Names before pronouns: Variation in pronominal reference and gender

Kirby Conrod
University of Washington

Abstract: I use a Twitter corpus of public discourse about Chelsea Manning, a transgender political figure, to examine the relationship between names and pronouns in gendering language. I argue that although names and pronouns both mark gender, which is a social relationship subject to changes, names change faster than pronouns. This is because names are a lexical category, and pronouns are a functional category; thus, changing pronouns requires a rearrangement of syntactic features, while changing names does not. I also show the social sensitivity of names and pronouns, demonstrating that Twitter users were more likely to misgender Manning while expressing negative sentiment about her or her transgender identity.

Keywords: transgender, pronouns, proper names, natural gender

1 Introduction

In this paper I explore the way gender is marked on proper names and pronouns in English using corpus methods and focusing on a transgender referent, with the aim of highlighting the flexibility and limits of gendered language. I compare proper names, a lexical category with thousands of entries, with pronouns, a functional category with a very limited inventory based on person, number, and gender features. By focusing on a transgender referent, I will show that names and pronouns do not vary at comparable rates. I will also show that pronoun use is a socially sensitive variable and depends strongly on sentiment and attitudes about gender ideologies.

1.1 Why pronouns? Why transgender referents?

English third person pronouns provide an interesting window into the workings of gender in language for a few reasons: first, over the course of its history English has moved from grammatical gender to natural gender (Curzan 2003), and third person pronouns are among the last syntactic elements to show any reflex of gender in English. Second, because the limited set of options restricts gender, a complex social relationship, into a binary (or trinary) choice; and, finally, because pronouns are grammatical, rather than lexical, and are thus semantically quite bleached aside from gender.

Because third person singular pronouns are gendered (and neither he nor she can be fully generic: Hughes and Casey 1986, Hyde 1984, Moulton et al. 1978, Gastil 1990), they are often static: for most people, only one set of pronouns will be used in reference to them throughout their life. This has gone largely unquestioned in the syntactic and semantic literature, but becomes more complicated when we try to account for the lived realities of transgender people.¹ If one is assumed to be a certain gender in childhood, but later lives as another gender, then the pronouns

---

¹ I owe thanks to my advisors, Professors Alicia Beckford Wassink and Barbara Citko, for their valuable mentorship and insights on this project. I am also grateful to Professor Alex d’Arcy of UVic for her encouragement and comments, and for the support of Rachael Tatman, Brent Woo, and Katy King.

Contact: kconrod@uw.edu, kirbyconrod.com

¹ I will use transgender as an umbrella term for anyone who lives outside of the social gender they were assigned at birth: this may include those who self-identify as transsexual, genderqueer, gender-nonconforming, transvestites, or a number of other categories. For the case of Ms. Manning, she uses the word transgender to describe herself.
used in reference to them will likely change over time. Likewise, if one’s gender presentation does not wholly fit into one or the other of the assumed binary categories, then one may hear a variety of different pronouns used to refer. This phenomenon may constitute an interesting thought experiment for linguists interested in the gendering of language, but the gendering of pronouns can also indicate a person’s level of (un)safety, especially since transgender people are often the target of violence based on their identity (Grant et al. 2011). Besides physical danger, using pronouns that do not match with a person’s gender identity constitutes misgendering, which has been shown to have negative psychological effects on transgender people (McLemore 2015).

In comparing the names and pronouns used to talk about a transgender referent, I am interested in two sides of the issue: on the one hand, how do the grammatical properties of pronouns and proper names differ from each other in liminal situations where gender seems to be unstable across different speakers? And on the other hand, what are the social forces that cause people to choose one or the other name or pronoun to refer to a transgender person? That is – is it the case that someone who dislikes a transgender person is more likely to misgender that person?

In order to answer this question within a fairly limited scope, I have focused this study on discourse (on Twitter) about a particular transgender referent. I focus on Chelsea Manning because of her prominence in online discussion, her public notoriety, and the public nature of her gender transition from start to finish.

Chelsea Manning has received much attention in the press and in public discourse centering around her transition from the gender assigned to her at birth towards her gender identity (Capuzzo 2015). Aside from her public identity as a transgender woman, Manning is a politicized figure due to her involvement in the release of military documents (Tate 2013). Her identity as a transgender woman in the public eye is compounded with her identity as a figure of political intrigue; this is reflected in the way that English speakers refer to her. Besides using her military rank and surname, many refer to Manning by her birth name, Bradley, which she no longer uses. I will be examining what names and pronouns speakers use to refer to Manning, how they compare with each other, and how they are related to the speaker’s feelings about Manning.

1.2 Research questions and hypothesis

This paper seeks to answer two discrete questions. I will pose these each separately with explicit hypotheses and predictions.

The first research question is about gender (and how it is encoded in names and pronouns) as a feature of language. It is almost universally assumed in syntax and semantics that names and pronouns must match in gender to be coreferential; however, those of us in the transgender community have experiences that conflict with this assumption. I will address this question by specifically building a small corpus of public discussion where names and pronouns might not match in gender.

Below I state in my alternative hypothesis that names and pronouns that are coreferential might not always match in gender. I make a related prediction: more people will use the “new” name (that is – the referent’s most recent preferred name) than the new pronoun. In Section 2, I discuss differences between lexical and functional categories that motivate this prediction.

**RQ 1:** Do names and pronouns always match in gender?

**Null Hypothesis 1:** Names and pronouns that are coreferential will match in gender.

**Alternative Hypothesis 1:** Names and pronouns will not always match in gender.

**Prediction:** More people will use the new name than the new pronoun.
The second research question is about why people use the pronouns and names that they use. As I suggested in the last section, there may be a relationship between a speaker’s feelings about a referent and how they choose to gender that referent. In this hypothesis I do not delimit what the nature of the relationship may be – I will explore that question in a more qualitative way, looking carefully at representative examples of the data found in my Twitter corpus.

**RQ 2:** Do pronouns and names people use about a person reflect their stance on that person?

**Null Hypothesis 2:** There will be no relationship between gender of names/pronouns and sentiment about the referent.

**Alternative Hypothesis 2:** There will be a relationship between the gender of names/pronouns and sentiment about the referent.

In this section, I have briefly laid out why reference to transgender referents is of particular interest to linguists, and explicated the research questions for this project. In the next section, I will explore in a little more detail the relevance of Chelsea Manning as a referent of interest and give more background on the properties of lexical and functional categories that would lead us to expect a difference between names and pronouns.

## 2 Background

### 2.1 Chelsea Manning

In this section, I will provide some background on Chelsea Manning to contextualize the study. Because this study is a very narrow case study focusing on a single referent, I will be cautious about generalizing too broadly. However, I do want to stress that the tweets and data I present below are not at all abnormal for public discourse, especially regarding transgender women; readers should be aware that some data presented display prejudice against transgender people.

Chelsea Manning was born in 1987 and began her military service in 2009. She worked in Iraq as an intelligence analyst and during that assignment released 750,000 military documents, some of which were classified or sensitive, to WikiLeaks. She was imprisoned from 2010 to 2017 and was court-martialed in 2013. She announced her gender identity as a woman in 2013, and much of the publicity about her case thereafter focused on her gender transition while she was still imprisoned. Mainstream media largely started using her new first name and feminine pronouns around 2013/2014 (Capuzza 2015). At the end of his presidency in January 2017, President Barack Obama commuted Manning’s sentence. She was released that May. Since her release she has been speaking in public, and in January 2018 she announced her candidacy for Senator in Maryland.

I collected the tweets on January 20, 2017. This was three days after her sentence was commuted, and the day that President Trump was inaugurated. Manning has been active on Twitter both during her imprisonment and since her release, though she became considerably more active on her verified Twitter account (handle @xychelsea) after she was free. Her tweets are often of a political nature, but also are largely focused on optimism and positivity. She frequently retweets tweets where people are misgendering her, suggesting she should still be imprisoned, or threatening violence, often with rainbow, heart, and sunglasses emoji in her response.

I here include the summary of her Twitter presence to contextualize the public discourse about her on Twitter: while many tweets about her include her name only, there are also many who direct their tweets directly at her by including her Twitter handle in the tweet. My decision to...
use Twitter as a source of data for discourse about a transgender person – and about Manning in particular – is in large part because she is a known Twitter presence and because much discussion about her takes place on Twitter. I also am using Twitter data as a source for “speechlike” text that is easily accessed in large volume through automated methods, following studies like Tatman (2016).

In the next section, I turn from the social context to the linguistic theory: I will give a (brief) overview of syntactic analyses of pronouns, and more generally of the difference between lexical and functional categories.

2.2 Lexical vs. Functional categories

My review of the difference between lexical and functional categories will be very brief and will be primarily based on the definition given by Abney (1987:64–65), here paraphrased:

Functional elements…
- are closed lexical classes
- are phonologically/morphologically dependent or reduced
- only permit one complement, which is not an argument
- are inseparable from their complement
- lack descriptive content

I will discuss the difference between names and pronouns focused primarily on the two bolded criteria above, since they are strong and mostly theory-neutral generalizations.

Firstly, closed lexical classes are classes where neologisms, productive morphological composition, and borrowings are typically not permitted. This is widely accepted to be true of English pronouns: the paradigm resists even explicit attempts to add additional items, e.g., neologisms like *thon* which are intended to be gender-neutral third person singular pronouns.2 Proper names, by contrast, are by necessity extremely amenable to new additions; new names like *Baelynn* are coined by creative English-speaking new parents all the time. Proper names can also be produced through morphological addition, as in gendered pairs like *Michael* and *Michaela*, or the use of diminutives or nicknames such as *Edward vs. Eddy*; likewise, names can be borrowed easily across languages – the English name *Kirby* is an Anglicization of the Irish *Ciarmhac*. Proper names therefore more probably constitute a lexical class, and pronouns a functional class.

The second qualification that is easily compared is the lack of “descriptive content.” This is especially clear with proper names vs. pronouns: *Kirby* has the descriptive content that specifies only a few people (who are named Kirby), while *he* does not describe any particular person or set of people, but can generally refer to any male person (or person whom the speaker claims is male).

For both criteria, pronouns pattern with functional elements and proper names behave more like lexical elements. Since functional categories have been claimed to be more stable and resistant to change (Muysken 2008, a.o.), we would therefore expect to see pronouns changing much less rapidly or frequently than proper names change in English. However, this is not a diachronic study – instead, this case study is focused on another sort of change over time.

---

2 It may be the case that there are some Englishes that do accept new pronouns more readily; transgender and non-binary communities online especially have adopted non-traditional pronouns to a far greater extent than has the larger body of English speakers. I focus mainly on Standard American (spoken) English here, so these other pronouns are out of the scope of this paper.
Over the period of a transgender person’s (public) transition, the words used to refer to that person typically will shift – this is not precisely a ‘language change’, since the elements of the language all maintain essentially their original meanings, but I propose that this shift is instead on the semantic side. That is, speakers are still using their original mental definitions of names and pronouns, but their conception of the referent is changing over time. This may be compared roughly to aging: as a person grows from infancy to adulthood, people use different words to refer to that person (toddler vs. tween vs. senior citizen). However, English does not have a functional class that differs with the age of a referent; thus, gendered pronouns are our only real example of this sort of lifetime change in English. In predicting that names will change faster than pronouns over the course of a single referent’s gender transition, I am likewise predicting that any grammatical element that denotes a property of a person that changes over time will change more slowly than lexical elements referring to that property.

In this section, I have detailed the context in which this study took place (with reference to Chelsea Manning) and the linguistic context in which my research questions are situated (with reference to functional categories). In the next section, I will detail my methods of data collection and treatment.

3 Methods

In this section, I will first go through how I collected and analyzed tweets, and then how I went about determining positive and negative sentiment in tagging those tweets.

3.1 Collecting and analyzing tweets

I used the R script GetTweets (Tatman 2016) to collect 10,000 tweets from the Twitter API on Jan 20, 2017 that included the words Bradley [OR] Chelsea Manning. I eliminated all duplicate tweets, leaving me with 3,117 tweets. I then pulled only tweets that had one of the two names above, plus a gendered third person singular pronoun such as he or she. (I did not include it or they for the sake of streamlining, and because using either it or they to refer to Manning would merit much more detailed investigation.)

I then hand-tagged tweets for whether the name was coreferential with the pronoun in the tweet; this task is generally not automatable but was necessary to filter out tweets where pronouns were referring to other people (former President Obama being a common example in this data set). This left me with 110 tweets that included a name and a pronoun that were coreferential.

I hand-tagged these 110 tweets for whether they displayed a positive, neutral, or negative sentiment about Manning. I will go into more detail on what kind of tweets got tagged as what in the next section. I used a trinary numeric code for tagging sentiment: positive tweets were tagged with “1”, neutral tweets were tagged with 0, and negative tweets were tagged with -1. This allowed me to later take the mean of sentiment for each sub-group of tweets; I discuss this in the Results section. I used Fischer's Exact Test to determine if there were a significant number of tweets where the pronoun and the name did not match in gender (e.g., Chelsea/he, Bradley/she).

3.2 How to tag sentiment

I tagged as negative any tweets that used words like “treasonous” or “traitor,” tweets that suggested that Manning’s sentence should not have been commuted, or tweets about the negative consequences of Manning’s release of military documents. Below are examples of tweets that I tagged as negative.
(1) **Examples of negative tweets (-1)**

a. Please stop calling him Chelsea...traitor ‘s name is Bradley Manning...just ‘cause he wears a wig doesn ‘t make him a girl

b. Can this Bradley Manning commutation by Obama be reversed? R there other charges he can face? This is a HUGE miscarriage of justice!

c. #ThankYouObama for caring about the civil rights of treasonous Chelsea Manning, and not about the people who may have died bc of her actions

I tagged as positive any tweets that used words like “hero(ic)” or “whistleblower,” tweets that suggested that Manning’s commutation was just or that she should have been released earlier (or not imprisoned in the first place), and tweets that focused on the positive consequences of her release of military documents. Below are examples of positive tweets.

(2) **Examples of positive tweets (+1)**

a. Just in case no one has figured this out, Bradley Manning was tortured into becoming something he wasn’t.

b. Chelsea Manning is Free! Send her a letter of love and celebration by signing

c. What Chelsea Manning did was heroic. She exposed unethical actions in government. She is a hero.

I tagged as neutral tweets that were neither clearly positive nor negative. This included tweets that stated the facts of her sentence commutation without (obvious) value judgment and tweets where the sentiment was simply unclear from the content of the tweet. Examples of tweets I tagged as neutral are below.

(3) **Examples of neutral tweets (0)**

a. Chelsea Manning spent 7 years is mostly solitary confinement. He didn’t pardon her he shortened her sentence.

b. BREAKING: President Obama commutes Chelsea Manning’s sentence! She will be released from prison in May.

c. Wonder if they’d do the same thing to Chelsea Manning if he were trying to get through.

Because I was the only person tagging the tweets, interrater reliability measures cannot be applied. In the interest of transparency and replicability I have made available an anonymized version of the full data set on my GitHub.3

---

3 Github link to full data set: (https://github.com/kirbyconrod/manningTweetsPronouns)
4 Results

4.1 Quantitative results

To revisit the first research question: these data did show that coreferential names and pronouns do not always match in gender. There were a significant number of tweets where the gender of the pronoun did not match the name (Chi squared/Fischer’s Exact, p < 0.05). Table 1 below shows the predicted values for tweets of each pronoun and name given the null hypothesis (that names and pronouns always match). Table 2 shows the actual values found in this data set.

<table>
<thead>
<tr>
<th>Table 1: Predicted values</th>
</tr>
</thead>
<tbody>
<tr>
<td>She, her</td>
</tr>
<tr>
<td>Chelsea</td>
</tr>
<tr>
<td>Bradley</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2: Actual values</th>
</tr>
</thead>
<tbody>
<tr>
<td>She, her</td>
</tr>
<tr>
<td>Chelsea</td>
</tr>
<tr>
<td>Bradley</td>
</tr>
</tbody>
</table>

Notice in Table 2 that there are significantly many tweets in which pronouns and names do not match in gender – but only in one direction. There were 16 tweets where Chelsea was coreferential with he, but no tweets where Bradley was coreferential with she. I discuss the implications of this in Section 5.

The second research question asked whether there was a correlation between sentiment and the gender (of names and pronouns). Figure 1 below shows the mean sentiment rating for each group of tweets.

![Figure 1: Sentiment and gendering of names and pronouns](image)

The sentiment for tweets that used a masculine name and masculine pronoun was very negative overall; tweets that used a feminine name but a masculine pronoun were also very negative. The only tweets that had a positive mean were tweets with a feminine name and a feminine pronoun. I discuss these results in detail in Section 5.

In the next section I will discuss a pattern that appeared in the data that was not measurable by quantitative means. I will provide examples for this pattern, then discuss it in Section 5.

4.2 Qualitative results

In many tweets in this corpus, the tweeters made explicit comments on Manning’s gender – but, moreover, many explicitly corrected others’ uses of either names or pronouns. It was not the case that tweeters only corrected others when telling them to use Manning’s preferred name and pronouns; tweeters also made explicit corrections towards masculine names and pronouns. An example of such a correction is shown below in (4). Another correction, shown in (5), also
included an instance of switching pronouns mid-tweet – *her* is first used, but then *he* is used to emphasize the tweeter’s declaration that Manning is male despite undergoing gender transition.

(4) Please stop calling him Chelsea… traitor’s name is Bradley Manning… just ‘cause he wears a wig doesn’t make him a girl

(5) Chelsea Manning can change *her* name legally but *he* is still a man

In addition to tweets like (4) and (5), many corrections (in both directions) involved depronominalization – a syntactic phenomenon where a pronoun is used as a noun. The tweets in (6)–(9) all show depronominalizations.

(6) It’s now Chelsea Manning not Bradley Manning, the guy had gender reassignment surgery. So now **he** is a **she**

(7) @[USER] Please don’t refer to Chelsea Manning as “him,” she is a “**her**”

(8) Bradley Manning is NOT a “**her**”. Don’t believe me? Check his DNA. It’s right there in “X & Y”

(9) [a Twitter thread: an exchange between two users]
   a. Oh great. Way to encourage TREASON Obama! Manning should have been SHOT! American soldiers have DIED because of him!
   b. him???
   c. No. I didn’t misspell anything. Bradley Manning is a **HE**, regardless of if he likes it or not.

The depronominalizations in (6)–(9) are an understudied syntactic phenomenon, meriting usually brief mentions in works on the syntax of pronouns (as in Cowper and Hall 2009: 108). However, the syntactic structure of depronominalizations is beyond the scope of this paper; I merely highlight their appearance in the Twitter corpus because of their unusual frequency. These depronominalizations all occurred in tweets where the tweeter was explicitly contesting not only the gender identity of the referent, but the manner of referring to the referent. This use of pronouns as specifically calling upon a manner of referring is similar to the use of names as predicates of what-someone-is-called identified by Matushansky (2015).

It is worth noting also that all tweets correcting others’ uses of names or pronouns highlighted Manning’s transgender status; this supports the premise that the gender of transgender people is up for public debate, and that discussions might yield eventual consensus on how someone is referred to. That this is possible at all suggests, as I stated in Section 1, that so-called “natural gender” denoted by pronouns is not in fact a static property, but instead indexes a social relationship that can change or be questioned.

5 Discussion

In this section, I will briefly discuss the relevant results that I presented in Section 4 and explore their significance in relation to larger linguistic questions.

In Section 1, I proposed that there would be a difference between the changes in names and pronouns, and that names would change faster than pronouns. The results in Section 4 support
this, suggesting that there is some structural linguistic difference between pronouns and names. On a very broad semantic level, names and pronouns do the same thing: they refer directly to a particular referent. And in fact the proposed syntactic structures for proper names and for referential pronouns are often quite similar (compare Matushansky 2015 or Longobardi 1994 with Elbourne 2003, for example). The difference, therefore, cannot be strictly due to their syntactic category or their (general) semantic behavior. Instead, I will propose that proper names contain an element that pronouns lack: a lexical root.

In the cartographic tradition, lexical roots (denoted usually by √) are category-agnostic lexical entries that contain rich semantic information, like a sort of “dictionary definition”. To enter a syntactic derivation, these category-less roots must combine with a syntactic object of a certain type – for √burn to become a verb, it must combine with a v and form a vP. However, √burn can just as easily combine with a n and form a nP. This goes a long way to explain a lot of apparent category-switching in languages like English, where burn can indeed function as either a noun or a verb without any overt morphological marking.

What I propose is that what makes pronouns a functional category rather than a lexical one is their lack of a √. A pronoun still contains all the projections one expects to find in DP: D, NUM, and n (and more, depending on who you ask), but has no √she combined with n. This differentiates them from proper names, which do have a √Chelsea merged with n. The structures can therefore be almost identical, and names and referential pronouns can do all the sorts of things that we expect DPs to do – but a pronoun consists entirely of structural information. This captures the frequent claim that pronouns are ‘bundles of features’ (phi-features: person, number, and gender) and explains why English has a very limited set of pronouns available. There are only so many ways the features can combine, and without any further lexical information coming from a root, there is no other way to generate new pronouns.

If the difference between lexical and functional categories is the inclusion of a lexical root, then some general differences are explained. Lexical items can “change faster” (i.e., be used in different situations, come to mean different things) because changing a lexical root involves rewriting something in the lexical entry. Changing functional categories in this way is not so easy: you must instead re-map feature bundles to morphemes, change your feature inventory, or otherwise meddle with syntactic features in order to see change.

In this very small case study, I was not looking at a language change: in using he about Manning in earlier years and using she later on, there is not any kind of restructuring of features. However, learning a new name for someone does not necessarily require you to change any of the syntactic features you associate with that person: you can think Chelsea → masculine by simply redefining your mental definition of √Chelsea. Changing the gender features you use to refer to a person require you to remap the syntactic features associated with a certain concept – thus, more people have adopted Manning’s new name (83 uses of Chelsea appear in this data set) than her new pronoun (only 67 uses of she). And in the case of much longer diachronic changes, we see a similar pattern: proper names “switch” gender (being assigned more to either men or women) much, much more frequently than we see a remapping of pronouns to gender.

The final important takeaway from these data is that pronouns – which are known to be sensitive to social relationships (Raymond 2016, Simpson 1997, a.o.) – can apparently shift to accommodate changes in the way gender is conceived of within a society. If “natural gender” were a static, immutable, and obvious property of a referent, then we would see no disagreement, no corrections, no debate, and no misgendering. But we see all of these in this very small data set, all focused on a single person. Manning is by no means alone in her gender transition, and she is by no means the person with the most complicated relationship to gender in the English-speaking world. I can report anecdotally that pronouns can shift as often as moment to moment, and these
data support the premise that gender is a social relationship which pronouns happen to grammaticalize in English in the third person. English is certainly not unique in grammaticalizing our social categories, and in future studies I hope that linguists will be sensitive to the wide array of possibilities for gender to change, be complicated, or be contested.

References


