgender individuals. Of course, cisgender people can also be misgendered, but in doing so their right to gender-self determination is not denied in the same way it often is for trans individuals. Furthermore, unlike nicknaming, deadnaming becomes harmful when it is directed towards a person who might harbor dissonance between a former name and their biosocial experience of gender (Turton, 2021). Whereas intentional misgendering of cisgender people can arise from individual expression of gender that does not properly accord to a culture's gender norms, intentional misgendering of transgender people often results from the belief that trans people do not/should not exist. The cisgender case is offensive while the transgender case is dehumanizing (Haslam, 2006). Transgender people, especially those of color, are already one of our most at-risk health populations (Geiger & Graf, 2019). Transgender people are significantly more likely than their cisgender peers to suffer from mental health issues such as depression and suicidal ideation (Olson, Durwood, Demeules, & McLaughlin, 2016). Empirical studies in the health sciences have found that misgendering can negatively affect the mental health of binary and nonbinary trans people,

particularly through the degradation of their self-conception and the accumulation of social stigma (Johnson, Auerswald, LeBlanc, & Bockting, 2019; McLemore, 2015). As observed more qualitatively by Borba and Milani (2017), S. E. Valentine and Shipherd (2018) find that gender-affirming language and behavior by healthcare providers can mitigate some of the negative mental health symptoms experienced by trans individuals. Olson et al. (2016) find that trans youth who are supported in their gender identity at home and school experience developmentally normal levels of depression and anxiety. Similarly, the usage of trans youth's gender-affirming names across multiple contexts is associated with lower depression and suicidal ideation (Russell, Pollitt, Li, & Grossman, 2018). These studies highlight the medical harm cisnormative language practices can have on the mental health of transgender individuals, and, correspondingly, the critical importance of promoting gender-affirming language practices.

Having illustrated a broader picture of cisnormativity's relationship to the regulation of trans identities through misgendering and deadnaming, I turn now to reviewing previous literature on these two language practices, which are rather unexplored in the literature. The studies that do exist – particularly, work by Kirby Conrod – clearly demonstrate the anti-trans ideological work these practices are used to perform. In a production study, negative implicit attitudes among language users – more closely related to conceptual than intentional misgendering – were shown to predict higher rates of misgendering regarding a trans film character (Conrod, 2018b). Additionally, only the trans characters in the film were misgendered. In Conrod (2017), they demonstrate through a small case study that Twitter users more readily used Chelsea Manning's gender-affirming name than pronouns. Together, this suggests that that pronominal misgendering may go unnoticed more often and may be more related to unconscious attitudes than for deadnaming, which tends to be more explicit

in its rejection of trans identity (Turton, 2021).

Furthermore, deadnaming and misgendering are often undergirded by negative sentiments. For trans individuals, a self-determined novel name is not only an assertion of their gender identity but a means of bridging their biosocial experience of gender with their gender expression by taking advantage of the gender-normative distribution of names in a given language (Sinclair-Palm, 2017). Generally, the use of a nominal in a referring expression reflects the language user's assumption – or assertion, as deadnaming critically illustrates – that the referent is indeed categorized or characterized by that form (McConnell-Ginet, 2003). As such, the language user's selection of a former, non-gender affirming name that is often the one assigned to their transgender referent in infancy reintroduces feelings of dissonance and gender invalidation. Turton (2021) explores Urban Dictionary entries of Caitlyn Jenner following her coming-out event, finding that intentional deadnaming indeed functions to invalidate Caitlyn's identity as a trans woman and to maintain cisgender hegemony. Commentators focus on the possibility of Caitlyn's surgical procedures – proposing that the status of 'woman' must be accomplished aesthetically, not innately – and use biological sex terms to distance her body away from a natural, coherent cisnormative ideal. Turton (2021) illustrates how deadnaming is not merely a use of something dead; rather, deadnaming is a purposeful act in which an transgender individual's gender history is reanimated to introduce dissonance between their affirming gender expression and a past expression that is no longer relevant. In this way, deadnaming denies trans individuals the agency to affirm their gender and reinforces the borders of cisnormativity. A similar finding is observed in Conrod (2017), as tweets that both misgender and deadname Chelsea Manning have a more negative sentiment than pronominally gender-affirming tweets.

Together, this chapter demonstrates how misgendering and deadnaming reflect cisnorma-

tive principles, negatively impact the mental health of trans individuals, and serve to deny individuals their agentive right to gender self-determination.

CHAPTER 5
A COMPUTATIONAL SOCIOLINGUISTIC APPROACH TO GENDER

In this chapter, I situate this project within the burgeoning field of computational sociolinguistics. In doing so, I describe the unique features of and approaches to computer-mediated language variation and change. More specifically, I re-introduce Twitter as a site for computational sociolinguistic analysis and describe what it can offer scholars of language, gender, and sexuality. Finally, I introduce the present study.

5.1 Computer-mediated language variation and change

Scholars have recently begun utilizing computational linguistic methods to understand social phenomena and societal issues. An increasing reliance on social media platforms for human interaction has led to the availability of large corpora of computer-mediated communication. These data allow us to study the content of social media platforms as data, exploring how users, like other producers of speech and text, manipulate language to perform ideological work, establish their identities, and interact with one another. Of particular interest to sociolinguists, social media corpora enable large-scale investigations of dialectal variation (Blodgett, Green, & O'Connor, 2016; Doyle, 2014) and language change across social networks (Eisenstein, O'Connor, Smith, & Xing, 2014; Goel et al., 2016). Additionally, demographic information (whether inferred, known in advance, or listed on the user's profile) enable more micro-level studies of stylistic work (Ilbury, 2020) or interaction in discourse (Zappavigna, 2012).

Many of these studies – representative of the relatively new field of computational soci-

olingustics (Nguyen et al., 2016) – make use of Twitter, a microblogging platform. While linguistic data from Twitter is more readily acquirable than from other social media platforms, which provide little data access to those pursuing academic and industry projects, such data also makes processing and sociolinguistic analysis more difficult. For one, Twitter data is extremely noisy: it can contain emojis, abbreviations, acronyms, links (to gifs, images, other tweets, etc.), and other users (through the use of Twitter handles, such as @ddlovato[1]). Because of this, traditional computational linguistic tools, such as part-of-speech taggers, coreference resolution, and named entity recognizers, often do not perform as well. It is thus difficult to cleanly capture and analyze many of the dimensions of language when relying on Twitter data.

The social dimension of language is particularly tricky, as the demographic information that helps drive sociolinguistic analysis is often unknown or inaccessible. Twitter makes available multiple open text fields for identification purposes, including the username, name, short biography, and location fields. However, as observed in my own data, Twitter users do not use these fields in uniform ways or in ways that facilitate easy processing on a large scale. Many computational studies have thus turned towards inferring demographic features through different means. For some demographics, like political affiliation, researchers rely on homophily in Twitter users' following habits (Demszky et al., 2019). For others, such as age, gender, and regional origin, researchers have either exploited census data detailing the gender-normative distribution of proper names (Prabhakaran, Reid, & Rambow, 2014) or trained models using pre-annotated data sets (Nguyen, Gravel, Trieschnigg, & Meder, 2013; Rao, Yarowsky, Shreevats, & Gupta, 2010).

In fact, using linguistic data for automated gender detection has become part of the

[1]This is the Twitter handle of Demi Lovato, one of the celebrities in this paper's data set.

so-called 'author profiling task' in natural language processing, in which researchers construct increasingly complex models to predict demographic qualities of the author through their text alone (Estival, Gaustad, Pham, Radford, & Hutchinson, 2007; Rangel & Rosso, 2013). These methods may appear necessary to conduct large-scale analysis of sociolinguistic variation and change, but they are fundamentally opposed to the precept of gender self-determination. While those ascribing gender to users, speakers, and/or referents in this manner may not intend to misgender, it can clearly lead to repeated instances of conceptual misgendering. The decision to train models on data classified using the gender binary ultimately constrains the set of possible conclusions to those that reify the same binary and the cisnormative assumptions supporting it. In this way, a 'hall of mirrors' effect emerges (Bamman, Eisenstein, & Schnoebelen, 2014), whereby large scale studies that infer gender identity fail to further our understanding of how gender is performed, expressed, or affirmed; instead, they merely reflect, reproduce, and further reify those sociocultural gender norms that are significant enough to emerge through computational methods. Both pursuits are worthwhile, but researchers have often conflated findings that more closely adhere to the latter in pursuit of the former. In a sweeping review of the treatment of gender as a variable in computational linguistic work, Larson (2017) demonstrates that studies inferring gender almost always rely on a binaristic classification of gender and frequently omit any sort of statement detailing the theory or framework of gender that their work rests upon. Though the great scale and scope offered by computational linguistic methods represents a promising future direction for scholars pursing the social dimension of language, we must interrogate our methods for the same gender biases that we oppose in other arenas. This involves finding ways to further gender self-identification as the gold standard in computational work and explicitly stating working theories of gender in any research that intends to study it.

There are many recent improvements in this regard. Cao and III (2021) illustrate that current systems of coreference resolution almost always utilize gender-binaristic data and gender inference. They subsequently introduce a novel data set with balanced gender-inclusive coreference relationships – including nonbinary *they* – with which to train coreference resolution models. Further, Bamman et al. (2014) demonstrates that it is possible to use linguistic patterns to infer and analyze the *performance* of gender identity without ever explicitly conflating the conceived gender a model assigns to a user and that user's biosocial gender in reality. To avoid inferring entirely, another possibility is to rely on a limited set of Twitter users with known gender identities (Ilbury, 2020), such as celebrities, or to cluster users according to the pronouns listed on their account and analyze how this aspect of gender expression might pattern behaviorally. Unlike Facebook (Zimman, 2015) and Instagram (Lewis, 2021), however, Twitter has not yet implemented a pronoun field, forcing users to list their pronouns in their biography or location field. An additional option is to investigate not how Twitter users express gender, but the extent to which they ratify the gender of other users whose gender identities are known in advance.

5.2 Research questions

The present study joins a recent line of work in computational sociolinguistics exploring dehumanization and social bias (Breitfeller, Ahn, Jurgens, & Tsvetkov, 2019; Mendelsohn et al., 2020; Voigt et al., 2017), turning specifically to the mechanisms underlying Twitter users' (dead)naming and pronominal (mis)gendering of transgender celebrities. These mechanisms can be understood both distributionally and lexically. In the distributional analysis, I seek to measure the prevalence of misgendering and deadnaming in the period before and after these celebrities' coming-out events (COEs), similarly to the Twitter corpus analysis performed

by Conrod (2017). More specifically, I want to understand how transgender celebrities' pronouns are affirmed relative to those of cisgender and transgender celebrities without publicly documented coming-out events.

RQ1: At what rate do Twitter users adopt transgender celebrities' gender-affirming names and pronouns following publicly documented coming-out events? How does these rates compare to cisgender celebrities?

I also examine the potential differences in affirming pronoun usage surrounding trans celebrities who use the binary pronouns *she* and *he* (trans-binary) and nonbinary transgender celebrities who specifically use nonbinary *they* (trans-nonbinary). This effort complements studies in psycholinguistics and sociolinguistics that have sought to document ongoing changes in nonbinary *they* (Ackerman, 2019; Arnold et al., 2022; Bjorkman, 2017; Camilliere et al., 2021; Conrod, 2020; Konnelly & Cowper, 2020).

RQ2: Do Twitter users adopt gender-affirming binary pronouns (*he* or *she*) at the same rate as they adopt gender-affirming nonbinary *they*?

Any distributional findings in this regard might be in part explained by cisnormative ideologies that inhibit users' adoption of gender-affirming pronouns and names. In the content analysis, I thus attempt to document the rhetorical fingerprints of cisnormativity surrounding deadnaming and misgendering.

RQ3: Does the lexical content that occurs in misgendering/deadnaming tweets appear driven by cisnormative ideologies? To what degree are misgendering/deadnaming tweets associated with negative sentiment? What lexical items are highly correlated with these practices?

Together, I aim to address these questions by utilizing large-scale linguistic evidence to interrogate ongoing practices of deadnaming, misgendering, and the cisnormative regulation

of trans identities.

CHAPTER 6
DATA

In this chapter, I introduce the seven celebrities that constitute my target and comparison data sets and explore some facets of their public gender expressions.

6.1 Seven celebrity corpora

Zimman (2020) declared the 2010s as the decade of transgender publicity, throughout which trans notions, theories, and experiences of gender began to emerge into mainstream American academia and popular culture. The latter half of the decade might also be termed as a time of transgender celebrity[1], as many prominent trans celebrities affirmed their experiences of gender in publicly documented coming-out events (COEs). Here, COE refers to the exact moment of declaration by the trans celebrities in this data set that their inner, self-realized experience of gender matched their social, identity-oriented expression of that gender. By grounding my analysis in the self-expressed gender of these celebrities, I hope to explore not how they construct their own identities, but how these identities and expressions of gender are ratified by Twitter users on a large scale. It is thus important to understand what information and messages Twitter users might have received regarding the gender identities of the four trans celebrities in the 'target' group and the three celebrities in the 'comparison' group.

The target group consists of four transgender celebrities with Twitter accounts and publicly documented COEs. **Caitlyn Jenner** had the earliest COE of any celebrity in the target

[1]For a more intensive discussion on the nature of trans celebrity, see Zimman (2019).

group. Born in 1949 (age 72), Caitlyn Jenner is an American athlete, socialite, and, most recently, conservative politician. In April 2015, Jenner publicly came out as a trans woman in a *20/20* interview with Diane Sawyer. The change was particularly shocking to those who previously saw Jenner as the 'textbook embodiment of hegemonic masculinity' (Turton, 2021). Indeed, Jenner is a decorated Olympic athlete, a notable part of the Kardashian clan – arguably the most famous celebrity family of the 21st century – and has wedded multiple glamorous women. Her public COE reached its peak in June 2015, when she appeared on the cover of *Vanity Fair* in part of an equally glamorous photo shoot with photographer Annie Leibovitz. It was at this time that she presented her new name, Caitlyn, and elected to use the binary pronouns *she/her*. Jenner quickly transitioned from fulfilling sociocultural male gender norms (athletic prowess, wealth, reproductive success) to more closely identifying with the cultural norms pursued by her daughters (style, femininity, beauty, sophistication). Because Jenner's COE took place over two significant events, two dates were centered in analysis. Tweets were scraped from the 6 months before (10/24/2014) the *20/20* Diane Sawyer interview (4/24/2015) and, following the *Vanity Fair* cover (6/1/2015), an additional 9 months (3/01/2016). Because of this, the analysis period for Jenner is slightly longer than that of the other target celebrities. The following terms were used in the query to scrape tweets: Caitlyn Jenner, #CaitlynJenner, @Caitlyn_Jenner, Bruce Jenner, #BruceJenner. N=2,613,733 tweets were collected in total.

The second target celebrity is **Sam Smith**. Born in 1992 (age 29), Sam Smith is a Grammy-award winning English musical artist. While Sam Smith is not the first nonbinary celebrity, they were arguably the most famous at the time of their COE. Following identification earlier in their music career as a gay man and subsequently as genderqueer, Smith publicly came out as nonbinary in a series of tweets on September 13th, 2019 (Smith, 2019).

In these tweets, they say:

> "I've decided I am changing my pronouns to THEY/THEM... I understand that there will be many mistakes and mis gendering but all I ask is you please please try. I hope you can see me like I see myself now... I am at no stage just yet to eloquently speak at length about what it means to be non binary but I can't wait for the day that I am. So for now I just want to be VISIBLE and open."

Clearly, Smith's decision to list *they/them* pronouns is essential in defining and expressing their nonbinary identity, which they are still in the process of locating. They also recognize the possibility that, due to the relative unfamiliarity many people have with nonbinary *they*, there will be more than a few instances of misgendering. This series of tweets marks the center of their analysis period. Tweets were scraped from the 6 months before (3/13/2019) these tweets (9/13/2019) and the 9 months following them (6/13/2020). The following terms were used in the query to scrape tweets: Sam Smith, #SamSmith, @samsmith. N=601,835 tweets were collected in total.

The third target celebrity is **Elliot Page**. Born in 1987 (age 35), Elliot Page is an Oscar-nominated actor and producer. When formerly presenting as a gay female, Page was one of the most visible lesbian actors in Hollywood. On December 1st, 2020, he also used Twitter to come out as transgender – more specifically, as transmasculine nonbinary (Page, 2020). Page lists both *he* and *they* pronouns, but prefers *he* over *they* (GLAAD, 2021). In the tweet, Page says:

> "I want to share with you that I am trans, my pronouns are he/they and my name is Elliot... I can't begin to express how remarkable it feels to finally love who I am enough to pursue my authentic self. I've been endlessly inspired by so

> many in the trans community... I ask for patience. My joy is real, but it is also fragile... To all trans people who deal with harassment, self-loathing, abuse, and the threat of violence every day: I see you, I love you and I will do everything I can to change this world for the better."

Page suggests that his listed pronouns are heavily implicated in his trans identity and allow him to pursue his authentic self. And, similar to Smith, Page asks for understanding and respect as he embraces his trans identity. Furthermore, Page recognizes his positionality in the community as a trans celebrity and takes a notably activist stance in combating trans discrimination. Tweets were scraped from the 6 months before (6/1/2020) this tweet (12/1/2020) and the 9 months following it (9/1/2021). The following terms were used in the query to scrape tweets: Elliot Page, #ElliotPage, @TheElliotPage, Ellen Page, #EllenPage, @EllenPage. N=267,027 tweets were collected in total.

The final target celebrity is **Demi Lovato**. Born in 1992 (age 29), Demi Lovato is an American musical artist and actor. Lovato rose to fame on the Disney channel, where they appeared in *Camp Rock* and *Sonny with a Chance*. Shortly after, they began a prolific music career and were eventually nominated for a Grammy. Before publicly coming out as nonbinary, Lovato at different times identified as sexually fluid, queer, and pansexual. In a video and series of tweets uploaded to their Twitter account on May 19th, 2021, Lovato stated that they identify as nonbinary and officially listed their pronouns as *they/them* (Lovato, 2021). In the video, Lovato says:

> "...I've had the revelation that I identify as nonbinary. With that said, I'll officially be changing my pronouns to they/them. I feel that this best represents the fluidity I feel in my gender expression and allows me to feel most authentic and true to the person I both know I am and still discovering... I am excited

> to share with you what this means to me, and what it may look like for other people. I want to make it clear that I'm still learning and coming into myself, and I don't claim to be an expert or a spokesperson."

Like Smith, Lovato draws a clear connection between their nonbinary identity and their listing of *they* pronouns. They also note that they are excited to explore their nonbinary identity, as well as more general expressions of nonbinary identity, with their fan base, but that they are not the definitive voice on nonbinary or trans experiences. Tweets were scraped from the 6 months before (11/19/2020) these posts (5/19/2021) and from the 9 months following them. The following terms were used in the query to scrape tweets: Demi Lovato, #DemiLovato, @ddlovato, Themi Lovato[2]. N=1,188,029 tweets were collected in total.

Collectively, these four individuals are perhaps the most visible binary and nonbinary transgender celebrities in Western society today that also have publicly documented COEs. This latter fact enables studies of this data to compare how Twitter users' linguistic patterns evolve as transgender celebrities come into their identities, from the usage of pronouns and names to more general discussion. While these four celebrities were initially part of a much longer list, it became clear in initial processing that the total N of tweets for a given celebrity had to be very large to facilitate meaningful analysis. This is partly due to the filtering process, which will be illustrated more clearly in the next chapter. Thus, I retained only those celebrities with high enough celebrity[3] for final analysis, though it is certainly possible that there are other transgender celebrities for whom meaningful computational analyses

[2]In exploratory analysis, I observed tweets containing this nickname. I decided to include them in the larger scrape to gather as much data as possible, and I do not by any means endorse its usage.

[3]I take celebrity here a measure of gross N of tweets generated over a fixed period which, I believe, is as good a measure as any.

could have been conducted.

This same consideration in part contributed to the selection of celebrities in the 'comparison' group, which consists of one cisgender woman (Doja Cat), one cisgender man (Tom Holland), and one transgender woman (Laverne Cox) with Twitter accounts but no publicly documented coming-out events. While a transgender man would have complemented these three celebrities nicely, I was not able to identify such an individual with a sufficiently high celebrity and a Twitter account.

The first comparison celebrity is **Doja Cat**, or Amala Dlamini, who is a Grammy-winning American musical artist and producer. Born in 1995 (age 26), Doja shot to fame with her single 'Say So', which was shortly followed with the popular album *Planet Her*. Doja describes the album as 'giving divine feminine' and uses it to explore topics of femininity, romance, and sexuality (Curvy, 2021). Doja was selected to be part of the comparison group because of this; her celebrity is in part defined by her (cis)gender expression of womanhood. Tweets were scraped in a nonspecific, 6-month time period after Doja fully entered the mainstream, from (7/1/2021) to (1/1/2022). The following terms were used in the query to scrape tweets: Doja Cat,#DojaCat, @DojaCat. N=1,585,396 tweets were collected in total.

Tom Holland is the second celebrity in the comparison group. Born in 1996 (age 25), Holland is an English actor who achieved massive fame beginning in 2016 for his portrayal of Spider-Man. Since then, Holland has emerged as one of the most popular leading actors of his generation, and was selected for this reason to be part of the comparison group. Like other 'leading man' actors before him, Holland's celebrity persona is taken to be emblematic of modern (cisgender) expressions of manhood. Tweets were scraped in a nonspecific, 6-month time period after Holland fully entered the mainstream, from (7/1/2021) to (1/1/2022). The following terms were used in the query to scrape tweets: Tom Holland,#TomHolland,

@TomHolland1996. N=557,482 tweets were collected in total.

The final celebrity in the comparison group is **Laverne Cox**. Born in 1972 (age 49), Cox is an Emmy-winning American actress and LGBTQ+ advocate. She rose to fame for her recurring role in the series *Orange is the New Black* in the early 2010s, and was the first transgender person to be nominated in an acting category at the Emmy Awards. Besides Jenner, Cox is perhaps the most well-known transgender woman and is definitively the most famous transgender actress. Cox was selected to to be part of the comparison group for this reason; while at times her results are aggregated with the transgender target group, the fact that she has no publicly documented COE enables this project to consider the effects of time and the public nature of the target celebrities' gender expressions on the language practices under analysis. Tweets were scraped in a 2-year period following her *Time* magazine cover photo, from (8/1/2013) to (8/1/2015)[4]. The following terms were used in the query to scrape tweets: Laverne Cox,#Laverne Cox, @Lavernecox. N=252,725 tweets were collected in total.

6.2 Data collection

All tweets were scraped using the Twitter API v2 in Python using the above queries. The code for all scrapes is available on my Github. Besides retweets, all types of tweets (main, replies, quotes, etc.) were included in the scrape query. I modified sample Python code provided by Twitter[5] to perform scrapes for each celebrity, which ocurred in multiple rounds between December 2021 and March 2022. The academic API license used to perform these scrapes was provided by my thesis advisor, Professor Rob Voigt, on behalf of the Computational Linguistics Lab (CoLLab) at Northwestern University.

[4] A quick Google Trends searched revealed that this was the most prolific time of Laverne Cox's career, thus far.

[5] urlhttps://github.com/twitterdev/Twitter-API-v2-sample-code

CHAPTER 7
METHODS

In this chapter, I describe the steps I took to process my data. I first characterize my social positionality as a researcher, given that it undoubtedly bears on the methods and analyses that I elected to use. Then, I detail the filtering process and outline its results.

7.1 Statement of positionality

Gender is a fraught concept to work with. Particularly because I am dealing issues related to gender minorities – specifically, transgender individuals – and because I am not a member of this population, I find it important to state the relevant parts of my background, intentions, and guiding principles that I brought to this project.

I am a white, queer, and cisgender student. I developed the idea for this project when I first noticed the listing of pronouns as an act of gender identification and, further, observed some of the discourse that was taking place around pronouns online and in the news. I wondered what this practice meant to different people – how it could be critical and meaningful to some, yet transgressive and inflammatory to others. This initial idea led me down a long path of computational work, gender theory, and discussions with many scholars/friends of diverse genders.

In this work, I intend to expose partially the inner workings of an ideology – cisnormativity – that is perpetuated by people with whom I share a gender identity and, I'm sure at some points, by me as well. I do not intend, with this project, to speak for transgender experiences. To that end, I have relied extensively on research by trans scholars in and

outside of linguistics. I thank them for their inspiring work, for creating space in academia for a topic such as this, and for advancing the interests of the LGBTQ+ community.

7.2 Pre-processing

Tweets were scraped in JSON file format. I ended up with multiple JSON files for each celebrity as additional rounds of scraping became necessary as the analysis expanded. JSON files were combined and converted into CSVs using Python.

The following fields of information were collected during the scrapes: tweet id, raw text, date/time created at, number of likes, number of retweets, number of replies, number of quotes, author id, author username, author name, author bio (if available), author location (if available), number of followers, number of following, and place id (if available). I used spaCy (Honnibal & Montani, 2017) to tokenize each tweet and add all hashtags to a separate field. Then, for each token in the tweet, I checked if it was contained in a set of pronouns specified for each celebrity according to their listed pronoun suite. Throughout this process, I tracked total counts of *he*, *she*, and *they* usages.

I then used regular expressions to standardize the tweets across the seven celebrities. Target pronouns (*their* in a tweet discussing Demi Lovato, for example) were replaced with a standard TARGET_PRONOUN token, as were wrong pronouns (WRONG_PRONOUN), celebrity names and deadnames (CELEB_NAME, CELEB_DEADNAME), and celebrity Twitter handles (TWITTER_HANDLE, DEAD_HANDLE). Once standardized, I calculated total counts of these tokens for each tweet, allowing me to easily remove tweets that contained no pronouns or celebrity names later on in the filtering process. I then used these counts to calculate the percent of affirming pronoun and name usage contained in each tweet.

In addition to other tweet- and user-level measures offered through the initial API scrape,

I present a novel approach to incorporating aspects of gender expression information into computational linguistic studies. While self-identification will always be the gold standard, large-scale studies unable to directly access gender information have almost always relied on binaristic, arguably cisnormative methods to infer users' gender identities (Larson, 2017; Nguyen et al., 2016). True to the aim of the present study, I offer an alternative made possible by the increasing prevalence of identity information that Twitter users include in their profile bios (Jones, 2021). I designed a regex to detect any combination of pronouns, two neopronouns (*xe* and *ze*), or the phrase 'any/all pronouns' listed in these fields for each user in the data set. I will use the term 'user bio' to refer to both of these fields, as by-hand analysis revealed that users tend to utilize the location field as an additional space for identification more readily than they use it to provide information detailing their geographic location. While I do not map these pronouns directly onto gender identities, this information on aggregate provides insight into how diversely gendered people interact with Twitter – specifically, in this case, through the discussion of transgender celebrities. Furthermore, the mere presence of pronouns in a user's bio is a meaningful social determiner of their gender ideology. With this method specifically, I make use of the increasingly high rate at which users list their pronouns in their bio to ask whether the presence of listed pronouns is a predictor of misgendering/deadnaming tweets. I hope that future researchers take advantage of this approach to understand the distributional and behavioral patterns underlying third-person pronoun listing before adopting inferential gender assignment methods that do not ascribe to gender self-determination.

7.3 Filtering

I had to devise an extensive filtering process for this project because of two main reasons. First, as previously stated, Twitter data is incredibly noisy. I had to ensure that these tweets were not created by bots, which can produce hundreds of duplicate tweets, and that all tweets captured by the API actually referred to the celebrity in question. Although the search queries only used the names and twitter handles of these celebrities, the full archive search on the Twitter API uses these terms rather greedily. Second, I had to ensure that all pronouns and names contained in the tweet did actually refer to the celebrity mentioned under analysis. This is the tricky task of coreference resolution, which has generated a considerable amount of research in natural language processing over the last two decades (Morton, 2000). While I approached this task with more of a goal- than performance-oriented mindset, I must note that no coreference resolution system is perfect, and any large-scale analysis, particularly when working with Twitter data, is prone to include noise in its final results. I tried my very best to reduce this noise – specifically, by removing from analysis instances of pronouns that did not directly refer to the relevant celebrity – through six layers of filtering. This was by no means a perfect method, and I take full ownership for discarding tweets that contained accurate coreferential relationships and for including tweets with inaccurate ones. Though this method, I am making a trade-off between precision (in this case, the likelihood that the data consists of tweets where any pronouns in the tweet *actually refer* to the celebrity under analysis) and recall (the likelihood that the data consists of *all tweets* where the pronouns refer to the celebrity under analysis). The most selective filter (ALTHAND) for the PRONOUN analysis maximizes precision at the cost of recall, while the least selective filter (PRON) likely contains almost all the tweets containing pronouns

that potentially refer to the celebrity under analysis (high recall) in addition to many tweets containing pronouns that refer to other entities (low precision).

Table 7.1: Increasing selectivity throughout the tweet filtering process. PRONOUN analysis occurs at ALTHAND filter level; NAME analysis occurs at NAME filter level.

Celebrity	Total Tweets	DUP	NAME	PRON	COREF	ALTENT	ALTHAND
Target							
Jenner	2,613,733	2,452,601	2,250,303	547,483	409,054	320,698	293,513
Lovato	1,188,029	933,103	892,300	161,624	105,751	88,125	50,513
Smith	601,835	523,171	509,644	83,278	51,461	43,641	26,619
Page	267,027	263,666	253,842	76,217	42,683	37,843	22,930
comparison							
Doja	1,585,396	1,498,778	1,365,809	264,372	170,880	149,071	84,201
Holland	557,482	531,435	504,546	112,417	76,088	49,081	32,472
Cox	252,725	238,466	218,372	30,207	23,667	20,026	17,221

The six filters were ordered according to increasing selectivity. That is, I am most confident that the tweets which passed the last filter contain only pronouns that genuinely refer to the celebrity discussed in the tweet. Each pre-processed CSV was submitted to the same filtering process using Python, with some individual parameters set to account for differences in gender identity and named entity recognition between celebrities. The filtering process led to the creation of 6 additional CSV files, one for each filter. The results of this process are displayed in Table 7.1.

Evidently, this process was highly selective, with only N=527,469 tweets retained after the final filter. However, the process was also generally successful, with each successive filter improving the average affirming pronoun rate for the comparison celebrities (Table 7.2. Each layer of filtering is described in more detail below. I also include example tweets, slightly adapted from the Jenner corpus, that would have been removed at each level of successive filtering.

Table 7.2: Percent change in mean affirming pronoun rate across most selective filters follows expected patterns.

Celebrity Group	COREF over PRON	ALTENT over COREF	ALTHAND over ALTENT
Comparison			
Cox	6.09%	1.64%	2.12%
Holland	5.95%	10.91%	4.36%
Doja	6.49%	1.84%	5.21%
Average	**6.18%**	**4.80%**	**3.90%**
Target, Pre			
Jenner	-39.79%	-25.57%	-5.92%
Page	-20.75%	-11.04%	-16.19%
Smith	-12.86%	-10.70%	-16.07%
Lovato	-20.15%	-8.28%	-3.69%
Average	**-23.39%**	**-13.90%**	**-10.47%**

Filter 1: DUP The goal of this filter was to remove duplicate tweets from the data set. This was especially important given the prevalence of Twitter bots, who repeatedly tweet the same exact text. Furthermore, tweets that contain the same text but include links to the same media object (another tweet, gif, etc.) will often vary slightly, as Twitter appears to randomize these links. Thus, I removed non-celebrity Twitter handles and any links contained in the tweet before checking for duplication *within* an author. I decided not to remove duplicate tweets between authors because there is a high probability that, given the general brevity of tweets, there would be meaningful – albeit identical – tweets from multiple authors. Furthermore, although quote tweets can contain nearly all of the same language as the tweet being quoted, these tweets, similar to retweets, often function as endorsements of the sentiments expressed and language used in the tweet that is quoted.

Example tweet: 'Check out the new KUWTK episode'. Assuming that this tweet and the next example tweet were written by the same author, the tweet is a duplicate

and would skew analysis.

Filter 2: NAME The goal of this filter was to remove tweets that did not contain either the name/deadname or Twitter handle/deadhandle of the celebrity. As stated earlier, this was necessary due to the overly-permissive nature of the Twitter API full archive search. In this filter, each word of the tweet was checked against the list of name and handle tokens standardized in pre-processing. If none of these tokens was present in the tweet, the tweet was removed. I completed the (dead)naming analysis (NAME) using the data from this level of filtering. I expect this level to be sufficient for the NAME analysis because proper names are a much clearer signal to work with methodologically than third-person pronouns.

Example tweet: 'Check out the new KUWTK episode'. This tweet does not contain the (dead)name or (dead) Twitter handle of the celebrity under analysis.

Filter 3: PRON The goal of this filter was to remove tweets that did not contain any pronouns. This is the obvious first step in purposeful coreference resolution: delimiting the set to those tweets that have potential for pronominal coreference. In this filter, tweets that had a pronoun percent of 'None' were removed.

Example tweet: 'I'm so proud of Caitlyn Jenner'. This tweet does not contain pronouns.

Filter 4: COREF The goal of this filter was to make use of a fast, publicly available, and accessible coreference resolution module that utilizes a neural net scoring model: neuralcoref (Clark & Manning, 2016)[1]. Some research has suggested that coreference

[1] `https://github.com/huggingface/neuralcoref`

resolution systems are trained on gender-biased data that lacks instances of nonbinary and singular *they* (Cao & III, 2021; Webster, Recasens, Axelrod, & Baldridge, 2018). As such, I made use of neuralcoref in a very discriminative fashion. Rather than rely only on detected instances of coreference between the celebrity tokens and pronouns, I removed tweets that contained coreferential relationships between pronouns and *anything that was not* a celebrity name or handle token. Thus, this filter removed tweets containing third-person pronouns which the model confidently believed referred to other entities – not pronouns that may or may not have referred to the celebrity in question.

Example tweet: 'Caitlyn Jenner: What do Kendall and Kylie think of their dad's new name?' This tweet contains a coreferential relationship between a pronoun (*their*) and an entity (Kendall and Kylie) that is not the celebrity under analysis.

Filter 5: ALTENT The goal of this filter was to remove tweets that mentioned recognizable named entities. This was necessary because some tweets that contained a target or comparison celebrity's Twitter handle were actually discussing other, analytically-irrelevant celebrities. This process was partially automated and partially completed by hand. First, I made use of spaCy's out-of-the-box named entity recognizer, which labels some tokens (almost entirely other celebrities) as 'PERSON' entities. If a tweet contained a recognized named entity, it was removed. Second, I randomly sampled 150 tweets for each celebrity that made it through the most selective filtering level of an earlier approach. Using this sample, I identified additional irrelevant celebrities that appeared one or more times for each celebrity's corpus. I automated regular expressions to search for these celebrities within the tweet and removed them in an attempt to rid the data set of any pronouns used to refer to any entities not under analysis.

Example tweet: 'Caitlyn Jenner wears jeans and sweater as Kris drives her #Porsche' This tweet contains an alternate entity identified either by the NER or the regex celebrity list. As such, there is a chance that any pronouns contained in the tweet refer not to the celebrity under analysis but to one of these 'alternate entities'.

Filter 6: ALTHAND Similarly to ALTENT, this filter attempted to rid the data of any pronouns used to refer to another Twitter handle that was not the celebrity's. As such, ALTHAND removed any tweet containing a Twitter handle that did not match the celebrity's handle or deadhandle. While this filter is highly exclusionary – meaning that it discarded all tweets that replied to or tagged any Twitter user besides the celebrity under analysis – it was also shown to improve the mean affirming pronoun rate for comparison celebrities. I thus completed the misgendering analysis (PRONOUN) using the data from this, the most selective, level of filtering.

Example tweet: '@USER is attracted to Bruce Jenner as a woman... Does that make her a lesbian?' This tweet mentions the handle of a Twitter user that is not the celebrity under analysis. As such, there is a chance that any pronouns contained in the tweet refer not to the celebrity under analysis but to one of these 'alternate handles'.

The success of this purpose-oriented approach to coreference resolution, and of these filters specifically, is illustrated in Table 7.2. I observe a consistent improvement across filters in the affirming pronoun rate for the comparison celebrities, which adheres to my intuition: pronominal misgendering would be particularly rare for the cisgender comparison celebrities. Similarly, we see a consistent decrease across filters in the same measure for target celebrities before their COEs (the PRE condition). Recall that the the affirming pronoun rate is calculated using the celebrity's self-selected pronouns; before their COE, we would

expect to see much lower rates of affirming pronoun usage among Twitter users because it is not the suite the celebrity listed at the time.

7.4 Lexical processing

The computational methods used here were designed to explore the lexical content of tweets that misgender/deadname or gender-affirm, enabling final analysis to probe whether the lexical correlations align with attitudes, frames, and ideological principles characteristic of cisnormativity. To do so, I utilized measures of both sentiment valence and lexical association. The first of these was the VADER sentiment analysis tool, developed by Hutto and Gilbert (2014). VADER is a lexicon- and rule-based sentiment analysis tool that is trained specifically to measure sentiments expressed in social media text. During pre-processing, each tweet was assigned four VADER sentiment scores on a scale of -1 to 1: positive, negative, neutral, and a compound of the three metrics. I later excluded from analysis tweets that exhibited a compound score of 0; after checking by hand, it appeared that a score of 0 was more reliably an indication that VADER was unable to detect sentiment rather than the tweet having a truly neutral sentiment.

The compound VADER sentiment scores were aggregated and subsequently averaged across time and language practice conditions. The PRE condition consists of tweets discussing the target celebrities before their COE, while the POST consists of tweets from afterwards. These measures were used as baselines. The DEADNAME condition consists of tweets that only used deadnames when discussing the trans-binary celebrities (Caitlyn Jenner and Elliot Page, who use binary pronouns). Meanwhile, the NAME-AFFIRM condition consists of tweets that only used the trans-binary celebrities' gender-affirming names when discussing them. The MISGENDER condition consists of tweets that used exclusively mis-

gendering pronouns when discussing both the target and comparison celebrities. The PRON-AFFIRM condition consists of tweets that used exclusively gender-affirming pronouns when discussing both celebrity groups. Finally, I calculated a simple discrepancy measure between the negative (DEADNAME & MISGENDER) and positive conditions (NAME-AFFIRM & PRON-AFFIRM) across all celebrities.

I then adopted a fightin' words approach (Monroe, Colaresi, & Quinn, 2008) to contrast the lexical patterns in gender affirming and misgendering/deadnaming tweets. Specifically, I computed the weighted log-odds-ratio, informative Dirichlet prior algorithm for each of the conditions previously outlined. This method returns the correlations of lexical items across a binaristic measure with a prior used to balance the distribution of items. For the priors, I used all tweets at a filtering level higher than the one used for the main inputs. For example, in the NAME analysis, I used all tweets that exclusively used gender-affirming names in Filter 3: NAME as input 1, all tweets that exclusively used deadnames as input 2, and all tweets in Filter 2: DUP as the prior. Following Monroe et al. (2008), I consider lexical items with correlations above $|1.96|$ to be statistically significant. Ultimately, the two methodological approaches outlined here – the distributional and content analyses – allow me to probe the prevalence and lexical patterning of misgendering/deadnaming, particularly as they pertain to the diversity of gender expressions contained within the celebrity corpora.

CHAPTER 8
RESULTS

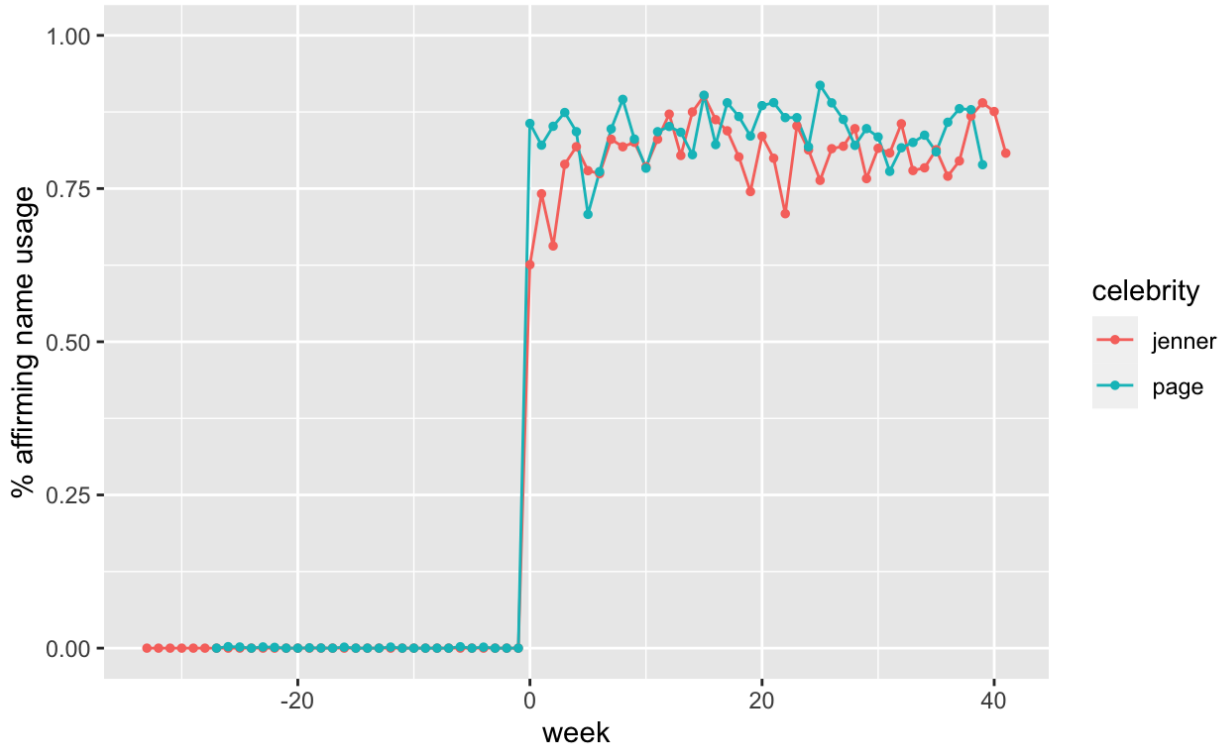
In this chapter, I share results from the distributional and content analyses. As part of the distributional analysis, I review descriptive statistic data and the results of several Automated Dickey-Fuller (ADF) tests which reveal the prevalence and chronological nature of deadnaming and misgendering before and after the celebrities' coming-out events (COE). Turning to the content analysis, I outline the results – and experimental validity – of two out-of-the-box computational linguistic methods before presenting results from a series of binary logistic regressions that makes use of an original lexical content measure.

8.1 Distributional analysis

Results from the NAME analysis (Figure 8.1, Table 8.2) demonstrated that users uptake the trans-binary celebrities' novel gender-affirming names post-COE at a stable rate of 82.47% on average. Inversely, we see that deadnaming occurs at a high and remarkably similar rate across both celebrities. Tweets were aggregated by week and name rates – the amount of affirming names divided by the total number of names in a given tweet – were averaged across each week. Deadnaming rates appeared very similar between Caitlyn Jenner and Elliot Page, with deadnaming occurring slightly less for Page. As expected, deadnaming rates in the weeks leading up to their COEs were near or at 0. Any nonzero name rate measure from this period resulted from to Twitter user speculation[1] (in the case of Jenner)

[1]For example, Wendy Williams asked fans to guess what Caitlyn Jenner's novel gender-affirming name would be, generating a few accurate guesses in response.

Figure 8.1: **Affirming name uptake for trans-binary celebrities across entire analysis period, with week 0 representing the coming-out event.**



or noise[2] in the data (in the case of Page), as confirmed by hand. Furthermore, gender-affirming name usages in the PRE condition were incredibly rare: 12 out 20,312 tweets for Page and 5 out of 651,112 tweets for Jenner. This is indicative of the success of the DUP and NAME filters.

NAME results were also submitted to Automated Dickey-Fuller (ADF) tests (Dickey & Fuller, 1979), across the GROSS (entire analysis period), PRE (before COE), and POST (after COE) conditions (Table 8.1). Tweets were aggregated by day and name rates were

[2]Despite the unique spelling of Elliot, a few tweets contained the proper name Elliot with the lexical item 'page' in immediate succession.

Table 8.1: NAME ADF tests for trans-binary celebrities in GROSS, PRE-COE, and POST-COE conditions. Test-statistic lower than $\tau 3$ rejects null hypothesis, suggesting stationarity.

Celeb	Test-statistic	$\tau 3$	AIC-lag
Jenner			
GROSS	-1.7632	-3.96	22
PRE	-4.2942**	-3.99	7
POST	-4.2179**	-3.98	5
Page			
GROSS	-1.6412	-3.98	3
PRE	-3.4358*	-3.43	9
POST	-4.9173**	-3.98	4

** indicates 99% confidence interval.
* indicates 95% confidence interval.

Table 8.2: Affirming name rate descriptive statistics for trans-binary celebrities in PRE-COE and POST-COE conditions.

	PRE				POST			
	Weeks	Mean	σ	IQR	Weeks	Mean	σ	IQR
Jenner	31	2.60E-06	1.05E-05	0	41	0.8063309	0.05605699	0.06156047
Page	26	0.0005781127	0.0008758778	0.001448508	39	0.8430067	0.04109404	0.0484912
Average		**2.90E-04**	**4.43E-04**	**7.24E-04**		**0.8246688**	**0.048575515**	**0.055025835**

averaged across each day. The null hypothesis of the ADF test is the presence of a unit root, or a statistical effect of time, in the data. The optimal lag length of the ADF test, in days, was determined using the Aikaike Information Criterion (AIC) measure. It is important to specify the lag length using reliable criteria to avoid biasing the test with a too short lag or reducing its power with a too long lag.

As expected, ADF test results in the GROSS condition did not reject the null hypothesis, indicating that the data was not stationary: time had a significant effect on the uptake of the celebrities' gender-affirming names over the full analysis period. Meanwhile, test statistics in the PRE and POST conditions were great enough to significantly reject the null hypothesis, indicating stationarity in the data. This suggests that the uptake of affirming names following

the COE was practically instantaneous, as visually evidenced by Figure 8.1. That is, the Twitter users who intended to affirm the transgender identities of these celebrities by using their gender-affirming names post-COE did so immediately over the span of days. Meanwhile, the practice of deadnaming was similarly persistent and stable across the analysis period. However, this analysis leaves open the question of how gender identity information diffuses across social networks at the user level in the immediate aftermath of a COE. It is also important to note that, by definition, the affirming name rate for comparison celebrities was 100%. Though the lack of a proper comparison for this aspect of the study is not ideal, the fact of the matter is that there is no similar naming practice for cisgender individuals that carries the gender identity-denying charge that deadnaming does for transgender individuals.

Results from the PRONOUN analysis (Figure 8.3, Table 8.4) demonstrated that affirming pronoun rates varied across the target group but were relatively consistent across the comparison group. Tweets were aggregated by week and pronoun rates were averaged across each week. It is important to note that two separate measures of affirming pronoun rates were considered for analysis. The first measure includes whatever pronoun suite might be alternate or irrelevant for a given celebrity (affirming pronouns over total *he/she/they* count). However, the results illustrated in Table 8.3 initiated a reconsideration of this metric due to the high prevalence of singular *they*[3] among the comparison celebrities (15.97% of tweets, on average).

This suggests that Twitter users already make use of singular *they* to reference binary gendered celebrities in a gender-unspecified manner. This level of noise in the data – particularly the high level of *they* – also explains why affirming pronoun rates in the PRE condition for target celebrities was not near or at zero (Table 8.5). In the case of Lovato and Smith,

[3]Undoubtedly, in addition to instances of plural and generic, indefinite antecedent *they* that slipped through the filtering process.

Table 8.3: Usage rates by pronoun form for comparison celebrities across entire analysis period. AFFIRM-RATE represents affirming pronoun usage, ALT-RATE represents misgendering usage, and THEY-RATE represents singular *they* and instances of plural *they* that slipped through the filtering process.

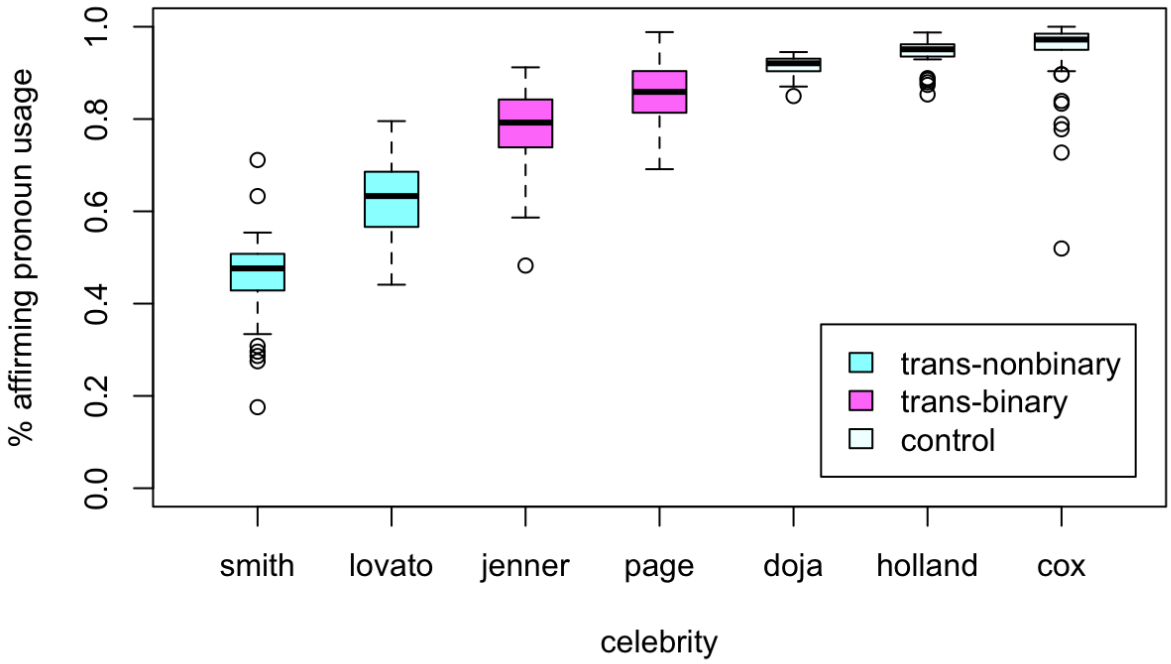
	Pronouns	AFFIRM-RATE	ALT-RATE	THEY-RATE
Cox	*she/her*	0.837	0.043	0.120
Doja	*she/her*	0.747	0.063	0.189
Holland	*he/him*	0.781	0.049	0.170
Average		**0.788**	**0.052**	**0.160**

for example, the high PRE affirming pronoun rate (usage of *they)* is the result of singular, as opposed to specifically nonbinary, *they* usage. We might also reasonably expect the comparison group – the cisgender celebrities, especially – to have their pronouns affirmed at rates close to 100%. Even at the most selective level of the filtering process, this was not the case. While any instance of misgendering can be harmful – and, indeed, by-hand analysis confirmed the incredibly rare but nonetheless nonzero presence of harmful *they*-misgendering tweets for Caitlyn Jenner – the large presence of singular and epicene *they* in the comparison data set led me to develop a more restrictive measure of affirming pronoun rates.

For the second measure, I excluded tweets that contained any pronouns part of the alternate or irrelevant suite for all celebrities in the data set. For example, all tweets containing *they* for Laverne Cox and Doja Cat were excluded from analysis, as *they* pronouns were at no point relevant to the expression of either celebrities' gender identities at any point in time. This measure elicited affirming pronoun rates for the cisgender celebrities that align much more closely with what I expected intuitively: pronominal misgendering, whether habitual or intentional, is a rarity for cisgender individuals. Elliot Page, who lists *he/they*, is the only celebrity in the data set for whom this second measure may not align correctly, as the exclusion of all tweets containing *they* will certainly dispose of some tweets where *they* was used to

affirm[4] Page's gender identity. However, the mean affirming pronoun rate across all tweets in Page's POST condition did not change too considerably from the first, *he/they*-inclusive measure (90.01%) to the second, *he*-only measure (85.89%). The rest of the PRONOUN results reported in this work make use of the second measure of affirming pronoun rate (affirming pronouns over the sum of affirming and directly misgendering pronouns).

Figure 8.2: **Mean affirming pronoun rate by week for all celebrities. Data for comparison celebrities spans entire analysis period, while data for target celebrities is POST-COE.**



Considerations of noise and proper pronoun measures dealt with, the results indicate a clear patterning in the extent to which Twitter users affirm the pronouns of trans-binary

[4]Indeed, average *they* usage increased for Page after his COE, from 20.17% in the PRE condition to 27.71% in the POST condition. This 7% increase is likely reflective of gender-affirming *they* uptake rather than epicene usages.

celebrities (Jenner and Page), trans-nonbinary celebrities (Smith and Lovato), and comparison celebrities (Doja, Holland, and Cox). Figure 8.2 displays the week-aggregated mean affirming pronoun rate for the comparison celebrities and for the POST condition of the target celebrities. This analysis was enabled by the statistical stationarity observed in the pronoun ADF tests for target celebrities in the around-9 month POST condition (Table 8.6) and across the entire comparison celebrity data sets (Table 8.7). Because there is no effect of time post-COE in any of these generally long-term data, I presume that these rates of misgendering may be generally unchanging.

Table 8.4: Affirming pronoun rate descriptive statistics for target celebrities in PRE-COE and POST-COE conditions.

		PRE				POST			
Celebrity	Pronouns	Weeks	Mean	σ	IQR	Weeks	Mean	σ	IQR
Trans-nonbinary									
Smith	*they/them*	26	0.1565974	0.04493363	0.05707767	39	0.4571644	0.09605299	0.07620958
Lovato	*they/them*	26	0.1602483	0.06448988	0.07431863	39	0.6289992	0.07665435	0.1133933
Average			**0.15842285**	**0.054711755**	**0.06569815**		**0.5430818**	**0.08635367**	**0.09480144**
Trans-binary									
Jenner	*she/her*	31	0.08032814	0.06572916	0.07834197	41	0.7782324	0.0865904	0.1009244
Page	*he/they*	26	0.06805317	0.06760755	0.0691342	39	0.8588556	0.06114524	0.08771423
Average			**0.074190655**	**0.066668355**	**0.073738085**		**0.818544**	**0.07386782**	**0.094319315**
Total average			**0.1163067525**	**0.060690055**	**0.0697181175**		**0.6808129**	**0.080110745**	**0.0945603775**

Table 8.5: Affirming pronoun rate descriptive statistics for comparison celebrities across entire analysis period.

Celebrity	Pronouns	Weeks	Mean	σ	IQR
Transgender					
Cox	*she/her*	104	0.955228	0.06346705	0.03491379
Cisgender					
Doja	*she/her*	27	0.9136034	0.02229997	0.02717804
Holland	*he/him*	27	0.9409226	0.03422071	0.02651157
Average			**0.927263**	**0.02826034**	**0.026844805**

Led by Cox, with a mean affirming pronoun rate of 95.52%, the comparison celebrities elicited more pronominally gender-affirming tweets than the trans-binary celebrities, who

in turn surpassed the trans-nonbinary celebrities. While the trans-binary celebrities' pronouns were affirmed around 10% less frequently than the comparison celebrities, averaging to 81.85%, more than half of the pronouns in the POST conditions misgender Sam Smith and only slightly less than half misgender Demi Lovato. Furthermore, I also observe a historical pattern of COE year within the two target groups. Smith (COE: 9/13/19) presented their nonbinary identity a little less than two years before Lovato did (COE: 5/19/21); Jenner (COE: 6/1/15) first presented her transgender identity a little over five years before Page did (COE: 12/1/20). That in both groups the higher affirming pronoun rate is observed for the celebrity with the chronologically later COE suggests that broad social shifts in notions of gender may be contributing to these patterns (Geiger & Graf, 2019; Zimman, 2020). Additionally, lower standard deviations were observed for the cisgender comparison celebrities than for Cox and the target celebrities (Tables 8.4 and 8.5). This itself indicates that affirming pronoun usage is more stable for the cisgender celebrities week-over-week and that the greater standard deviation values for Cox and the target celebrities may be attributable to a higher tendency for Twitter users to misgender these celebrities.

Table 8.6: PRONOUN ADF tests for target celebrities in GROSS, PRE-COE, and POST-COE conditions. Test-statistic lower than $\tau 3$ rejects null hypothesis, suggesting stationarity.

	GROSS			PRE			POST		
Celebrity	Test-statistic	tau3	AIC-lag	Test-statistic	tau3	AIC-lag	Test-statistic	tau3	AIC-lag
Trans-nonbinary									
Smith	-2.2066	-3.98	10	-7.6915**	-3.99	1	-6.0576**	-3.98	5
Lovato	-2.3188	-3.98	7	-6.0167**	-3.99	1	-7.8767**	-3.98	1
Trans-binary									
Jenner	-2.2708	-3.96	7	-7.3385**	-3.99	1	-5.247**	-3.98	3
Page	-1.5109	-3.98	11	-3.4398**	-3.43*	11	-10.9851**	-3.98	1

** indicates 99% confidence interval.

PRONOUN results were also submitted to Automated Dickey-Fuller (ADF) tests (Dickey & Fuller, 1979), across GROSS (entire analysis period), PRE, and POST conditions (Tables

Table 8.7: PRONOUN ADF tests for comparison celebrities across entire analysis period. Test-statistic lower than $\tau 3$ rejects null hypothesis, suggesting stationarity.

Celebrity	Test-statistic	Tau3	AIC-lag
Cox	-16.5326**	-3.96	1
Doja	-6.5577**	-3.99	1
Holland	-4.3105**	-3.99	8

** indicates 99% confidence interval.

8.6 and 8.7). The optimal lag length of the ADF test was again determined using the Aikaike Information Criterion (AIC). For the target celebrities, ADF test results in the GROSS condition did not reject the null hypothesis, indicating that the data was not stationary: time had a significant effect on the uptake of the celebrities' gender-affirming pronouns over the full analysis period. As expected, results in the PRE and POST conditions for all celebrities were significant enough (CI≤0.01) to reject the null hypothesis, indicating stationarity. This suggests that post-COE, there was no sort of significant gradual uptake period in affirming pronoun usage, as visually evidenced by Figure 8.3. Rather, uptake happens practically immediately (in the span of hours and minutes, not days) and affirming usage rates remain stable thereafter.

Crossover analysis for the trans-binary celebrities reveals the interrelated nature of misgendering and deadnaming. Table 8.8 shows that misgendering tweets were also more likely to deadname, and vice versa. This pattern is likely related to a gender expectancy effect (Doherty & Conklin, 2017), whereby users who are already using an injurious nominal form are more likely to use the pronominal form that matches their lexical entry for that name (Konnelly & Cowper, 2020). In other words, users who are already misgendering/deadnaming have a greater tendency to assume and assign the (pro)nominal form that makes the coreferential relationships coherent in terms of gender (McConnell-Ginet, 2003).

Figure 8.3: **Affirming pronoun uptake for target celebrities across entire analysis period, with week 0 representing the coming-out event.**

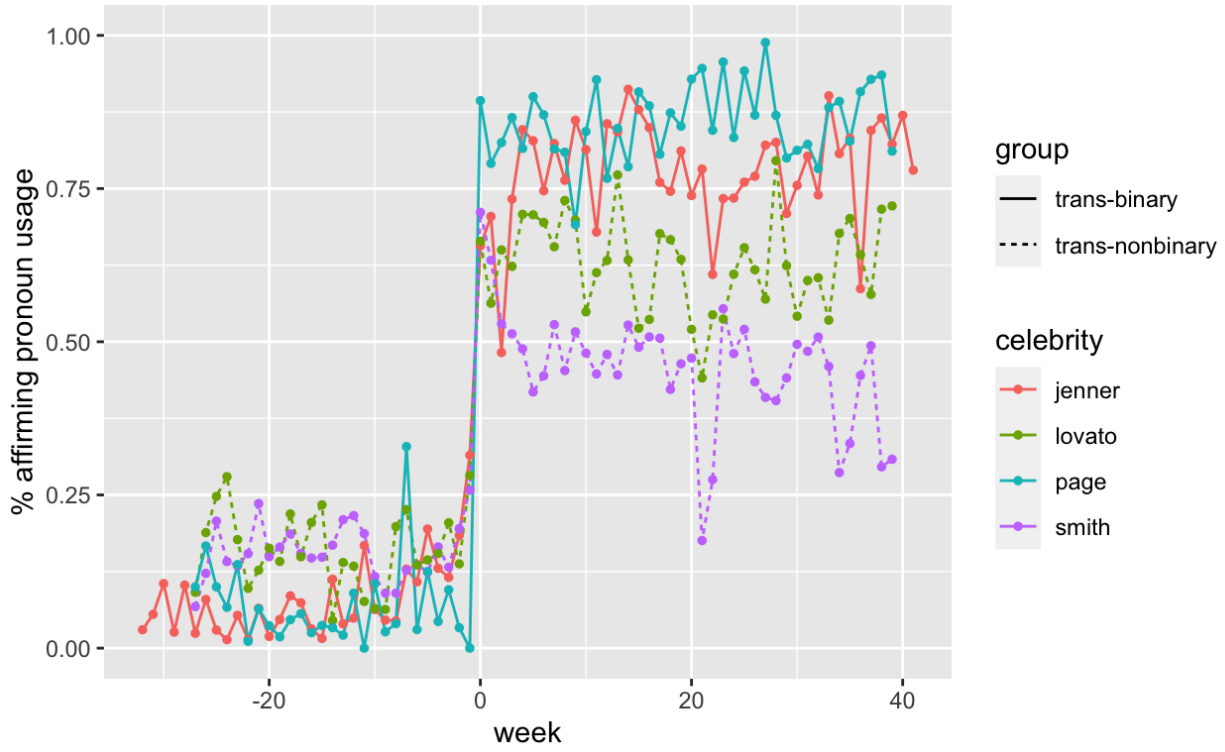


Table 8.8: Crossover NAME and PRONOUN results for trans-binary celebrities across Aff (rate=1.0), Mix (0.0<rate<1.0), and Mis/Dead (rate=0.0) conditions.

	Mean affirming name rate			Mean affirming pronoun rate		
	Pr-Aff	Pr-Mix	Pr-Mis	N-Aff	N-Mix	N-Dead
Jenner	0.9182809	0.5605776	0.3199732	0.8937198	0.6055047	0.1753081
Page	0.9560607	0.6942159	0.3768535	0.9642512	0.7036747	0.2822468
Average	**0.9371708**	**0.62739675**	**0.34841335**	**0.9289855**	**0.6545897**	**0.22877745**

8.2 Content Analysis

In the content analysis, I review results from VADER sentiment scores (Hutto & Gilbert, 2014), fightin' words lexical associations (Monroe et al., 2008), and binary logistic regres-

sions designed to test both tweet-level (likes, replies, lexical features, etc.) and user-level (pronouns in bio, followers, etc.) factors. While these measures are not directly indicative of cisnormativity *in se*, they reveal specific patterns of sentiment and lexical disparity between groups that encourage an interpretive analysis within the ideological framework of cisnormativity. That is, while ideologies are not quantitatively measurable, I elect to explain the quantitative results from this content analysis through consideration of the ideological principles of cisnormativity outlined in Section 2.4.

All tweets from the NAME filter in the POST condition that did not have a VADER score of 0.0 were included[5] for the DEADNAME VADER analysis. Results demonstrate that deadnaming carries more negative sentiment than name-affirming tweets. While there is no experimental comparison for these results, they are similar in size to the misgendering vs. gender-affirming pronoun discrepancies observed for the same celebrities in the PRONOUN VADER analysis results (Table 8.11).

Table 8.9: Trans-binary VADER sentiment scores across PRE-COE and POST-COE conditions in the NAME analysis.

	PRE	POST		
Celebrity		NAME-AFFIRM	DEADNAME	DISCREPANCY
Jenner	0.1096007	0.1987948	0.08348442	0.11531038
Page	0.2794252	0.3664013	0.1395359	0.2268654
Average	**0.19451295**	**0.28259805**	**0.11151016**	**0.17108789**

A greater discrepancy between deadnaming and affirming tweets was observed for Page (0.226) than for Jenner (0.115). This is perhaps explainable through the observation that the average sentiment in tweets discussing Page before his COE (0.279) was higher than for Jenner (0.11) or the observation that Page was deadnamed less after COE than Page – while

[5]There remains the possibility that some of the excluded tweets were truly balanced in sentiment, though by-hand analysis revealed that such tweets constituted a very small portion of the data.

deadnaming was rarer for Page, it seems that it may have also been more severe.

Table 8.10: NAME lexical associations for trans-binary celebrities by DEADNAME (name rate=0.0) and NAME-AFFIRM (rate=1.0) in the POST condition.

JENNER				PAGE			
DEADNAME		NAME-AFFIRM		DEADNAME		NAME-AFFIRM	
Score	Term	Score	Term	Score	Term	Score	Term
-396.828	CELEB_DEADNAME	373.035	CELEB_NAME	-81.295	CELEB_DEADNAME	103.708	CELEB_NAME
-334.94	jenner	98.865	TARGET_PRONOUN	-52.909	TWITTER_HANDLE	21.253	TARGET_PRONOUN
-119.679	WRONG_PRONOUN	58.185	cait	-32.688	WRONG_PRONOUN	17.628	coming
-78.138	TWITTER_HANDLE	40.55	espys	-21.577	DEAD_HANDLE	14.451	star
-63.243	😂	38.142	via	-11.606	woman	14.365	transgender
-58.531	woman	36.505	new	-9.624	sex	12.713	trans
-50.98	interview	35.278	trans	-9.287	actress	12.039	announces
-37.761	shit	34.167	photos	-8.673	😂	10.95	love
-33.796	21stepsmichael	33.371	speech	-8.544	formerly	10.644	happy
-33.55	nigga	33.099	costume	-8.507	mental	10.642	academy
-32.22	sawyer	32.949	vanity	-8.453	female	10.418	umbrella
-32.199	diane	32.055	manslaughter	-8.363	⭐	10.277	actor
-31.708	sex	31.955	espy	-8.252	male	9.949	portner
-30.996	caitlin	31.795	fair	-8.17	science	9.906	oprah
-30.773	change	31.712	2015	-7.845	gender	9.851	deadname
-29.538	like	31.086	kylie	-7.366	women	9.739	juno
-28.062	fuck	30.766	first	-7.26	born	9.28	news
-27.665	man	30.499	laverne	-7.148	girl	9.208	emma
-27.522	😭	30.331	halloween	-6.924	🤣	9.104	came
-26.601	niggas	30.248	cox	-6.899	gay	8.924	divorce
-25.849	female	28.02	awards	-6.724	recast	8.583	shares
-25.338	lol	27.802	kevinonliner	-6.459	#mugclub	8.236	shirtless
-24.134	watching	26.742	charge	-6.366	change	8.139	first
-22.422	smh	26.565	glamour	-6.199	biology	7.941	deadnaming
-21.361	really	26.288	face	-5.97	degeneres	7.459	interview
-21.095	got	26.086	kris	-5.832	illness	7.354	top
-20.121	still	26.047	.@	-5.726	biological	7.346	abs
-19.485	ass	25.988	stewart	-5.67	mean	7.161	surgery
-19.426	u	25.819	jon	-5.664	eliot	7.146	stan
-19.399	becoming	25.593	award	-5.583	child	7.114	terfs
-19.368	dude	25.57	candis	-5.466	latinx	7.087	today

Note: correlation scores above ±1.96 are considered significant (Monroe et al., 2008). Full fightin words' results are in the Appendix.

All tweets from the NAME filter in the POST condition that were either entirely deadnaming (0.0) or name-affirming (1.0) were included for the DEADNAME lexical analysis. The reason for excluding tweets with mixed name usage is that it is much harder to attribute intentionality to tweets that do not solely affirm or reject the gender identities of these celebrities. Results show that deadnaming is significantly correlated with lexical items

that relate to gender identity and indicate a direct targeting of the celebrity. The top 30 correlations are displayed in Table 8.10. Across Jenner and Page, significantly DEADNAME-correlated include items such as: WRONG_PRONOUN, TWITTER_HANDLE[6], woman, sex, and female. Notable DEADNAME-correlated lexical item groups include those relating to gender/sex (sex, biological, gender, male, woman, science, female, biology, born), change (becoming, change, formerly), and humor (lol, 🤣, 😂, smh). Furthermore, these results indicate a very strong correlation between deadnaming tweets and the TWITTER_HANDLE of both celebrities and the DEAD_HANDLE of Page[7].

Meanwhile, tweets in the NAME-AFFIRM condition are correlated with terms that relate to the general nature of Jenner and Page's celebrity. In the case of Jenner, NAME-AFFIRM-correlated terms reference her Vanity Fair cover (cait, photos, vanity, glamour), her appearance at the ESPN 'ESPY' award show (espys, espy, awards, award), other members of the Kardashian clan (kylie, kris), and her transgender identity specifically (trans, laverne, cox). NAME-AFFIRM-correlated tweets for Page also generally discuss aspects of his celebrity, including discussion of his acting gigs (umbrella, academy, star, actor), his divorce from his wife (emma, portner, divorce), general sentiments of support (love, happy, news, stan[8]), and his transgender identity specifically (coming, transgender, trans, announces, deadname, came, shares, deadnaming).

All tweets from the ALTHAND filter for the comparison celebrities and all tweets in the POST condition for the target celebrities that did not have a VADER score of 0 were included for the PRONOUN VADER analysis. Results from the pronoun analysis demonstrate that, on average, the sentiment discrepancy between trans-binary celebrities is higher than

[6]Caitlyn Jenner did not have a Twitter account before her COE.

[7]Twitter handle tokens were not included in the calculation of name percentages.

[8]'Stan' is Twitter-speak for an extreme fan, and can be used as a noun or a verb: 'I am an Elliot Page stan' or 'I stan Elliot Page'.

Table 8.11: Target celebrity VADER sentiment scores by pronoun rate across PRE-COE and POST-COE conditions in the PRONOUN analysis.

	PRE	POST		
Celebrity		PRON-AFFIRM	MISGENDER	DISCREPANCY
Trans-nonbinary				
Smith	0.2928862	0.2272077	0.2123156	0.0148921
Lovato	0.1510316	0.2426199	0.02633819	0.21628171
Average	**0.2219589**	**0.2349138**	**0.119326895**	**0.115586905**
Trans-binary				
Jenner	0.2130364	0.2760341	0.1148471	0.161187
Page	0.3086618	0.4471038	0.1217256	0.3253782
Average	**0.2608491**	**0.36156895**	**0.11828635**	**0.2432826**
Total Average	**0.241404**	**0.298241375**	**0.1188066225**	**0.1794347525**

Table 8.12: Comparison celebrity VADER sentiment scores by pronoun rate across entire analysis period in the PRONOUN analysis.

Celebrity	PRON-AFFIRM	MISGENDER	DISCREPANCY
Transgender			
Cox	0.4609984	0.3538285	0.1071699
Cisgender			
Doja	0.2106753	0.122494	0.0881813
Holland	0.2287218	0.1480358	0.080686
Average	**0.21969855**	**0.1352649**	**0.08443365**
Total Average	**0.3001318333**	**0.2081194333**	**0.0920124**

for trans-nonbinary celebrities, who more closely pattern with the cisgender celebrities (Table 8.11, Table 8.12). However, individual measures were more complex. Tweets discussing Page exhibited the greatest discrepancy between PRON-AFFIRM and MISGENDER tweets (0.325), while those discussing Smith exhibited the lowest (0.015). These results are complicated by an unexpectedly high difference between PRON-AFFIRM and MISGENDER tweets for the cisgender comparison celebrities. That tweets discussing cisgender celebrities

should exhibit a discrepancy at all between the two conditions suggest that this finding is likely the result of systematic noise in the data, suggesting that the filtering process may be over-selecting tweets with more negative lexical content. Nevertheless, tweets mentioning Cox, Jenner, Lovato, and Page all exhibited higher discrepancy scores than the cisgender average (0.084). Additionally, I observed a similar effect pattern within target groups from the distributional PRONOUN analysis: tweets discussing Lovato (later COE) exhibit a much greater sentiment discrepancy between misgendering and affirming than for Smith (earlier COE), whereas tweets discussing Page (later COE) exhibit a greater discrepancy than for Jenner (earlier COE). This indicates that greater identity recognition – as indicated by higher mean affirming pronoun rates – may historically co-occur with more intentional and negative misgendering, and that these effects could be related to the historical time periods in which these celebrities presented their transgender identity (Geiger & Graf, 2019; Zimman, 2020). These results are further contextualized by findings from the binary regression analyses detailed below.

First, it is critical to note that I also conducted PRONOUN lexical analyses using data from the ALTHAND filter for all celebrities in the data set. However, these results were more spurious in their ideological fingerprint when compared to the NAME analysis. This is perhaps due to the fact that the generally smaller size of the data set and more aggressive, precision-oriented filtering allowed noise to more directly bias the results of the tests. The top 30 correlations from each of these lexical tests can be found in the Appendix.

These findings are suggestive of different lexical patterns between tweets that misgender the trans-binary celebrities versus those that misgender the trans-nonbinary celebrities. However, I wanted to test more concretely whether the presence of words in specific, ideologically-charged lexical categories predicted misgendering/deadnaming and gender-affirming lan-

guage practices at the tweet level. I thus drew upon the exploratory results from Jenner and Page's NAME lexical analysis to develop lexicons representing key words in eight relevant categories that might better indicate the different ideological aspects that are associated with these language practices. For example, the biological essentialism terms reflect the cisnormative principle that the man-woman gender binary corresponds to the male-female sex binary, to which individuals are naturally assigned at birth. The binary gender terms reflect the precept that the gender of all individuals can be classified using the man-woman binary. Finally, the hate speech terms represent enforcement of the cisnormative belief that an individual cannot change their gender identity over a lifetime.

I used these features to conduct confirmatory regression analyses described below. While the effects of these variables may be rather circular in the NAME binary logistic regression, they gain methodological validity in the PRONOUN regressions, as the categories were not determined using results from lexical analyses of these same data. Ultimately, these regression analyses enable identification of recurring lexical items that co-occur with misgendering and deadnaming, allowing me to analyze whether the patterns correspond to cisnormative ideological frames (Borba & Milani, 2017; Ericsson, 2018; Hornscheidt, 2015; Turton, 2021) while controlling for user- and tweet-level factors. The lemmas used in each of the eight lexical categories are listed below:

Transgender identity terms: [transgender, trans, pronoun, non, binary, nonbinary, misgender, misgendere, misgendering, enby, nb, transphobic].

Binary gender terms: [woman, girl, male, female, man, boy, masculine, feminine, dude, chick, guy].

Gender/sex terms: [gender, gendered, sex].

LGBTQ+ terms: [straight, lesbian, gay, sexuality, lgbt, lgbtq, queer].

Coming-out event terms: [come, out, revealing, reveal, announce, journey, formerly, transition, change, declare, identifie, identify, unveil, identity, embrace, introduce, news].

Biological essentialism terms: [science, biological, surgeon, surgery, chest, penis, ball, pussy, tit, dick, chromosome, implants, vagina, implant, boob, breast, tuck, surgically, mutilate, remove, operation].

Hate speech terms: [faggot, illness, psychotic, mental, delusional, crazy, tranny, bizarre, fag, disorder, disgusting, transvestite, bitch].

Pride and support terms: [proud, pride, support, happy, joy, celebrate, beautiful, gorgeous, amazing, love, happy, congrat, congratulation, equality, confidence, respect].

The raw text from each tweet was lemmatized using spaCy's lemmatizer, after which each lemma was checked against all eight lists to determine the presence of these categories in a given tweet. Though tweets were initially assigned a continuous score throughout this process (+1 for each lemma identified in the tweet), these measures were later binarized (score $>$ 0 *or* score = 0) to facilitate easier comparison across effects and groups. Other independent variables were also included in the regression models. These included tweet-level characteristics (the tweet's VADER compound score and the amount of both likes and replies the tweet received) as well as user-level characteristics (whether or not the user listed pronouns in their bio, whether or not the user's bio contained a LGBTQ+ or trans pride flag, the number of followers the user has, and how many accounts the user follows). As a more rigorous test of the interrelated nature of deadnaming and misgendering for the trans-binary celebrities suggested descriptively by Table 8.8, NAME regression models also included a

continuous affirming pronoun rate measure while PRONOUN models included a continuous affirming name rate measure. To identify the clearest instances of misgendering and gender-affirming, tweets that had a 'mixed' affirming name/pronoun rate ($0.0 < \text{rate} < 1.0$) were excluded from analysis, thus making the regression models truly binary.

Table 8.13 displays the results from the NAME binary logistic regression. To combine the trans-binary data sets, I recorded the number of observations from the smaller data set (Page, N=136,848) and randomly sampled the same number from the larger data set (in this case, Jenner). I find that all of the devised lexical categories beside LGBTQ+ terms significantly predict whether a tweet will deadname or affirm the trans-binary celebrity under discussion. Furthermore, significant effects are observed for the continuous VADER compound variable, the continuous affirming pronoun rate variable, and all the user characteristic variables. The strongest significant binary predictor of name-affirming tweets was the presence of pronouns in the user's bio (β=1.93586113, $p\leq$0.001) while the strongest binary predictor of deadnaming tweets was the presence of hate speech terms in the tweet (β=-0.91228621, $p\leq$0.001). Meanwhile, the strongest significant continuous predictor of name-affirming tweets was, perhaps to be expected, the rate at which the tweet used gender-affirming pronouns (β=1.17696486, $p\leq$0.001). These results strengthen the observations from the NAME lexical analysis. I find that the presence of transgender, COE, and pride terms significantly predict name-affirming tweets while the presence of binary gender, gender/sex, biological essentialiasm, and hate speech terms significantly predict deadnaming tweets. Furthermore, tweets written by users with either LGBTQ+ or trans flags in their bio were significantly more likely to use the celebrity's affirming pronoun suite (β=0.82469496, $p\leq$0.001). Lastly, tweets written by users with more expansive social networks (as measured by number of followers (β=0.21774302, $p\leq$0.01) and following (β=0.07520546, $p\leq$0.05) were significantly

more likely to use gender-affirming names.

Table 8.13: NAME binary logistic regression in POST-COE condition. Data aggregated from equal number of tweet-level observations for trans-binary celebrities.

Variable	Estimate	σ	t-statistic	p-value	Sig.
Intercept	1.98947755	0.0266899	74.5404664	<0.01	***
Tweet characteristics					
Vader compound	0.0796929	0.01746908	4.5619409	<0.01	***
% affirming pronouns	1.17696486	0.0142066	82.8463608	<0.01	***
Likes	0.33183225	0.22213733	1.4938158	0.14	
Replies	-0.08052594	0.04161068	-1.9352228	0.05	
User characteristics					
Pronouns in bio	1.93586113	0.07700463	25.1395435	<0.01	***
Flag in bio	0.82469496	0.18013275	4.5782623	<0.01	***
Followers	0.21774302	0.07566461	2.8777393	<0.01	**
Following	0.07520546	0.03382711	2.2232304	0.03	*
Lexical characteristics					
Trans terms	0.70927027	0.06525659	10.8689441	<0.01	***
Binary gender terms	-0.73926508	0.03943366	-18.7470565	<0.01	***
Gender/sex terms	-0.36986305	0.09186218	-4.0262819	<0.01	***
LGBTQ+ terms	-0.06125005	0.07794996	-0.7857611	0.43	
COE terms	0.10707955	0.04000839	2.6764274	0.01	**
Biological terms	-0.2342998	0.07369388	-3.1793658	<0.01	**
Hate speech terms	-0.91228621	0.11766785	-7.7530624	<0.01	***
Pride terms	0.2812753	0.05532936	5.0836537	<0.01	***

*** = p-value≤0.001; ** = p-value≤0.01; * = p-value≤0.05.

Table 8.14 displays the results from the PRONOUN binary logistic regression for the target celebrities, separated by pronoun group. To combine the trans-binary data sets, I recorded the number of observations from the smaller data set (Page, N=16,248) and randomly sampled the same number from the larger data set (Jenner). To combine the trans-nonbinary data sets, I also recorded the number of observations from the smaller data set (Smith, N=24,778) and randomly sampled the same number from the larger data set (Lovato). I find that all of the devised lexical categories beside LGBTQ+ terms significantly

predict whether a tweet will misgender or affirm the trans-binary celebrity under discussion, while LGBTQ+ terms was also a significant predictor for the trans-nonbinary celebrities.

For the trans-binary celebrities, PRONOUN results somewhat mirror those observed in the NAME analysis: the strongest significant binary predictor of gender-affirming tweets was the presence of pronouns in the user's bio (β=1.10367387, p<0.001) while the strongest binary predictor of misgendering tweets was the presence of binary gender terms in the tweet (β=-0.72769207, p<0.001). Meanwhile, the strongest significant continuous predictor of gender-affirming tweets was, again as expected, the rate at which the tweet used the celebrity's gender-affirming name (β=1.32883589, p<0.001). I find that the presence of transgender, COE, and pride terms significantly predict gender-affirming tweets while the presence of binary gender, gender/sex, biological essentialiasm, and hate speech terms significantly predict misgendering tweets. Furthermore, tweets written by users with either LGBTQ+ or trans flags in their bio were significantly more likely to use the celebrity's affirming name (β=0.63054446, p<0.001). Lastly, tweets written by users with more expansive social networks (as measured by number of followers) were significantly more likely to use gender-affirming names (β=1.00872449, p<0.01) – however, the users following measure did not reach significance (p=0.19).

For the trans-nonbinary celebrities, PRONOUN regression results differ from those in the trans-binary model. The strongest significant binary predictor of gender-affirming tweets by far was the presence of transgender terms (β=1.454819239, p<0.001) while the strongest binary predictor of misgendering tweets was also the presence of binary gender terms in the tweet (β=-0.931703764, p<0.001). I find that the presence of transgender, gender/sex, COE, and pride terms significantly predict gender-affirming tweets while the presence of binary gender, LGBTQ+, biological essentialiasm, and hate speech terms significantly pre-

dict misgendering tweets. Furthermore, tweets written by users with either listed pronouns (β=1.295147401, p<0.001) or at least one of the pride flags (β=0.274252286, p<0.01) present in their bio were significantly more likely to use gender-affirming pronouns. Lastly, tweets written by users with more expansive social networks (as measured by number of followers (β=0.535186954, p<0.001) and following (β=0.083737484, p<0.001)) were again significantly more likely to use gender-affirming pronouns.

Table 8.14: PRONOUN binary logistic regression for target celebrities in POST-COE condition. Separate data aggregated from equal number of tweet-level observations for trans-binary and trans-nonbinary celebrities, respectively.

	Trans-binary					Trans-nonbinary				
Variable	Estimate	sigma	t-statistic	p value	Sig.	Estimate	sigma	t-statistic	p value	Sig.
Intercept	1.91759944	0.04180581	45.8692122	<0.01	***	-0.189469764	0.01808674	-10.4756196	<0.01	***
Tweet characteristics										
Vader compound	0.10686288	0.02765654	3.8639281	<0.01	***	0.075053419	0.01498919	5.007171	<0.01	***
Likes	0.20799056	0.16636045	1.2502405	0.21		0.33183225	0.22213733	1.4938158	0.14	
Replies	-0.03024703	0.03968788	-0.7621225	0.45		-0.08052594	0.04161068	-1.9352228	0.05	
% affirming names	1.32883589	0.02035711	65.2762416	<0.01	***					
User characteristics										
Pronouns in bio	1.10367387	0.09827727	11.2302049	<0.01	***	1.295147401	0.0398894	32.4684612	<0.01	***
Flag in bio	0.63054446	0.24686576	2.5541997	0.01	*	0.274252286	0.0882783	3.1066784	<0.01	**
Followers	1.00872449	0.32946064	3.0617451	<0.01	**	0.535186954	0.09891463	5.4105946	<0.01	***
Following	0.0608671	0.04694862	1.2964619	0.19		0.083737484	0.02692641	3.1098648	<0.01	***
Lexical characteristics										
Trans terms	0.43258834	0.09399203	4.602394	<0.01	***	1.454819239	0.04937553	29.4643773	<0.01	***
Binary gender terms	-0.72769207	0.06195385	-11.7457125	<0.01	***	-0.931703764	0.06021494	-15.4729666	<0.01	***
Gender/sex terms	-0.90097379	0.14434656	-6.2417408	<0.01	***	0.402150152	0.08843967	4.5471695	<0.01	***
LGBTQ+ terms	-0.10918252	0.12875746	-0.8479704	0.40		-0.73855538	0.10753757	-6.8678825	<0.01	***
COE terms	0.54612187	0.05608507	9.737384	<0.01	***	0.375299092	0.03896724	9.6311447	<0.01	***
Biological terms	-0.62340956	0.12073804	-5.1633235	<0.01	***	-0.425722984	0.14436106	-2.9490153	<0.01	**
Hate speech terms	-0.54065085	0.20467209	-2.6415466	0.01	**	-0.487915408	0.08898243	-5.483278	<0.01	***
Pride terms	0.91314678	0.0840951	10.8585021	<0.01	***	0.156960211	0.04088543	3.8390259	<0.01	***

*** = p-value≤0.001; ** = p-value≤0.01; * = p-value≤0.05.

Table 8.15 displays the results from the PRONOUN binary logistic regression for the comparison celebrities, separated by gender identity. Explicitly running identical analyses on the cisgender comparison celebrities in this way helps confirm that the above results from the target group are not just artifacts from the filtering process. Meanwhile, the results from the Cox regression allow us to assess the extent to which a publicly-documented COE

influences the prevalence and lexical patterning of these language practices. To combine the cisgender celebrity data sets, I recorded the number of observations from the smaller data set (Holland, N=26,289) and randomly sampled the same number from the larger data set (Doja). I find that only two of the eight devised lexical categories (binary gender and pride/support) significantly predict whether a tweet will misgender or affirm the cisgender celebrity under discussion, while transgender and COE terms were also significant predictors for Cox.

Table 8.15: PRONOUN binary logistic regression for comparison celebrities across entire analysis period. Cisgender data aggregated from equal number of tweet-level observations for Doja and Holland.

	Cisgender					Transgender (Cox)				
Variable	Estimate	σ	t-statistic	p-value	Sig.	Estimate	σ	t-statistic	p-value	Sig.
Intercept	2.80685673	0.02909685	96.4660157	<0.01	***	2.98645984	0.0601559	49.64533218	<0.01	***
Tweet characteristics										
Vader compound	0.09917044	0.01966408	5.0432278	<0.01	***	0.23064974	0.04274314	5.39618106	<0.01	***
Likes	-0.04419116	0.01978701	-2.2333414	0.03	*	0.3281389	0.20304829	1.61606337	0.11	
Replies	0.13629956	0.09025561	1.5101507	0.13		-0.18359409	0.05611176	-3.27193599	<0.01	**
User characteristics										
Pronouns in bio	-0.32999631	0.03862558	-8.5434654	<0.01	***	0.6307902	0.13321704	4.73505663	<0.01	***
Flag in bio	0.13274082	0.12652395	1.049136	0.29		0.01067962	0.27108565	0.03939575	0.97	**
Followers	0.72693585	0.35509653	2.04715	0.04		0.82339668	0.52227689	1.57655203	0.11	
Following	0.29407724	0.06763837	4.3477872	<0.01	***	-0.05879762	0.02646524	-2.22169186	0.03	*
Lexical characteristics										
Trans terms	0.90396013	0.71701588	1.2607254	0.21		0.41072029	0.2039688	2.01364269	0.04	*
Binary gender terms	0.35855797	0.06483355	5.530439	<0.01	***	-0.45535585	0.13505659	-3.37159291	<0.01	***
Gender/sex terms	0.30270006	0.45933381	0.658998	0.51		0.0984495	0.46479164	0.21181427	0.83	
LGBTQ+ terms	-0.03920849	0.25956957	-0.151052	0.88		-0.44654638	0.47861384	-0.93299931	0.35	
COE terms	0.13257942	0.07137453	1.8575173	0.06		0.42141218	0.1507108	2.79616443	<0.01	**
Biological terms	-0.31621868	0.17581548	-1.7985827	0.07		-0.58232413	0.35730259	-1.62977863	0.10	
Hate speech terms	0.14408252	0.15774102	0.9134119	0.36		-0.32244556	0.43514765	-0.74100264	0.46	
Pride terms	0.6664182	0.07976184	8.355101	<0.01	***	0.55918634	0.12535067	4.4609759	<0.01	***

*** = p-value≤0.001; ** = p-value≤0.01; * = p-value≤0.05.

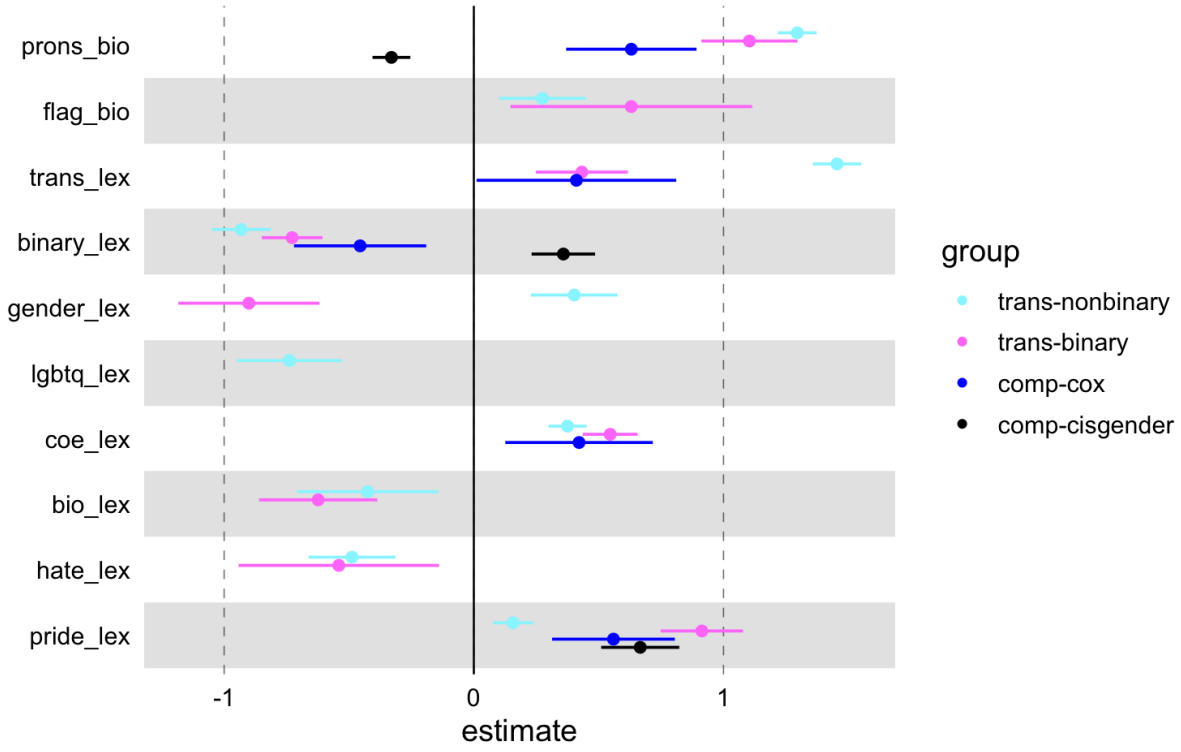
That binary gender and pride/support terms predict gender-affirming tweets for the cisgender celebrities is to be expected. Given that misgendering tweets in this case are generally mentioning entities other than the celebrities at hand, the significance of gender terms indicates discussion of these celebrities' (cis)gender identities and messages of support in tweets that actually discuss them. This latter fact perhaps explains the higher-than-expected

VADER discrepancy discussed above and the observation that the continuous VADER compound measure significantly predicted gender-affirming tweets. However, the significance of both the presence of pronouns and the number of following is somewhat inexplicable. One clue was unearthed through by-hand random sampling of misgendering tweets in the Doja corpus. I observed that tweets heuristically tagged as misgendering, in this case, frequently resulted from Twitter users' quoting lines from Doja Cat songs, which generally make heavy use of *he* suite pronouns and contain many expletives. Why users who list pronouns in their bio might be more likely to quote Doja's songs may relate to fandom demographics, as more loyal fans may be more likely to quote her lyrics. Future work could explore the demographic distribution of listed pronouns on Twitter and other social media platforms to shed light on this finding. Nonetheless, the fact that six of the eight binary lexical measures did not reach significance for the cisgender comparison group indicates that misgendering is rarely intentional for these celebrities and does not reflect principles of cisnormativity.

Results from the Cox model pattern more closely with the target celebrities than with the other comparison celebrities. For Cox, the strongest significant binary predictor of gender-affirming tweets was, like the trans-binary celebrities, the presence of pronouns in the user's bio (β=0.6307902, p<0.001) while the strongest binary predictor of misgendering tweets was the presence of binary gender terms in the tweet (β=-0.45535585, p<0.001). I additionally find that the presence of transgender, COE, and pride terms significantly predict gender-affirming tweets. Furthermore, tweets written by users with either listed pronouns (β=0.6307902, p<0.001) or one of the pride flags (β=0.01067962, p<0.01) present in their bio were significantly more likely to use gender-affirming pronouns.

Results from the distributional and content analyses reveal that considerable differences exist in gender-affirming reference by celebrity group and that misgendering/deadnaming

Figure 8.4: **Effect of binary measures on gendered pronominal usage for all celebrity groups. Positive coefficients represent tweets that were more likely to affirm, while negative coefficients represents those more likely to misgender. Only significant measures (p-value≤0.05) are displayed.**



tweets co-occur with ideologically-imbued lexical categories for trans celebrities but not for cisgender comparison celebrities. In the distributional analysis I show that users pronominally affirm the trans-nonbinary group (celebrities who use nonbinary *they)* at rates lower than they affirm the trans-binary group (who use the binary pronouns *she* or *he*). However, both of these groups are pronominally affirmed less than the comparison group – celebrities without publicly-documented coming-out events (COEs) who use binary pronouns (Figure 8.2. Additionally, these results demonstrate that deadnaming persists at a stable rate in around 17.5% of tweets post-COE (Figure 8.1) and that deadnaming and misgendering are

closely interrelated for the trans-binary celebrities (Table 8.8). In the content analysis I find significant correlations between misgendering/deadnaming and lexical terms representative of hate speech, biological essentialism, and binary gender terms across the target celebrity groups (Figure 8.4). Inversely, gender-affirming tweets significantly correlated with transgender, COE, and pride/support terms for target celebrities. As to be expected, only binary gender and pride/support terms significantly predicted gender-affirming pronominal usage for the cisgender comparison celebrities. Across both analyses, a general effect of historical time is observed, whereby Twitter users linguistically ratified the gender identities of target celebrities who came out chronologically later at higher rates (Table 8.4) but also utilized more negative sentiment in tweets that misgendered/deadnamed the later-COE celebrities (Tables 8.11 & 8.9). In the next chapter, I contextualize these quantitative results within the ideological framework of cisnormativity.

CHAPTER 9
DISCUSSION

These findings represent the first large-scale social media corpus analysis of nonbinary *they*, pronominal misgendering, and deadnaming. This study highlights the important progress that has been made; swaths of Twitter users, the large majority in some cases, began to use transgender celebrities' gender-affirming names and third-person pronouns immediately after their coming-out event (COE). Furthermore, discussion surrounding these celebrities in tweets that used gender-affirming language was more positive in sentiment, utilized language of support and pride, and in some cases appeared to be quite standard celebrity fare, mentioning appearances at award shows or new albums and TV roles. The transgender moment is here (Zimman, 2020) and, with it, the cultural emergence and linguistic recognition of a new wave of trans celebrities.

However, deadnaming and pronominal misgendering remain prevalent, stable, and malicious after these celebrities' presented their transgender identities. Twitter users affirm trans celebrities who list binary pronouns (Caitlyn Jenner and Elliot Page) at a somewhat lower rate than they affirm cisgender comparison celebrities who list the same pronouns (Doja Cat and Tom Holland). Further, Twitter users affirm trans celebrities who list nonbinary *they* at a much lower rate than than the trans-binary celebrities and at about half the rate of the comparison celebrities. The high affirming pronoun rate for Laverne Cox (95.52%), however, suggests that it is not just transgender identity itself that initiates such high rates of misgendering and deadnaming. Rather, it appears that the novel presentation of transgender identity from individuals who previously identified as cisgender men and women trigger

strong, even virulent reactions that are in part channeled into specific discursive practices. By comparing the distribution and lexical patterning of tweets discussing both transgender and cisgender celebrities, I present evidence that cisnormative ideologies attempt to prevent individuals perceived to transgress an immutable gender binary from full constitution in society (Zimman, 2020). Because being linguistically ratified as trans depends on the extent to which others affirm trans identity through the use of gender-affirming names and pronouns (Bucholtz & Hall, 2005; Zimman, 2017), the present findings demonstrate that trans identities – in this case, trans celebrities – are not yet completely and consistently ratified by this sample of Twitter users.

In the NAME analysis, I observe that deadnaming occurred at a high, stable rate immediately following the trans-binary celebrities' COE. Deadnaming patterns were remarkably similar across Jenner and Page in prevalence (80.63% and 84.3% of tweets, respectively) and in lexical content. For both celebrities, the practice of deadnaming was significantly correlated with lexical items that discussed their transition through binary-gendered language. Additionally, terms indicative of biological essentialism significantly predicted deadnaming tweets, as these celebrities' identities invoked discourses of the precept that the man-woman gender binary must correspond to the male-female sex binary among cisnormativity-enforcing Twitter users. This suggests that these users interpreted the celebrities' trans identities as violations of a coherent gender-sex correspondence. These patterns follow findings by Turton (2021), who demonstrated that commentators utilized biological sex terms in the context of deadnaming to distance Caitlyn Jenner's body away from a 'natural' cisnormative ideal. Meanwhile, I find terms that specifically discussed and celebrated the Page and Jenner's trans identities were significantly more likely to appear in tweets that used their gender-affirming names. This suggests that the mere use of transgender-related terms indexes a

more expansive gender ideology that accounts for and supports the transgender experiences of Caitlyn Jenner and Elliot Page.

Deadnaming already operates to make these celebrities' identities adhere to a cisnormative understanding of gender – one that relies on a fixed, stable, and natural gender binary. The significant correlations between deadnaming and binary-gender terms (man, woman, girl, etc.) suggest that these users find the act of changing one's name to be an explicit violation of the celebrities' former gender identities – perhaps, a violation of gender itself. In combination with cisnormative discourses, their past names are weaponized in the present to deny their these celebrities' the agency to assert their own genders. The following examples from the Page corpus illustrate these phenomena more concretely. Items in bold demarcate those picked up by the lexical category measures.

> "*Ellen Page is a* ***mentally ill woman*** *who* ***mutilated*** *her body. And this* ***disturbed*** *individual is a role model for the left.. I really sucks bearing witness to societal decay of this magnitude. God save America. #FuckYourPronouns*"

> "*There are only 2* ***genders****, only* ***women*** *can have kids..****men*** *who think they are* ***girls*** *need to see a shrink.. Ellen Page is still a* ***girl*** *she just needs* ***breast implants*** *now*"

The prevalence of humor terms in the fightin' word analysis (lol, 🤣, 😂, smh) also suggests that misgendering users treat these celebrities' trans identities as negligible and, indeed, laughable violations of cisnormativity. This finding extends work by Turton (2021) in demonstrating how deadnaming is used to illegitimize trans claims to self-definition through the rejection of self-determined name usage. Additionally, the presence of listed pronouns and trans/LGBTQ+ pride flags strongly predicted name-affirming tweets. This is unsurprising:

when an individual lists their pronouns, they promote an understanding that gender identity can not be assumed or assigned by others; rather, it is a personal expression that demands linguistic ratification in interaction. Similarly, the inclusion of a trans or LGBTQ+ pride flag suggests at the very least allyship or, more likely, membership within these communities. Conversely, deadnaming tweets appear to be driven by a Twitter population that less commonly expresses allyship with these communities and does not explicitly list the pronouns that in part constitute their gender identity, instead reifying cisnormativity through hate speech, a binary understanding of gender, and biological essentialism. This pattern provides evidence to the proposal that those who have the most to gain in preserving gendered power structures are more likely to reject trans experiences in the maintenance of cisnormativity. By choosing not to use the celebrities' gender-affirming names, these users withhold Page and Jenner's transness and relocate them to their proper place within an uncrossable gender binary.

Furthermore, deadnaming appears to at least partly consist of hate speech that is directly targeted at the celebrities themselves. Results from the fightin' words analysis show a strong association between TWITTER_HANDLE and deadnaming tweets for Jenner (-78.138), while for Page deadnaming tweets were highly correlated with both TWITTER_HANDLE (-52.909) and DEAD_HANDLE (-21.577). These users do not simply reject these celebrities' identities in their own, self-contained discourse but maliciously impose their denial directly onto the celebrities. In other words, deadnaming is intended to be heard by the celebrity being referenced. This is particularly concerning considering the fact that Twitter revised their blanket ban on 'dehumanizing speech' policy to include instances of 'targeted misgendering or deadnaming of transgender individuals' in October 2018 (Robertson, 2018). The use of a former name becomes harmful when it impacts transgender individuals who actively express

gender dissonance between a former name and their sense, experience, and expression of gender (Turton, 2021). Deadnaming carries not only symbolic power (Bourdieu, 1979) but is a truly injurious speech act; research has demonstrated that the usage of trans youth's gender-affirming names leads to better mental health outcomes (Russell et al., 2018). If this project alone can identify such a magnitude of deadnaming, much of it directly targeted at trans individuals and consisting of hate speech and biological essentialism, surely Twitter can take a more responsive approach to banning instances of deadnaming on the platform as part of its approach to limiting dehumanizing speech.

In the PRONOUN analysis, I observe that pronominal misgendering occurs at higher rates for trans celebrities who use nonbinary *they* exclusively (trans-nonbinary) than for trans celebrities who make use of at least one suite of binary pronouns (trans-binary). This study mirrors previous production work on nonbinary *they* by observing that, in discussions of trans individuals, *they* is used at much lower rates than the binary pronouns *she* or *he* when controlling for gender identity (Arnold et al., 2022). However, tweets discussing both trans groups trail behind the cisgender comparison celebrities in affirming pronoun usage. This in itself provides further evidence that social factors are motivating language users along the three-stage ongoing reconfiguration in English's pronominal system proposed by Konnelly and Cowper (2020). If all language users were at Stage 2, where masculine and feminine gender features remain fully contrastive but not obligatorily specified, equal rates of gender-affirming pronoun usage for the trans-binary and cisgender celebrities would be expected. Meanwhile, if all language users were at Stage 3 – at which point gender features have lost their contrastive and obligatory status – equal rates of gender-affirming pronoun usage across all analysis groups would be expected, but especially between the trans-nonbinary and trans-binary celebrities.

However, this is not the case. The data suggest that, while the majority of Twitter users remain at Stage 2, a small portion remain at Stage 1, where gender features are contrastive and cannot be respecified, as they readily use cisgender celebrities' affirming pronouns while either struggling to scrub their former lexical entries for Caitlyn Jenner and Elliot Page in instances of habitual or intentional misgendering. Meanwhile, an even greater portion of users fail to produce nonbinary *they* when referring to the celebrities who list such pronouns. Indeed, Twitter users appear to use trans-nonbinary celebrities' former pronouns (in acts of misgendering) at rates generally similar to their use of the celebrities' gender-affirming pronouns post-COE (Lovato: 37.1% *she*, 62.9% *they*; Smith: 54.28% *he*, 45.72% *they*). If the majority of language users were at Stage 3, where gender features are completely optional modifiers, we would expect to see much lower rates of misgendering binary pronoun usage. One possible explanation for this observation is the spread of COE information throughout a Twitter user's network: if a user was not aware of these celebrities COE, there is little chance they would use nonbinary *they* to refer to them. However, the fact that users pronominally affirmed Page's identity at the highest rate among the target group despite Page generating the smallest tweet corpus and having the least amount of Twitter followers (1.9m) suggest that this explanation is unlikely. Clearly, something is blocking the production of nonbinary *they*.

I argue that cisnormative ideologies inhibit Twitter users from affirming the pronouns of transgender celebrities and from advancing along the stages of the tripartite pronoun reconfiguration proposed by Konnelly and Cowper (2020). For transgender celebrities, instances of hate speech, biological essentialism, and the inclusion of binary gender terms strongly predicted misgendering tweets. This suggests that cisnormative-enforcing users who author misgendering tweets respond to expansive expressions of gender by reifying their under-

standing of an immutable man-woman binary. By using binary gendered terms in tweets that misgender trans celebrities, these users attempt to recodify trans identities and make them align with the gender/sex assigned to them at birth. Similarly, the presence of terms representative of biological essentialism indicates that these celebrities' identities are subjugated through cisnormative comparison to biologically 'natural' men and women. The significant presence of hate speech further indicates how these celebrities' COEs are interpreted as transgressive acts. However, the opposite effect of gender/sex terms suggest that the transgender identities of Page and Jenner appear explicitly at odds with cisnormative understandings of sex and gender, while the nonbinary identities of Smith and Lovato are ratified and affirmed through their rejection of the gender binary. Additionally, the presence of transgender identity terms was a much stronger predictor of gender-affirming tweets for the trans-nonbinary celebrities than for the trans-binary group. When considered in tandem with the observation of much higher rates of gender-affirming pronouns for the latter group, it appears that trans-binary gender-affirming tweets depend less on the actual nature of these celebrities' identities to nonetheless ratify them in the process of reference. In other words, trans-nonbinary gender affirming tweets depend more heavily on the presence of transgender identity and gender/sex terms more generally to successfully elicit nonbinary *they*.

Furthermore, it appears that as the dialogic ratification of transgender identity through the use of gender affirming third-person pronoun usage increases, the negative ideological and attitudinal charge behind instances of misgendering increases as well. Some evidence for this proposal comes from the VADER sentiment analysis, where we observe the highest discrepancy in sentiment between misgendering and gender-affirming tweets for Elliot Page, who also exhibits the highest affirming pronoun rate among the target celebrities. Meanwhile, Smith exhibits the lowest – an almost nonexistent – discrepancy measure and the

lowest affirming pronoun rate. Even with the reliability of this metric called into question by the somewhat large discrepancies observed for cisgender celebrities, the size of these effects (Smith=0.015, cisgender comparisons=0.084, Page=0.325) still indicates a meaningful pattern. Between the trans-nonbinary group, we also see that the sentiment of tweets discussing Smith in the PRE condition was much higher than for Lovato on average (0.292 vs. 0.151). Despite this indicating that Smith might be the more well-liked celebrity of the two, users affirmed Smith's pronouns at a considerably lower rate than Lovato's. Elsewhere, comparing effect sizes across the regression analyses for the target groups reveal greater effect sizes related to both hate/biological essentialism and COE/pride terms for the trans-binary celebrities than for the trans-nonbinary group, suggesting that greater pronominal gender-affirmation is simultaneously associated with more support and more hate – together, more overall recognition.

I observe an additional pattern of historical time within the respective target groups. Smith presented their nonbinary identity a little less than two years before Lovato did, while Jenner first presented her transgender identity a little over five years before Page did. At the same time, we observe higher affirming pronoun rates for Lovato than for Smith, and for Page than for Jenner. This complements other research demonstrating the increasing visibility of transgender as well as specifically nonbinary identities in society (Geiger & Graf, 2019; Zimman, 2020). However, it also appears that the recognition and more explicit cisnormative regulation of binary trans identities[1] currently surpasses that of nonbinary identities.

This study illustrates the partially differential treatment of binary and nonbinary celebrities within cisnormativity. Like other ideologies, cisnormativity is a totalizing vision: individuals that do not fit the organizing principle (in this case, a stable, natural, and coherent

[1]Here, this term comprises trans masculine and feminine individuals as well.

gender binary) must be modified or ignored (Gal & Irvine, 2019). To fit the cisnormative schema, users recodify the former identities of trans-binary celebrities (Jenner and Page) through deadnaming and more explicit instances of intentional misgendering. Though the trans-binary celebrities are aggressively deadnamed and there are greater signs of cisnormative regulation in pronominal misgendering tweets, Twitter users affirm their gender identity through pronouns at rates only slightly lower than comparison celebrities. Meanwhile, users seem to render nonbinary identities invisible through high rates of pronominal misgendering. Compared to the trans-binary celebrities, the trans-nonbinary results are more indicative of habitual misgendering with smaller lexical effect sizes for both negative (hate speech, biological essentialism) and positive (coming-out-event terms, expressions of pride and support) lexical categories. Furthermore, the tweets that do pronominally affirm the gender identity of these celebrities seem far more dependent on the specifically trans and nonbinary nature of their identities, as evidenced by the much larger positive effect of transgender lexical items. However, notions of gender change rapidly (Zimman, 2015), and we observe historical differences in the regulation of transgender identities within both the trans-binary and trans-nonbinary analysis groups.

Change in both the reconfiguration of English's pronominal system (Konnelly & Cowper, 2020) and in the language practices that support certain theories of gender (Conrod, 2020; Jones, 2021; Turton, 2021) are ongoing. Like other gender-inclusive language reforms (Bodine, 1975), these innovations are triggering broad discourse in society and seem to directly challenge 'crystallized' structures of cisnormativity (Borba, 2019). Ultimately, the different distributional and lexical patterns that emerge in this study between the four analysis groups (trans-binary, trans-nonbinary, comparison-cisgender, and comparison-cox) reveal in more detail the mechanisms by which trans identities are denied, dehumanized, and erased

through deadnaming and misgendering.

CHAPTER 10
CONCLUSION

10.1 Summary of findings

This study presents the first large-scale social media corpus analysis of natural instances of (dead)naming, (mis)gendering, and nonbinary *they* surrounding celebrities of diverse genders on Twitter. Expanding on previous work that has explored these practices from an ideological perspective (Conrod, 2020; Turton, 2021), I demonstrate how the deadnaming and pronominal misgendering of transgender celebrities is intimately connected to binary conceptions of gender and cisnormative power structures in English-speaking society. These findings align with previous work on gender-inclusive language reforms in other Western cultures (Bonnin & Coronel, 2021; Borba, 2019; Sendén et al., 2015).

I also provide strong evidence that rates of gender-affirming pronoun use in English are dependent on the gender identity, expression, and perception of the referent and that lower rates observed for transgender individuals may actually be improving over historical time. Drawing upon data consisting of hundreds of thousands of tweets, I find that Twitter users most readily affirm cisgender and transgender comparison celebrities without documented coming out events (COEs) – in this case, Doja Cat, Tom Holland, and Laverne Cox. Next, users pronominally affirm Caitlyn Jenner and Elliot Page, who have documented COEs and use binary pronouns – the trans-binary group – at rates slightly lower than that of the comparison group. Lastly, I show that users pronominally affirm Sam Smith and Demi Lovato, who have documented COEs and use nonbinary *they* – the trans-nonbinary group –

at rates considerably lower than the other two analysis groups.

Furthermore, results from Automated Dickey-Fuller (ADF) tests indicate that while there is a significant period of affirming pronoun uptake across the entire analysis period, there is no such effect in the time immediately following these celebrities' COEs. In other words, the rate at which tweets use the gender-affirming pronouns of the trans-binary and trans-nonbinary celebrities does not change significantly, in either direction, over the entire measured time period following the presentation of their transgender identities. An additional effect of chronological time is observed in both pairs of trans-binary and trans-nonbinary celebrities, as I show that the celebrity with the historically later COE also has a higher mean affirming pronoun rate post-COE.

Combining considerations from the content and distributional analysis, I argue that cisnormativity – a social and linguistic ideology that centers an immutable, coherent, and stable man-woman gender binary – inhibits language users from affirming the listed pronouns of Caitlyn Jenner, Elliot Page, Demi Lovato, and Sam Smith at similar rates to Cox, Doja, and Holland. Users who perceive these celebrities to transgress their notions, experience, and beliefs surrounding gender respond in turn by misgendering and deadnaming them at remarkably high rates. By deconstructing the mechanisms of cisnormativity implicated in these language practices – particularly, the presence of hate speech, relationship to biological essentialism, and reliance on binary gender identity terms – I hope to shed light on the ways in which cisnormativity functions as an organizing principle, recodifying trans expressions of gender back within an uncrossable gender binary and inhibiting Twitter users from advancing along the stages of an ongoing change in English's pronominal system.

10.2 Limitations

This study suffered from several limitations, many related to the nature of working with large amounts of social media data. For one, Twitter data is typically subject to a high level of noise – tweets that are largely irrelevant to the task at hand and hard to account for in processing and analysis. Though the extensive, purpose-oriented coreferential filtering process removed much of this noise, it is evident that a considerable amount remained. In part, this reflects the imperfect state of coreference resolution systems – particularly regarding their performance on social media text data and with natural instances of singular/nonbinary *they*. This noise biased certain analyses and led to some inexplicable and/or uninterpretable results.

Second, out-of-the-box computational linguistic tools – even those designed especially for social media text – tend to perform less well on large-scale Twitter corpora. As the amount of data approaches computational levels, it becomes more difficult to determine from what data certain results arise from. Finally, the high number of tweets required to elicit meaningful results limited my ability to compose a balanced comparison set that included a transgender man with no publicly documented COE. Additionally, it is worth re-stating the fact that there is no sufficient comparison for the high rates of deadnaming experienced by Elliot Page and Caitlyn Jenner in this data set. Addressing these latter two shortcomings would add more empirical validity to my results.

10.3 Future work

There are several promising directions for future work. First, I plan to make anonymize the user- and tweet-level data and make the entire filtered corpus publicly available. I truly believe the analyses conducted here represent the so-called tip of the iceberg – a corpus of this

size yet with such specific focus certainly provides scholars of language, gender, and sexuality a great deal of meaningful data to work with. Second, future studies could analyze naturalistic productions of gendered pronouns and (dead)naming in other contexts to identify the extent to which they mirror the patterns observed here. Similarly, there are many other possible linguistic manifestations of cisnormativity that researchers could pursue in computational, sociolinguistic, and psycholinguistic work to continue identifying and dismantling its mechanisms. Third, there is practically no research on the social and pragmatic factors that drive individuals to list their pronouns (cf. Jones, 2021). Whether as part of in-person introductions or in a user's Twitter bio, future work should explore the relationship between gender identity/ideology and the act of listing pronouns. Finally, the relationship between an individual's linguistic behavior in computer-mediated communication and their speech patterns is extremely understudied. The two domains variably lend themselves to influences from the many dimensions of communication – future work could begin to untangle how speech and text relate, especially in the context of identity-driven language practice.

10.4 Ethical considerations

In the pursuit of understanding how misgendering and deadnaming enforce cisnormativity, it is unfortunate that this paper recirculates the broad attitudes driving these language practices. However, I hope it is evident that any such discussion of misgendering and deadnaming is quotational – a mention, not a use – and that the insight this paper offers into identifying how these practices are operationalized makes up for the fact that it must reprint them.

REFERENCES

Ackerman, L. (2018). *Our words matter: acceptability, grammaticality, and ethics of research on singular 'they'-type pronouns* (Unpublished doctoral dissertation). Newcastle University.

Ackerman, L. (2019). Syntactic and cognitive issues in investigating gendered coreference. *Glossa*, *4*, 1-27. doi: 10.5334/GJGL.721

Ackerman, L., Riches, N., & Wallenberg, J. (2018). Coreference dependency formation is modulated by experience with variation of human gender. In *Annual meeting of the linguistic society of america.*

Ansara, Y. G., & Hegarty, P. (2014). Methodologies of misgendering: Recommendations for reducing cisgenderism in psychological research. *Feminism and Psychology*, *24*, 259-270. doi: 10.1177/0959353514526217

Arnold, J. E., Marquez, A., Li, J., & Franck, G. (2022, 2). Does nonbinary they inherit the binary pronoun production system? *Glossa Psycholinguistics*, *2*, 1-14. Retrieved from `https://escholarship.org/uc/item/1x84d0dh` doi: 10.5070/G601183

Balhorn, M. (2004, 6). The rise of epicene they. *Journal of English Linguistics*, *32*, 79-104. doi: 10.1177/0075424204265824

Bamman, D., Eisenstein, J., & Schnoebelen, T. (2014). Gender identity and lexical variation in social media. *Journal of Sociolinguistics*, *18*(2), 135–160.

Baron, D. E. (1981). The epicene pronoun: The word that failed. *American Speech*, *56*(2), 83–97.

Baron, N. S. (1971). A reanalysis of english grammatical gender. *Lingua*, *27*, 113–140.

Bauer, G. R., Hammond, R., Travers, R., Kaay, M., Hohenadel, K. M., & Boyce, M. (2009). "i don't think this is theoretical; this is our lives": how erasure impacts health care for transgender people. *Journal of the Association of Nurses in AIDS Care*, *20*(5), 348–361.

Bjorkman, B. M. (2017). Singular they and the syntactic representation of gender in english. *Glossa*, *2*. doi: 10.5334/gjgl.374

Blodgett, S. L., Green, L., & O'Connor, B. (2016). Demographic dialectal variation in social media: A case study of african-american english. *arXiv preprint arXiv:1608.08868*.

Bodine, A. (1975). Androcentrism in prescriptive grammar: singular 'they', sex-indefinite 'he', and 'he or she". *Language in Society*, *4*, 129-175. Retrieved from `https://about.jstor.org/terms`

Bonnin, J. E., & Coronel, A. A. (2021, 3). Attitudes toward gender-neutral spanish: Acceptability and adoptability. *Frontiers in Sociology*, *6*. doi: 10.3389/fsoc.2021.629616

Borba, R. (2019). Gendered politics of enmity: Language ideologies and social polarisation in brazil. *Gender and Language*, *13*, 423-448. doi: 10.1558/genl.38416

Borba, R., & Milani, T. M. (2017, 9). The banality of evil: Crystallised structures of cisnormativity and tactics of resistance in a brazilian gender clinic. *Journal of Language and Discrimination*, *1*, 7-33. doi: 10.1558/jld.33354

Bourdieu, P. (1979). Symbolic power. *Critique of Anthropology*, *4*, 77-85.

Bradley, E. D. (2020, 3). The influence of linguistic and social attitudes on grammaticality judgments of singular 'they'. *Language Sciences*, *78*. doi: 10.1016/j.langsci.2020.101272

Bradley, E. D., Salkind, J., Moore, A., & Teitsort, S. (2019, 3). Singular 'they' and novel pronouns: gender-neutral, nonbinary, or both? *Proceedings of the Linguistic Society*

of America, *4*, 1-7. doi: 10.3765/plsa.v4i1.4542

Breitfeller, L., Ahn, E., Jurgens, D., & Tsvetkov, Y. (2019). Finding microaggressions in the wild: A case for locating elusive phenomena in social media posts. In *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (emnlp-ijcnlp)* (pp. 1664–1674).

Buch, L. A. (2017). *The effect of gender neutral pronouns on non-binary gender identity experience* (Unpublished doctoral dissertation). Northcentral University.

Bucholtz, M., & Hall, K. (2004). Theorizing identity in language and sexuality research. *Language in society*, *33*(4), 469–515.

Bucholtz, M., & Hall, K. (2005, 8). Identity and interaction: A sociocultural linguistic approach. *Discourse Studies*, *7*, 585-614. doi: 10.1177/1461445605054407

Butler, J. (1990). *Gender trouble*. Routledge.

Calder, J. (2020). Language, gender and sexuality in 2019: Interrogating normativities in the field. *Gender and Language*, *14*, 429-454. doi: 10.1558/GENL.18634

Cameron, D., & Kulick, D. (2003). *Language and sexuality*. Cambridge University Press.

Camilliere, S., Izes, A., Leventhal, O., & Grodner, D. J. (2021). They is changing: Pragmatic and grammatical factors that license singular they. In *Proceedings of the annual meeting of the cognitive science society* (p. 1542-1548).

Cannoot, P., & Decoster, M. (2020). The abolition of sex/gender registration in the age of gender self-determination: An interdisciplinary, queer, feminist and human rights analysis. *International Journal of Gender, Sexuality and Law*, *1*(1).

Cao, Y. T., & III, H. D. (2021). Toward gender-inclusive coreference resolution: An analysis of gender and bias throughout the machine learning lifecycle. *Computational*

Linguistics, *47*, 615-661. Retrieved from https://doi.org/10.1162/COLI doi: 10.1162/COLI

Clark, K., & Manning, C. D. (2016). Deep reinforcement learning for mention-ranking coreference models. *arXiv preprint arXiv:1609.08667*.

Conrod, K. (2017). Names before pronouns: Variation in pronominal reference and gender. In *Northwest linguistics conference, university of british columbia.*

Conrod, K. (2018a). Changes in singular they. In *Cascadia workshop in sociolinguistics.*

Conrod, K. (2018b, 10). Pronouns and misgendering. In *New ways of analyzing variation.*

Conrod, K. (2019). *Pronouns raising and emerging* (Unpublished doctoral dissertation). University of Washington.

Conrod, K. (2020, 9). Pronouns and gender in language. In K. Hall & R. Barrett (Eds.), . Oxford University Press. doi: 10.1093/oxfordhb/9780190212926.013.63

Corbett, G. G. (1991). *Gender*. Cambridge University Press.

Curvy, D. (2021). Doja cat's new dimension. *V Magazine*. Retrieved 2022-03-15, from https://vmagazine.com/article/doja-cats-new-dimension/

Curzan, A. (2003). *Gender shifts in the history of english*. Cambridge University Press.

Darwin, H. (2017, 8). Doing gender beyond the binary: A virtual ethnography. *Symbolic Interaction*, *40*, 317-334. doi: 10.1002/symb.316

Davenport, C. (2020). *The use of themself to refer to a singular antecedent* (Undergraduate Thesis). Retrieved from https://scholarworks.wm.edu/honorstheses/1538

Delphy, C. (1993). Rethinking sex and gender. In *Women's studies international forum* (Vol. 16, pp. 1–9).

Demszky, D., Garg, N., Voigt, R., Zou, J., Gentzkow, M., Shapiro, J., & Jurafsky, D. (2019). Analyzing polarization in social media: Method and application to tweets on 21 mass

shootings. *arXiv preprint arXiv:1904.01596*.

Dickey, D. A., & Fuller, W. A. (1979, 6). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*, 427-431. doi: 10.1080/01621459.1979.10482531

Doherty, A., & Conklin, K. (2017, 4). How gender-expectancy affects the processing of "them". *Quarterly Journal of Experimental Psychology*, *70*, 718-735. doi: 10.1080/17470218.2016.1154582

Doyle, G. (2014). Mapping dialectal variation by querying social media. In *Proceedings of the 14th conference of the european chapter of the association for computational linguistics* (pp. 98–106).

Eckert, P. (1989). The whole woman: Sex and gender differences in variation. *Language variation and change*, *1*(3), 245–267.

Eisenstein, J., O'Connor, B., Smith, N. A., & Xing, E. P. (2014). Diffusion of lexical change in social media. *PloS one*, *9*(11), e113114.

Ericsson, S. (2018). The language of cisnormativity: Children and parents in interaction with a multimodal app. *Gender and Language*, *12*, 139-167. doi: 10.1558/genl.31827

Ericsson, S. (2021). Gender and sexuality normativities: Using conversation analysis to investigate heteronormativity and cisnormativity in interaction. In J. Angouri & J. Baxter (Eds.), .

Estival, D., Gaustad, T., Pham, S. B., Radford, W., & Hutchinson, B. (2007). Author profiling for english emails. In *Proceedings of the 10th conference of the pacific association for computational linguistics* (Vol. 263, p. 272).

Foertsch, J., & Gernsbacher, M. A. (1997). In search of gender neutrality: Is singular they a cognitively efficient substitute for generic he? *Psychological science*, *8*(2), 106–111.

Gal, S. (1989). Between speech and silence: The problematics of research on language and gender. *IPrA Papers in Pragmatics*, *3*(1), 1–38.

Gal, S., & Irvine, J. T. (2019). *Signs of difference: Language and ideology in social life*. Cambridge University Press.

Geiger, A. W., & Graf, N. (2019, 9). *About one-in-five u.s. adults know someone who goes by a gender-neutral pronoun* (Tech. Rep.). Pew Research Center. Retrieved from `https://pewrsr.ch/2LmF98V`

GLAAD. (2021). Elliot page, oscar-nominated star of umbrella academy, speaks out about being transgender. *glaad*. Retrieved 2022-03-15, from `https://www.glaad.org/blog/elliot-page-oscar-nominated-star-umbrella-academy-speaks-out-about-being-transgender`

Goel, R., Soni, S., Goyal, N., Paparrizos, J., Wallach, H., Diaz, F., & Eisenstein, J. (2016). The social dynamics of language change in online networks. In *International conference on social informatics* (pp. 41–57).

Hall, K. (1995). Lip service on the fantasy lines. In K. Hall & M. Bucholtz (Eds.), (pp. 183–216).

Hall, K., Borba, R., & Hiramoto, M. (2021). Relocating power: The feminist potency of language, gender and sexuality research. *Gender and Language*, *15*, 1-10. doi: 10.1558/genl.19523

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, *10*, 252-264.

Hekanaho, L. (2020). *Generic and nonbinary pronouns: Usage acceptability and attitudes* (Unpublished doctoral dissertation). University of Helsinki.

Hernandez, E. E. (2020). *Pronouns, prescriptivism, and prejudice: Attitudes toward the*

singular 'they', prescriptive grammar, and nonbinary transgender people (Unpublished master's thesis). Purdue University.

Hobaica, S., Schofield, K., & Kwon, P. (2019). "here's your anatomy… good luck": transgender individuals in cisnormative sex education. *American Journal of Sexuality Education*, *14*(3), 358–387.

Honnibal, M., & Montani, I. (2017). *spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing.* (To appear)

Hord, L. C. R. (2016). Bucking the linguistic binary: Gender neutral language in english, swedish, french, and german. *Proceedings of Western Interdisciplinary Student Symposium on Language Research*, *3*. Retrieved from `https://ir.lib.uwo.ca/index.php/wpl_clw`

Hornscheidt, L. (2015). Trans_x_ing linguistic actions and linguistics. *Könskonstruktioner och språkförändringar: Rapport från den åttonde nordiska konferensen om språk och kön. Huddinge: Södertörn University*, 29–46.

Hutto, C. J., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the eighth international aaai conference on weblogs and social media* (p. 216-225). Retrieved from `http://sentic.net/`

Ilbury, C. (2020). 'sassy queens': Stylistic orthographic variation in twitter and the enregisterment of aave. *Journal of Sociolinguistics*, *24*(2), 245–264.

Irvine, J. T., & Gal, S. (2000). Language ideology and linguistic differentiation. In P. V. Kroskrity (Ed.), (pp. 35–83). School of American Research Press.

Johnson, K., Auerswald, C., LeBlanc, A. J., & Bockting, W. O. (2019). 7. invalidation experiences and protective factors among non-binary adolescents. *Journal of Adolescent Health*, *64*(2), S4.

Jones, J. J. (2021). A dataset for the study of identity at scale: Annual prevalence of american twitter users with specified token in their profile bio 2015–2020. *PloS one*, *16*(11), e0260185.

Konnelly, L., & Cowper, E. (2020). Gender diversity and morphosyntax: An account of singular they. *Glossa*, *5*, 1-19. doi: 10.5334/GJGL.1000

Kosnick, K. (2019, 4). The everyday poetics of gender-inclusive french: Strategies for navigating the linguistic landscape. *Modern and Contemporary France*, *27*, 147-161. doi: 10.1080/09639489.2019.1588869

Koyama, E. (2003). The transfeminist manifesto. In R. Dicker & A. Piepmeier (Eds.), (p. 244-259). Northeastern University Press. Retrieved from `www.eminism.org`

Kramarae, C. (1981). *Women and men speaking: Frameworks for analysis.* ERIC.

Labov, W. (1966). Hypercorrection by the lower middle class as a factor in linguistic change. *Sociolinguistics*, 84–113.

Lakoff, R. (1973). Language and woman's place. *Language in society*, *2*(1), 45–79.

Larson, B. N. (2017). Gender as a variable in natural-language processing: Ethical considerations. In *Proceedings of the first worksop on ethics in natural language processing* (p. 1-11). Association for Computational Linguistics.

Lascotte, D. K. (2016, 2). Singular they: An empirical study of generic pronoun use. *American Speech*, *91*, 62-80. doi: 10.1215/00031283-3509469

Lewis, S. (2021). Instagram launches feature to let users add pronouns to profile. *CBS News*. Retrieved 2022-03-15, from `https://www.cbsnews.com/news/instagram-pronouns-profiles/`

Lindqvist, A., Gustafsson Sendén, M., & Bäck, E. A. (2016). Vem tycker om hen? *Språk och stil*, *26*, 101–129.

Lovato, D. (2021). Every day we wake up, we are given another opportunity... *Twitter*. Retrieved 2022-03-18, from `https://twitter.com/ddlovato/status/1394913210856185858` (@ddlovato)

McConnell-Ginet, S. (2003). 'what's in a name?' social labeling and gender practices. In (p. 69-97).

McConnell-Ginet, S. (2014). Gender and its relation to sex: The myth of 'natural' gender. In G. G. Corbett (Ed.), (Vol. 6, p. 3-39).

McElhinny, B. (2003). Theorizing gender in sociolinguistics and linguistic anthropology. *The handbook of language and gender*, 21–42.

McInroy, L. B., & Craig, S. L. (2015). Transgender representation in offline and online media: Lgbtq youth perspectives. *Journal of Human Behavior in the Social Environment*, *25*(6), 606–617.

McLemore, K. A. (2015, 1). Experiences with misgendering: Identity misclassification of transgender spectrum individuals. *Self and Identity*, *14*, 51-74. doi: 10.1080/15298868.2014.950691

Mendelsohn, J., Tsvetkov, Y., & Jurafsky, D. (2020, 8). A framework for the computational linguistic analysis of dehumanization. *Frontiers in Artificial Intelligence*, *3*. doi: 10.3389/frai.2020.00055

Monroe, B. L., Colaresi, M. P., & Quinn, K. M. (2008). Fightin' words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Analysis*, *16*, 372-403. doi: 10.1093/pan/mpn018

Morgenroth, T., & Ryan, M. K. (2018, 7). Gender trouble in social psychology: How can butler's work inform experimental social psychologists' conceptualization of gender? *Frontiers in Psychology*, *9*, 1-9. doi: 10.3389/fpsyg.2018.01320

Morton, T. S. (2000). Coreference for nlp applications. In *Proceedings of the 38th annual meeting of the association for computational linguistics* (pp. 173–180).

Nguyen, D., Doğruöz, A. S., Rosé, C. P., & de Jong, F. (2016, 9). Computational sociolinguistics: A survey. *Computational Linguistics*, *42*, 537-593. doi: 10.1162/COLI_a_00258

Nguyen, D., Gravel, R., Trieschnigg, D., & Meder, T. (2013). 'how old do you think i am?' a study of language and age in twitter. In *Proceedings of the international aaai conference on web and social media* (Vol. 7).

Noll, J., Lowry, M., & Bryant, J. (2018, 10). Changes over time in the comprehension of he and they as epicene pronouns. *Journal of Psycholinguistic Research*, *47*, 1057-1068. doi: 10.1007/s10936-018-9577-4

Olson, K. R., Durwood, L., Demeules, M., & McLaughlin, K. A. (2016, 3). Mental health of transgender children who are supported in their identities. *Pediatrics*, *137*. doi: 10.1542/peds.2015-3223

Page, E. (2020). Hi friends, i want to share with you that i am trans... *Twitter*. Retrieved 2022-03-18, from `https://twitter.com/theelliotpage/status/1333820783655837701` (@theelliotpage)

Papadopoulos, B. (2018). *Morphological gender innovations in spanish of genderqueer speakers* (Undergraduate Thesis). University of California, Berkely.

Parks, J. B., & Roberton, M. A. (1998). Contemporary arguments against nonsexist language: Blaubergs (1980) revisited. *Sex roles*, *39*(5), 445–461.

Popič, D., & Gorjanc, V. (2018, 12). Challenges of adopting gender–inclusive language in slovene. *Suvremena Lingvistika*, *44*, 329-350. doi: 10.22210/suvlin.2018.086.07

Prabhakaran, V., Reid, E. E., & Rambow, O. (2014). Gender and power: How gender and gender environment affect manifestations of power. In *Proceedings of the 2014 confer-*

ence on empirical methods in natural language processing (emnlp) (pp. 1965–1976).

Rangel, F., & Rosso, P. (2013). Use of language and author profiling: Identification of gender and age. *Natural Language Processing and Cognitive Science*, *177*.

Rao, D., Yarowsky, D., Shreevats, A., & Gupta, M. (2010). Classifying latent user attributes in twitter. In *Proceedings of the 2nd international workshop on search and mining user-generated contents* (pp. 37–44).

Ring, T. (2018). Tucker carlson says there's no transgender community. *The Advocate*. Retrieved 2022-03-18, from `https://www.advocate.com/media/2018/4/04/tucker-carlson-says-theres-no-transgender-community`

Robertson, A. (2018, 11). Twitter has banned misgendering or 'deadnaming' transgender people. *The Verge*.

Russell, S. T., Pollitt, A. M., Li, G., & Grossman, A. H. (2018, 10). Chosen name use is linked to reduced depressive symptoms, suicidal ideation, and suicidal behavior among transgender youth. *Journal of Adolescent Health*, *63*, 503-505. doi: 10.1016/j.jadohealth.2018.02.003

Sanford, A. J., & Filik, R. (2007, 2). 'they' as a gender-unspecified singular pronoun: Eye tracking reveals a processing cost. *Quarterly Journal of Experimental Psychology*, *60*, 171-178. doi: 10.1080/17470210600973390

Schaefer, S. L. (2013). *A progressive prescription: Epicene pronouns and feminism in the chicago manual of style* (Unpublished master's thesis). University of Wisconsin at Oshkosh.

Sendén, M. G., Bäck, E. A., & Lindqvist, A. (2015, 7). Introducing a gender-neutral pronoun in a natural gender language: The influence of time on attitudes and behavior. *Frontiers in Psychology*, *6*. doi: 10.3389/fpsyg.2015.00893

Sheydaei, I. (2021). Gender identity and nonbinary pronoun use: Exploring reference strategies for referents of unknown gender. *Gender and Language*, *15*, 369-393. doi: 10.1558/genl.18871

Silverstein, M. (1985). Language and the culture of gender: At the intersection of structure, usage, and ideology. In *Semiotic mediation* (pp. 219–259). Elsevier.

Simpson, L., & Dewaele, J. M. (2019, 3). Self-misgendering among multilingual transgender speakers. *International Journal of the Sociology of Language*, *2019*, 103-128. doi: 10.1515/ijsl-2018-2014

Sinclair-Palm, J. (2017). 'it's non-existent': Haunting in trans youth narratives about naming. *Occasional Paper Series*, *2017*, 1-13.

Sluchinski, K. (2019). Genderless narratives: The pragmatics of ta in chinese social media. In *The 2019 annual conference of the canadian linguistic association: 2019 congress of the humanities and social sciences. http://cla-acl. ca/actes-2019-proceedings.*

Smith, S. (2019). Today is a good day so here goes. i've decided i am... *Twitter*. Retrieved 2022-03-18, from `https://twitter.com/samsmith/status/1172519872464662530` (@samsmith)

Stryker, S. (2008). Transgender history, homonormativity, and disciplinarity. *Radical History Review*, *2008*(100), 145–157.

Turton, S. (2021). Deadnaming as disformative utterance: The redefinition of trans womanhood on urban dictionary. *Gender and Language*, *15*, 42-64. doi: 10.1558/genl.18816

Valentine, D. (2007). *Imagining transgender*. Duke University Press.

Valentine, S. E., & Shipherd, J. C. (2018, 12). A systematic review of social stress and mental health among transgender and gender non-conforming people in the united states. *Clinical Psychology Review*, *66*, 24-38. doi: 10.1016/j.cpr.2018.03.003

Vergoossen, H. P., Renström, E. A., Lindqvist, A., & Sendén, M. G. (2020, 9). Four dimensions of criticism against gender-fair language. *Sex Roles*, *83*, 328-337. doi: 10.1007/s11199-019-01108-x

Verguiero, V. (2016). *Por inflexões decoloniais de corpos e identidades degênero inconformes: Uma análise autoetnográfica dacisgeneridade como normatividade* (Undergraduate Thesis). Universidade Federal da Bahia.

Vincent, B., & Manzano, A. (2017). History and cultural diversity. In *Genderqueer and non-binary genders* (pp. 11–30). Springer.

Voigt, R., Camp, N. P., Prabhakaran, V., Hamilton, W. L., Hetey, R. C., Griffiths, C. M., … Eberhardt, J. L. (2017). Language from police body camera footage shows racial disparities in officer respect. *Proceedings of the National Academy of Sciences*, *114*(25), 6521–6526.

Wales, K., & Katie, W. (1996). *Personal pronouns in present-day english*. Cambridge University Press.

Webster, K., Recasens, M., Axelrod, V., & Baldridge, J. (2018). Mind the gap: A balanced corpus of gendered ambiguous pronouns. In L. Zettlemoyer (Ed.), *Transactions of the association for computational linguistics* (p. 605-618). Retrieved from `http://direct.mit.edu/tacl/article-pdf/doi/10.1162/tacl_a_00240/1567674/tacl_a_00240.pdf` doi: 10.1162/tacl_a_00240/1567674/tacl_a_00240.pdf

Zappavigna, M. (2012). *Discourse of twitter and social media: How we use language to create affiliation on the web* (Vol. 6). A&C Black.

Zimman, L. (2009). 'the other kind of coming out': Transgender people and the coming out narrative genre. *Gender & Language*, *3*(1).

Zimman, L. (2014). The discursive construction of sex: Remaking and reclaiming the gendered body in talk about genitals among trans men. In J. Davis & J. Raclaw (Eds.), . Oxford University Press.

Zimman, L. (2015). Facebook, the gender binary, and third-person pronouns. In A. Northover (Ed.), (p. 137-142). Oxford University Press.

Zimman, L. (2017). Transgender language reform. *Journal of Language and Discrimination*, *1*, 84-105. doi: 10.1558/jld.33139

Zimman, L. (2019). Trans self-identification and the language of neoliberal selfhood: Agency, power, and the limits of monologic discourse. *International Journal of the Sociology of Language*, *2019*, 147-175. doi: 10.1515/ijsl-2018-2016

Zimman, L. (2020). Transgender language, transgender moment: Toward a trans linguistics. In K. Hall & R. Barrett (Eds.), . Oxford University Press. doi: 10.1093/oxfordhb/9780190212926.013.45

TRANSGRESSING THE BINARY: GENDERED LANGUAGE PRACTICES ON TWITTER

Approved by:

Rob Voigt
Department of Linguistics
Northwestern University

Annette D'Onofrio
Department of Linguistics
Northwestern University

Date Approved: May 11, 2022

APPENDIX A
TERMS USED IN ALTENT FILTER

Tom Holland 'Zendaya'

Doja Cat 'The Weeknd'

Laverne Cox 'John Legend'

Caitlyn Jenner 'Annie Leibovitz', 'Kris Jenner', 'Kardashian', 'Kim Kardashian', 'Kylie Jenner', 'Denrele Edun'

Elliot Page 'Heath Ledger', 'Chris Pratt', 'Desmond Elliot'

Sam Smith 'Lady Gaga', 'Demi Lovato', 'SZA'

Demi Lovato 'Ariana Grande', 'Neil DeGrasse Tyson', 'G-Eazy'

APPENDIX B

TWITTER BIO FIELD DEMOGRAPHIC DATA

Table B.1: User demographic data for target celebrities, ordered chronologically by time of COE.

Celebrity	**Jenner**			**Smith**			**Lovato**			**Page**		
Users with pronouns in bio	5.37%			12.41%			18.51%			35.80%		
Users with flags in bio	0.87%			2.30%			2.62%			5.39%		
	Type	Count	Frequency	Type	Count	Frequency	Type	Count	Frequency	Type	Count	Frequency
	she	31825	59.59%	she	20757	54.02%	she	28276	54.26%	she	21764	37.84%
	he	10984	20.57%	he	7662	19.94%	he	9508	18.25%	he	12375	21.52%
	she/they	4477	8.38%	she/they	3441	8.95%	she/they	5362	10.29%	she/they	7496	13.03%
	they	3298	6.18%	they	3077	8.01%	they	3648	7.00%	they	7120	12.38%
	he/they	1675	3.14%	he/they	1935	5.04%	he/they	2543	4.88%	he/they	5008	8.71%
	any	384	0.72%	any	670	1.74%	any	1224	2.35%	any	1567	2.72%
	he/she/they	340	0.64%	he/she/they	473	1.23%	he/she/they	701	1.35%	he/she/they	1065	1.85%
	he/she	319	0.60%	he/she	258	0.67%	he/she	461	0.88%	he/she	582	1.01%
	they/ze	25	0.05%	he/they/xe	39	0.10%	they/xe	80	0.15%	he/they/xe	119	0.21%
	they/xe	15	0.03%	they/xe	29	0.08%	he/they/xe	73	0.14%	they/xe	112	0.19%
	he/they/xe	11	0.02%	he/xe	22	0.06%	he/xe	55	0.11%	he/xe	83	0.14%
	he/ze	9	0.02%	she/they/xe	9	0.02%	she/they/xe	43	0.08%	xe	31	0.05%
	she/ze	9	0.02%	they/ze	7	0.02%	she/xe	32	0.06%	she/they/xe	30	0.05%
	she/xe	7	0.01%	xe	7	0.02%	xe	21	0.04%	she/xe	26	0.05%
	ze	7	0.01%	she/xe	6	0.02%	he/she/xe	13	0.02%	they/ze	25	0.04%
Total distinct users	**53408**			**38426**			**52112**			**57513**		

APPENDIX C

ADDITIONAL FIGHTIN' WORDS LEXICAL ANALYSES

Table C.1: PRONOUN lexical associations for Elliot Page post-COE.

MISGENDER		PRON_AFFIRM	
Score	Term	Score	Term
-30.391	WRONG_PRONOUN	19.116	CELEB_NAME
-15.896	CELEB_DEADNAME	14.092	transgender
-12.682	TWITTER_HANDLE	13.413	announces
-8.53	you	13.245	star
-7.565	a	11.445	TARGET_PRONOUN
-6.621	man	9.496	academy
-6.516	are	9.308	umbrella
-5.713	woman	8.966	happy
-5.546	not	7.775	for
-5.475	declares	7.093	nominated
-5.428	actress	6.858	oscar
-4.952	or	6.826	actor
-4.94	do	6.628	coming
-4.542	n't	6.511	love
-4.319	female	6.319	out
-4.128	named	5.611	proud
-4.111	it	5.566	juno
-4.079	DEAD_HANDLE	5.339	top
-4.016	re	5.307	first
-3.998	be	5.173	trunks
-3.863	no	5.154	swim
-3.844	someone	4.949	so
-3.832	if	4.837	joy
-3.824	n't	4.831	shares
-3.823	that	4.786	surgery
-3.783	pronouns	4.615	since
-3.719	were	4.575	i
-3.714	does	4.448	abs
-3.687	your	4.414	support
-3.539	sex	4.41	after

Note: correlation scores above ± 1.96 are considered significant (Monroe et al., 2008).

Table C.2: PRONOUN lexical associations for Demi Lovato post-COE.

MISGENDER		PRON_AFFIRM	
Score	Term	Score	Term
-33.909	WRONG_PRONOUN	42.53	TARGET_PRONOUN
-6.114	oin	12.136	pronouns
-5.175	's	11.921	non
-5.152	TWITTER_HANDLE	11.877	binary
-4.719	was	9.954	changes
-4.501	kill	7.942	nonbinary
-4.5	hacked	7.525	announces
-3.995	attention	7.206	misgender
-3.843	is	7.026	are
-3.829	attempting	6.759	as
-3.709	into	6.279	spider
-3.677	has	5.989	tattoo
-3.674	needs	5.981	changing
-3.629	s	5.86	CELEB_NAME
-3.615	met	5.67	out
-3.593	night	5.462	comes
-3.534	last	5.337	eazy
-3.407	a	5.303	ago
-3.321	demonic	5.211	posted
-3.232	demon	5.207	sober
-3.216	does	5.095	california
-3.169	drugs	5.055	today
-3.136	lost	4.749	singer
-3.129	woman	4.598	re
-2.954	me	4.576	misgendered
-2.941	heroin	4.44	picture
-2.742	girl	4.364	longer
-2.722	spoon	4.166	're
-2.707	when	4.156	identifies
-2.611	wants	4.154	misgendering

Note: correlation scores above ± 1.96 are considered significant (Monroe et al., 2008).

Table C.3: PRONOUN lexical associations for Sam Smith post-COE.

MISGENDER		PRON_AFFIRM	
Score	Term	Score	Term
-37.707	WRONG_PRONOUN	40.956	TARGET_PRONOUN
-6.173	philippines	14.557	pronouns
-5.866	voice	8.069	changes
-4.859	's	8.007	are
-4.551	was	5.105	tested
-4.441	man	4.968	re
-4.292	is	4.705	coronavirus
-4.213	noticed	4.437	respect
-4.007	filipino	4.424	non
-3.978	a	4.406	binary
-3.961	kid	3.908	nonbinary
-3.817	blind	3.857	changing
-3.815	singing	3.785	're
-3.362	sings	3.769	#imready
-3.239	from	3.716	misgendered
-3.217	mansion	3.552	use
-2.962	sounds	3.529	replies
-2.905	does	3.474	preferred
-2.826	talent	3.397	upcoming
-2.813	#thelastdance	3.258	disclosure
-2.777	songs	3.228	.@
-2.775	descubrir	3.219	die
-2.774	sing	3.156	comments
-2.761	usado	3.153	crashing
-2.733	shazam	3.127	nb
-2.686	looks	3.047	coming
-2.677	has	2.999	to
-2.668	s	2.97	proud
-2.666	para	2.959	using
-2.635	needs	2.908	misgender

Note: correlation scores above ±1.96 are considered significant (Monroe et al., 2008).

Table C.4: PRONOUN lexical associations for Caitlyn Jenner post-COE.

MISGENDER		PRON_AFFIRM	
Score	Term	Score	Term
-153.487	WRONG_PRONOUN	148.131	TARGET_PRONOUN
-95.972	CELEB_DEADNAME	93.211	CELEB_NAME
-89.136	jenner	24.877	cait
-27.231	a	21.822	new
-24.476	😂	21.446	beautiful
-21.397	man	21.051	am
-19.744	woman	20.826	reveals
-16.296	wanted	20.799	photos
-16.207	aborted	20.618	vanity
-15.806	oldest	19.767	fair
-14.429	nigga	18.579	debut
-13.95	was	16.928	first
-13.151	this	16.416	cover
-12.931	still	15.255	looks
-12.847	god	14.931	gorgeous
-12.742	kanye	14.39	espys
-12.557	shit	14.047	shows
-12.534	lesson	13.96	sexy
-12.366	daughter	13.864	in
-12.363	lost	13.719	trans
-12.218	if	13.698	marriage
-12.071	transforms	13.688	name
-11.95	outing	13.657	formerly
-11.932	into	13.539	TWITTER_HANDLE
-11.856	imma	13.466	happy
-11.829	sex	13.379	clip
-11.769	or	13.188	steps
-11.757	timbaland	13.184	stunning
-11.481	sick	12.845	style
-11.48	did	12.35	mini

Note: correlation scores above ±1.96 are considered significant (Monroe et al., 2008).