Pronouns Raising and Emerging

Kirby Conrod

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Barbara Cikto, Co-chair

Alicia Beckford Wassink, Co-chair

Edith Aldridge

Program Authorized to Offer Degree:
Linguistics

University of Washington

Abstract

Pronouns Raising and Emerging

Kirby Conrod

Co-Chairs of the Supervisory Committee:
Professor Barbara Cikto
Linguistics

Professor Alicia Beckford Wassink
Linguistics

This dissertation revisits the question of the syntactic and semantic status of pronouns, incorporating new syntactic, sociolinguistic, and pragmatic data to support an analysis of n-to-D head movement in the nominal domain. The support for pronouns originating in n comes from predicative pronouns, including pronominal relative clauses (1) and depronominalizations (2).

(1) he who is without sin
(2) That person is a she.

I compare predicative pronouns with variable grammatical restrictions on singular *they* using data from two sociolinguistic studies that I conducted. I show that there is an effect of speaker age on production and perception of definite, specific uses of singular *they* (dsT) as in (3), while definite generic (4) and epicene uses (5) are more broadly accepted.

(3) Jayden forgot their homework.
(4) The ideal student never forgets their homework.

(5) Every student should do their homework.

I take the sociolinguistic variability in singular *they*, particularly the differences related to age, as evidence of an ongoing change in the grammar of English towards increasing use and acceptance of the type shown in (3). The inclusion of dsT in the grammar predicts intraspeaker sociopragmatic variation in pronoun use, and this prediction is borne out. Through variable rankings of pragmatic constraints I show that dsT enables speakers to include or exclude gender features from pronominal choices in order to achieve strategic discourse goals in various contexts.

The n-to-D head movement analysis that I propose accounts for predicative pronouns, as well as differences in grammaticality of dsT and its related discourse-sensitivity, by separating pronouns into sub-classes depending on how far head raising proceeds. For predicative pronouns, external determiners (overt or covert) block head movement completely, and pronouns stay in n. For epicene pronouns like (5) and definite generic antecedents like (4), the pronoun raises from n to an intermediate functional projection (Num) but is merged with a variable D. Finally, referential (specific) pronouns like (3) are formed through movement from n through Num to D, where the pronoun combines with a phase head D that is linked to a discourse referent. Because phase edges are sensitive to discourse context, it is only when a pronoun moves to D that it is evaluated for context-appropriateness relative to the referent picked out by D.

This system of context-appropriateness necessitates analyzing the natural gender features of pronouns as less like noun classes and more like honorifics, in that they signify social relationships rather than grammar-internal categories. An honorific analysis of gendered pronouns more robustly explains the sociopragmatic variation found in natural language use, and is more generalizable cross-linguistically.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Barbara Citko, Alicia Beckford Wassink, and Edith Aldridge: Thank you for the chocolate and long talks and warmth and brilliance and kindndess that you have shared with me throughout this process. I could not have possibly dreamed of better mentors than you, each one of you individually and together. You have profoundly bettered my life much beyond the effect on my work. Thank you, thank you.

Hayden Stern, love and light of my life, you are a brain genius for finding me in the middle of this and convincing me to have an extracurricular life and mind. You inspire me and you remind me to eat snacks and you make my non-syntax-related hours worth being awake and alive and present for. Thanks for marrying me, thanks for putting music and art and love into the dark corners of my life.

My small flock of trans/nonbinary and affiliated linguists, thanks for finding me and being amazing friends in addition to your beautiful insight and help throughout this. Lauren Ackerman, Caitl Light, Edwin Howard, Brooke Larson, Leah Velleman, Lex Konnelly, Bronwyn Bjorkman, your collaboration and friendship have gotten me through times where I would have otherwise been crumpled. A very very big thank you to Mallory Tucker, without whose dedication I would not have made it through my own mountain of data. My friends at UW, your support has been warm and inspiring, Brent Woo, Rachael Tatman, Katy King, Alli Germain, Stevie Peterson, Ben Jones, Amie De Jong, Alli Germain, Rik Koncel-Kedziorski, Katie Manlove, Nathan Loggins, Kellianne Bennett, Nataliya Griggs, Nicole Chartier, Molly FitzMorris, Anna Moroz, Amandalynne Paullada, Rob Squizzero, Alec Sugar. Those other faculty and staff at UW who were not directly involved in reading this document over and over also provided me invaluable support both

material and emotional, and I am grateful to Richard Wright, Emily Bender, Joyce Parvi, Mike Furr, and Misha Burgess for all they have done to make my time here lovely. Thank you to the very dear mentors and friends at UC Santa Cruz who made me a linguist in the first place—especially Jim McCloskey, whose bright and vibrant teaching I will forever be trying to emulate. Linguists who saw me and saw my work and did what they could to magically put me in front of audiences, thank you for your inspiring support and belief in me—Rusty Barrett, Kira Hall, Lal Zimman, Alex d'Arcy, and many others, I am indebted to you. Linguists and humans alike who saw something in my work that sparked their own, who spoke with me online and who were beautiful and inspiring in their own right, I am grateful for the cheer and cheerleading.

My friends in Seattle who have made sure I eat and see the sunshine once in a while, who let me bother your cats and let me drag you to brunch with singleminded fury, thank you–Erin Murphy, Sean Jarvis, Cate Uselton, Rachael Tatman, Conner Kasten, Alder Truelove, Nick Woods, Max Kristofferson, thank you. Thank you as well to my friends farther away who still cheer me on, listen to my kvetching, and praise my memes–Anna Leonid, Devin Tankersley, Alison O'Connor-Korb, Neirin Necropticon, Kate Lubeck, Denzelle Marcovicci, and all the rest of you who can't be named in print, thank you, I love you.

Thank you to Batman for your love and support through the first half of my time here, and may your memory be a blessing. Finally, thank you very much to the cats who tolerated a great deal of botheration from me, which was very helpful for my mental health. Nim, Willie Behemoth, Galois Woods, and my own darling Toes, thank you.

DEDICATION

Magpie, who loves excellently and makes good language

Chapter 1

FOUNDATIONS

1.1 Introduction

This dissertation aims to examine the gender features of third person pronouns in English from a syntactic and sociolinguistic point of view, with particular attention to potential development around a gender-neutral innovation in the English pronominal system, singular *they*. A primary goal of the work I present here is to provide evidence for a more nuanced understanding of gender features as they are instantiated in syntax, to be supported by sociolinguistic and pragmatic study of the way that sociocultural and conversational context can influence those features.

Recent work on the grammaticality of singular *they* has made reference to apparent variation, but this variation is currently understudied and not well-explained in these works (cf Ackerman, Riches, and Wallenberg 2018; Bjorkman 2017; Konnelly and Cowper 2017 i.a.). Bjorkman (2017) has observed that speakers report different grammaticality/acceptability judgments for definite, specific uses of *they* in the singular (1).

(1) ? That syntax professor$_1$ loves their$_1$ job.

The fact that speakers of English vary in their reports of grammaticality for sentences like (1) brings up several questions. First, why or how do problems with pronouns 'mismatching' antecedents cause (or get reported as) ungrammaticality—is there a specific mechanism in the grammar that rules out sentences like (1) for some speakers, but not for others? Such a grammatical mechanism must be linked to the makeup of pronouns, and must be linked to the morphology, syntax, and semantics (at least), in order to capture the specific effect of pronominal coreference.

This question has a place in a long lineage of morphosyntactic and semantic investigations of pronouns, and in order to answer it I need to address the syntactic category of pronouns, the semantic makeup of pronouns as it relates to syntactic structure, and how the morphology can fall out of the syntactic structure in a principled way. In Chapter 2, I introduce data that help to support a complex categorical analysis of pronouns, intended to account for both noun-like and determiner-like behavior.

The second question that arises from the varying judgments of (1) is whether those judgments are socially mediated, or socially meaningful, in any way. Because singular *they* has received a great deal of attention in public discourse, it seems likely that varying uses of this pronoun are connected to salient social meanings. In Chapter 3 I show data from two sociolinguistic studies, aimed at production and acceptability of definite, specific singular *they* respectively. These experiments give evidence that not only is this use of singular *they* socially meaningful, but also that its use is increasing over time. The change over time is evidence that Bjorkman's (2017) intuition (that uses like (1)) are new and increasing) is correct, and that a grammatical account needs to also explain how this change proceeds.

The metalinguistic comments around singular *they* also suggest that speakers have both conscious and unconscious opinions about its use, and that it will be sensitive to pragmatic context. In Chapter 4 I analyze the use of singular *they* in alternation with other third person singular English pronouns, showing instances where speakers change the pronoun they use about a referent in a single conversation. In order to explain this intraspeaker variation in its context, I show how different sociopragmatic considerations can be variably ranked. For example, whether a speaker prioritizes specificity over the drive to only include relevant details predicts whether they use *they* or *she* in certain contexts.

After showing the variation that is enabled by, and intrinsic to, the English pronominal paradigm, in Chapter 5 I return to the question of how to incorporate flexibility and context-dependence into a theory of syntax. The desiderata for any syntactic account of

pronouns are that 1.) N-like and D-like behavior should be equally well explained, and differentiable by the syntactic structure; 2.) that the parameter which controls acceptability of elements like singular *they* be fine-grained and socially sensitive enough to support the kind of interspeaker variation shown in Chapter 3; 3.) that the morphosyntactic features differentiating forms of third person pronouns be flexible and discourse-sensitive in a way that robustly explains and predicts the pragmatic variation shown in Chapter 4; and 4.) that any syntactic account which correctly describes these qualities is also contingent upon logical extensions of an internally-consistent theoretical framework and its mechanisms. The account I give in Chapter 5, based upon head movement and incorporating insights from use conditional semantics, fills these criteria. I show how pronouns can be separated into subcategories (referential, variable, and predicative) that are distinguished by their empirical properties, and which empirical properties are a direct consequence of how far head movement takes place.

Because of the data around which this dissertation is centered, the gender of pronouns is a significant concern for syntax, sociolinguistics, and pragmatics. English pronouns are said to reflect natural, rather than grammatical, gender, but how natural gender is determined has previously been under-explored in linguistics. In order to build a model of gender that can correctly and inclusively account for non-normatively gendered speakers and ways of speaking, I follow Butler (2011) in analyzing gender as a constantly-negotiated and updated social behavior that is at least partly built from speech-acts. This is consistent with Ackerman's (2019) three-tiered model of gender, differentiating biosocial gender from either grammatical or conceptual gender. This means that I need to justify whether there should be formal features for gender in the syntax at all; I argue that there should, on analogy with other morphosyntactic properties cross-linguistically that participate in syntactic operations like agreement. However, the features that I include for gender are not reflective of "biological sex" or static categories of referents; instead I compare gender marking to phenomena like honorific pronouns that encode social relationships. This chapter gives further details on that comparison.

In order to set up the foundation for this dissertation, in the following sections I review accounts of pronouns from the perspective of several different disciplines within linguistics. After positioning myself in syntactic and sociolinguistic theoretical landscapes, I discuss analyses of the semantic contribution and composition of pronouns, followed by their syntactic structure (in the context of DP structure somewhat more generally). In looking at the syntax of pronouns I also discuss the syntactic location of formal gender, person, and number features (with special attention to gender). I then turn towards the more social side of pronouns, looking at sociopragmatic analyses of pronouns (including the T/V alternation and some cross-linguistic comparison) and in particular looking at use and acceptability of singular *they* in English. Finally, I discuss why the traditional account of gendered pronouns is not compatible with more contemporary understandings of social gender, and review some discussion of sociolinguistic theory that helps me to account for pronoun production behaviors that are complex with regard to gender.

1.2 Theoretical position

The purpose of this section is to make explicit my position in the theoretical landscapes of syntax and sociolinguistics, particularly highlighting what literature informs my work and what core assumptions I use to build my questions and hypotheses.

1.2.1 Syntactic theory

In my analysis I work within the Minimalist Program, somewhat loosely interpreted (Chomsky 1995, 2000, 2001, 2004, 2008). I assume that the mental faculty of language consists of (at least) a lexicon in which words or morphemes are stored, a minimal syntactic structure-builder which combines lexical items (LIs) through a basic combination mechanism (Merge), and two post-syntactic components, one of which controls gestural/articulatory movements (the PF, or physical form) and one of which is responsible for interpretation of sentences (the Logical Form, LF).

The narrow syntax, which is strictly responsible for combining LIs into phrases and sentences, operates through two basic functions. A syntactic structure starts with an unordered set of LIs, the Numeration. Merge can combine either a LI from the Numeration with already-existent structure (external Merge), or Merge can combine an LI from within the structure, connecting it to another part of the structure (internal Merge). Agree is an operation in the narrow syntax where a dependency is formed between syntactic objects in order to value features (Chomsky 2000, 2001). The syntactic structure is cyclically "sent" to the LF/PF component when particular functional heads, phase heads, are merged into the derivation (Chomsky 2001, 2008).

I am working in a version of Minimalism that decomposes lexical items into category-free roots and categorizing heads; lexical words like *shoe* and *blast* are stored in the lexicon without information about syntactic category, and only "become" nouns or verbs upon being Merged with a categorizing functional head (n or v). I follow accounts of lexical decomposition along the lines of Borer (2005).

1.2.2 Sociolinguistic theory

This section provides the theoretical foundation upon which I am building my sociolinguistic research questions and hypotheses. This dissertation uses insights from both variationist and interactional sociolinguistics, in order to examine and explain interspeaker variation and intraspeaker variation respectively. My variationist framework is built on the work of Weinreich, Labov, and Herzog (1968) and subsequent investigations (for applications of this theory for syntactic variation, see e.g. Adger and Smith 2010; D'Arcy, Haddican, Richards, Tagliamonte, and Taylor 2013). These studies typically share the assumption that language change over time can be studied through synchronic data; that language variation can result in language change but does not always do so; and that language variation is mediated by socially meaningful relationships and identities. Chapter 3 describes experiments in the production and perception of a particular English pronoun, definite specific singular *they*, and analyzes these data from a variationist perspec-

tive. In that chapter I explain my specific assumptions about how diachronic change can be inferred from synchronic data (the Apparent Time Hypothesis) and how those assumptions stand despite evidence of individual variation and possible change over time among individual speakers. Chapter 3 details how the Apparent Time Hypothesis, based on assumptions of a critical language acquisition period and grammatical stability in adults, states that variation related to age in synchronic data can imply diachronic change (cf Sankoff and Blondeau 2007; Weinreich et al. 1968).

The interactional sociolinguistic tradition that I work in is built primarily on investigations of cooperation and interaction through politeness, such as that based on Goffman (1967) and P. Brown and Levinson (1987). These works follow foundations by Gumperz (1982), and follow the basic premise that particular linguistic behaviors are embedded in the larger social and interactional behaviors that speakers are engaging in; under this model, language is one of several tools that speakers use to communicate their intentions and achieve their goals. This approach is particularly useful for investigations of specific speech acts in context; I will take an interactional approach primarily in Chapter 4 in order to analyze why a particular speaker uses a particular pronoun in a particular context.

In the next section I review foundational accounts of pronouns from generative semantics and (morpho)syntax. I first discuss some accounts of what pronouns contribute to the semantic meaning of an utterance, then discuss the necessary background for a full syntactic account of pronouns. The review of syntactic analyses will start with a broad explanation of the DP hypothesis and clausal structure in the nominal domain, followed by investigations into the categorical status of pronouns and their component formal features (especially gender).

1.3 Semantic contribution of pronouns

This section presents three different analyses of the semantic contribution of pronouns. In this dissertation I focus primarily on referential pronouns rather than pronouns acting

as bound variables[1] (or anaphora). Example (2) contains a referential pronoun with a linguistic antecedent, while the pronoun in (3) is a variable and has a semantic interpretation bound by a quantifier.

(2) This is Jay. **He** is my friend.

(3) Every male professor needs to be respectful to **his** TAs.

While (3) is not the focus of this dissertation, the analysis I put forth in Chapter 5 should be compatible with pronouns of this type. This section first reviews Elbourne's (2013) analysis of pronouns as definite descriptions, then the presuppositional account of (the features of) pronouns; finally, I present an alternative account of semantics that attempts to explain the conditions of appropriateness as they combine with the meaning of an utterance.

1.3.1 Pronouns as Definite descriptions

Elbourne (2013) posits that pronouns, whether they are independently referential or bound variables, constitute definite DPs that are (largely) devoid of lexical content, but are semantically composed and interpreted in the same way as other definite descriptions. In order to account for the context-sensitivity of pronominal reference (as well as to describe other types of definite descriptions) Elbourne shows how introducing a situation variable s to the denotation of definite descriptions allows for the binding of referents within relevant context as determined by the discourse. Thus, Elbourne suggests that pronouns are constituted of an iota operator (which provides uniqueness) and the phi-features (person/number/gender) of the pronoun, and that the construal of the iota+phi is bound by the situation s. The iota operator in (4a) below denotes a function where there is a unique entity x and x is a cat. (4b) shows that the same operator can be applied to the phi features of a pronoun.

[1] The term variable is used by semantics and sociolinguistics very differently; in this context I refer to the fact that the construal of a pronoun when it is a bound variable is dependent on a local binder, rather than fixed to a particular referent.

(4) a. $\iota x.x$ cat(x) (=*the cat*)

b. $\iota x.x$ FEMALE(x) (=*she*)

The inclusion of the situation variable s in Elbourne's denotations restricts the uniqueness requirement of the iota operator, so that it is not necessary for there to be only one unique cat in the universe in (4a), but rather one unique cat in the relevant situational context. I do not follow Elbourne's exact formalization of the s variable, but I do discuss the importance of pragmatic context more extensively in Chapter 4.

In his discussion of pronouns modified by relative clauses, Elbourne posits a null NP complement to the D pronouns (which are definite determiners with phi-features), meaning something like *person*; it is this null NP to which the relative clause adjoins in his analysis. I discuss pronominal relative clauses in Chapter 2, and show how a nominal analysis of non-referential pronouns allows a more parsimonious unification between pronominal relative clauses and other instances where pronouns can be modified or restricted (e.g. depronominalizations like *a she*).

While Elbourne (2013) analyses pronouns as D heads with null NP complements, his semantic analysis would be equally workable if pronouns were specialized NPs that combine with a null definite D. I more extensively discuss the syntactic internal structure of pronouns in Section 3 of this chapter.

1.3.2 Presupposition

Semantic presupposition (as opposed to pragmatic presupposition) is a type of entailment where a proposition A *must* be true in order for a proposition B to be semantically coherent. This very strict model of presupposition has several issues, including defeasibility where the supposed entailment relationship is easily manipulated by context or world-knowledge (Levinson 1983). The 'presupposition' of (5a) in (5b), for example, can be cancelled by the additional phrasing in ((5)c). *Regret* typically presupposes its complement, but sufficient context lets hearers interpret (5c) without the presupposition of

(5b).

(5) a. John doesn't regret that he did a PhD (Levinson 1983:187)

b. *John did a PhD.*

c. At least John won't have to regret that he did a PhD.

Pragmatic presupposition is a more flexible model where a statement B will be felicitous/appropriate dependent upon the shared assumption between interlocutors that a statement A is true; this shared assumption, or common ground, is updated throughout a discourse (ibid.).

It is very commonly assumed that phi-features on pronouns (especially gender features) contribute presuppositional material, rather than propositional, to the meaning of an utterance (Sudo 2012, i.a.). One argument for a presuppositional analysis of gender on pronouns is the transparency to sentential negation, as shown in (6):

(6) a. She went to the store.

b. She did not go to the store.

In (6a), the pronoun *she* is taken to mean that the subject is feminine[2] – but negating the sentence as in (6b) does not negate this (i.e. 6b does not suggest that the subject is NOT feminine). Under the semantic model of presupposition, (6a-b) presupposes that the subject of the sentence must be a woman in order for either statement to be felicitous, as shown below:

(7) a. She went to the store / She didn't go to the store

b. Proposition: there is a person x such that x went to the store

c. *Presupposition:* Person x is feminine.

[2] I complicate this 'meaning' of gendered pronouns throughout Chapters 2, 4, and 5; for now I use the word feminine as an abbreviation for what the actual social relationship is denoted by she.

Under a pragmatic model of presupposition, however, (7c) needn't necessarily be true for (7a) to be felicitous: instead, both interlocutors must have the mutual understanding and awareness that (7c) is true (whether it is or isn't, in actuality). In Chapter 4 I show how speakers use conversational and conventional implicatures to introduce gendered meaning into the discourse; the data are incompatible with a strict semantic presuppositional account of pronominal gender features such as Sudo's (2012), but rely more on discourse considerations including relationships between interlocutors and referents.

1.3.3 Use conditional semantics

Gutzmann and McCready (2014) propose treating gender features of pronouns in terms of appropriateness, rather than truth conditions. In Section 1.5.4 below I detail how appropriateness is evaluated in relation to Japanese honorific marking; the same analysis can be applied to the gender features of pronouns. Their proposal has in common with presuppositional accounts the two-dimensional nature of semantic denotations, where the gender features of pronouns are analyzed by a separate system from the propositional semantics (cf. also Potts 2007). Potts, along with Gutzmann and McCready, depart from Sudo's (2012) analysis in not requiring the gender feature of pronouns to be evaluated for 'truth,' but rather for whether the pronoun has been used appropriately in a given context. This line of approach is similar to that of pragmatic presuppositions (rather than semantic ones), and is also much more readily applied to pronominal encoding of complex social relationships other than gender, such as honorifics. In Section 1.5.4 I discuss an application of this system, use conditional semantics, to Japanese honorific marking; in Chapter 5, I describe Gutzmann and McCready's analysis of English pronouns in more depth, and show how it fits into my syntactic proposal.

All the accounts of the semantic interpretation of gender features of pronouns agree that (free) pronouns are definite, and that context is necessarily part of how pronouns get construed; the accounts differ with respect to the relative strength/importance of discourse context, as well as with respect to how the discourse context can be formally mod-

eled with respect to other components of the grammar (e.g. how or whether discourse features should be included in the narrow syntax or transmitted to the LF component). In Chapter 2 I show how the semantics of referential pronouns are comparable to proper names, and in Chapter 5 I expand on this to take up Gutzmann and McCready's analysis of pronouns as referential descriptions. This is a departure from Elbourne's (2013) analysis, and Chapter 5 shows exactly how the difference between *referential* and *definite* descriptions can be formalized and consequently extended to pronouns.

1.4 Syntactic makeup of pronouns

This section reviews syntactic accounts of pronouns. I first give background on the nominal domain, including the DP Hypothesis (Abney 1987) and other functional projections in DP. I then show different accounts of the syntactic category of pronouns, the first being that pronouns are universally determiners, and the second being that pronouns vary in their category (either across or within languages). I then explore proposals for the location of phi-features in DP, both for pronouns and other DPs; this includes accounts where phi-features project their own functional head, are separately located on different functional projections (e.g. gender on *n*, person on D), and accounts where gender itself appears on multiple different heads (e.g. natural gender on D, grammatical gender on *n*).

1.4.1 DP structure

In this section, I briefly discuss the expanded functional projection in the nominal domain as it is relevant to pronouns and *n*-to-D raising. This includes, first, an overview of the DP Hypothesis and its reasoning, followed by a proposal for one intermediate functional projection, NUM, between *n* and D.

The DP Hypothesis (Abney 1987) is a proposal that nouns are universally selected by a functional category, D, and that D is the highest head of the nominal domain. One advantage of the DP Hypothesis (DPH) is the regularization of syntactic structures between domains in such a way as to show parallels between verbs and nouns (lexical categories)

and tense/Infl and D(eterminers) (functional categories). Abney's argument was in part driven by the parsimony of finding direct parallels between the nominal domain and the clausal domain, and this parallelism is likewise supported by later work on phasehood of DP and CP (see also Bernstein 2001; Szabolcsi 1987).

In the DP domain where D is the head, there may also be other functional projections that constitute grammatical components of nominal structures. To illustrate with one example, Ritter (1995) argues for an intervening projection above N and below D, NumP, which allows for a full differentiation of nominal features onto different layers of syntactic structure.

(8) [DP $D_{[definiteness]}$ [NumP $Num_{[number]}$ [NP $N_{[gender]}$]]] (Ritter 1995:418)

Ritter also notes that "*the [person feature] implies definiteness, but definiteness does not imply person*" (Ritter 1995:421). Thus, person features also appear on D in this model. This more expanded nominal hierarchy also allows Ritter to differentiate between pronouns that are Num heads (Hebrew pronouns, which show gender agreement) and D pronouns (as in Abney 1987; Postal 1966). Ritter highlights a general property of pronouns cross-linguistically: first- and second-person pronouns appear to behave differently than third person pronouns. Ritter notes that third person Hebrew pronouns can and do co-occur with articles and demonstratives, suggesting that they are not D heads themselves:

(9) (Ritter 1995:420)

a. ha-hu
the-he
that.MASC

b. *ha-ani
The-I
*that.1sg

For Ritter, $1^{st}/2^{nd}$ person pronouns are, as previous accounts proposed, DPs that consist only of D heads specified for phi (person/number/gender) features; however, third person pronouns are Num heads that can appear as complements to non-pronominal determiners such as definite articles.

1.4.2 Pronouns as D(eterminers)

This section addresses the question of the syntactic category of pronouns; I first discuss the popular understanding that pronouns are universally D heads, then compare it with accounts where pronouns can consist of varying levels of functional structure (such that some consist of NPs where others are DPs, etc.). The two 'split' accounts (where pronouns are not a single unified category) that I most closely review both suggest that this split exists both within and across languages.

Postal (1966) quite famously challenged the status of pro-"nouns" (including his title, which referred to "so-called pronouns") with evidence that pronouns behave like articles and demonstratives in distribution and co-occurrence. He notes that pronouns and other determiners (including articles, demonstratives) appear in complementary distribution, as shown in (10). He also showed evidence that pronouns can, under certain circumstances, act as determiners selecting lexical nouns, as in (11) below.

(10) * the he

(11) we honest policemen (Postal 1966:71)

This account of pronouns is very widely accepted, and pronouns are assumed to be determiners in semantic accounts such as Elbourne's (2013). Some criticisms of Postal's reliance on data like (11) come from Delorme and Dougherty (1972), who instead explore the possibility that *we linguists* constitutes a NP *we* with an appositive NP *linguists* adjoined to it. I do not deeply discuss the arguments for the appositive NP analysis, except to suggest that if *linguists* is a type of a reduced relative modifier, it is likely to

be restrictive (like pronominal restrictive relative clauses, which I discuss in Chapter 2) rather than non-restrictive (as Delorme and Dougherty assume).

1.4.3 Different pronouns are different sizes

This section reviews 'split' accounts of pronominal category, where pronouns are not considered as a uniform phenomenon but rather are a class that consists of sub-types with different categories (and predictable behavior based on these categories). This includes Ritter's (1995) account of Hebrew pronouns, which I discussed above, where 1^{st} and 2^{nd} person pronouns constitute different functional heads in the nominal hierarchy. These accounts vary in terms of how *many* sub-categories make up pronouns as a class: Ritter (1995) and Cardinaletti (1994) identify two types, while Déchaine and Wiltschko (2002) identify three. The data that I give in Chapter 2 would be compatible with a two-type account, but I show in Chapter 5 that a three-type account can capture more fine-grained and consistent differences in pronoun behavior.

Déchaine and Wiltschko (2002) propose a three-way division of types of pronouns, such that different types of pronouns can constitute DPs, *phi*Ps (an intermediate nominal projection), or NPs. Their typology can be summarized by the divergent behaviors that contribute to their classification, shown in Table 1.1.

Example	Halkomelem independent pronouns	Shuswap independent pronouns	Japanese *kare*
Proform	pro-DP	pro-phiP	pro-NP
Internal syntax	D syntax, morph complex	neither D nor N syntax	N syntax
Distribution	argument	either arg/predicate	predicate
Semantics	definite	-	constant
Binding status	R-expression	variable	-

Table 1.1: Sub-types of pronouns, from Déchaine and Wiltschko 2002: 419

Table 1.1 shows the consistent differences between the three types of pronouns that Déchaine and Wiltschko identify. Pro-DPs pattern as full DPs do, meaning that they have

a morphologically complex syntax and are distributed in sentences where full DPs can appear (argument positions); semantically they are definite and behave as R-expressions, meaning that they refer directly to entities with or without antecedents. Pro-phiPs show what Déchaine and Wiltschko call "neither D nor N syntax," in this case meaning that they cannot act as predicates as NPs would (in Shuswap at least; 2002:415), but also they can be modified by articles (which DPs wouldn't). The main defining property of phiPs for Déchaine and Wiltschko is that they can act as bound variables (while pro-DPs cannot, since they are referential). Pro-NPs show consistent noun-like syntax (in that they can be modified by any nominal modifier), appear in predicate positions (and never argument positions), and are generally not subject to binding conditions.

Déchaine and Wiltschko give as an additional example of a pro-NP the English *one*, as in *the big one*; in their system $1^{st}/2^{nd}$ person pronouns are pro-DPs (they invoke the example of *we linguists*) and 3^{rd} person pronouns are phiPs. Their analysis of 3^{rd} person pronouns in English as phiPs is largely dependent on the apparent unavailability of **they linguists* combined with the fact that 3^{rd} person pronouns (but not 1^{st} or 2^{nd}) can act as bound variables in English.

Déchaine and Wiltschko's split analysis is similar to some earlier accounts of pronouns that attempt to explain both determiner-like and noun-like properties of pronouns. Pesetsky (1978) shows that there is a correlation between which languages allow category-switching of various syntactic categories and which languages allow D-like behavior of pronouns. For languages that do allow category-switching, pronouns can switch to be determiners; they are otherwise NPs. For Pesetsky, pronouns in the "*We linguists*" construction are in fact determiner pronouns, but it is not necessarily the case that *all* pronouns are determiners; they are usually inserted in N as the head of an NP.

Cardinaletti (1994) also gives a mixed account of pronoun category, though hers relies on head movement rather than category-switching (though these may be two theoretical mechanisms for describing the same set of empirical phenomena). Cardinaletti draws an analogy between the N-to-D raising in pronouns that she proposes and Longobardi's

(1994) N-to-D raising analysis of proper names in Italian. While Longobardi specifically shows that pronouns in Italian do not appear to pattern with proper names (e.g. names can co-occur with a definite article while pronouns cannot), Cardinaletti shows that different types of pronouns in Italian show different patterns. For Cardinaletti, pronouns in Italian can be *strong* or *weak*: "strong" pronouns may appear in any position, and can be stressed; whereas "weak" pronouns (like clitics) cannot be stressed and must appear in a derived position. Cardinaletti gives the divergent structures of strong/weak pronouns as in (12)):

(12) a. **Strong**: [$_{DP}$ D lui_i [NP N t_i]]

b. **Weak**: [$_{DP}$ D *il*]]

Strong pronouns are full DPs which contain NPs and have N to D raising; weak pronouns consist only of functional projections (just a D head for clitics, or a D head with a "support" morpheme for weak non-clitic pronouns).

In support for her analysis of N to D movement (rather than suggesting that strong pronouns enter the derivation at D), Cardinaletti notes first that strong pronouns are never homophonous with determiners, and, second, that pronouns can potentially be preceded by an adjective in exclamatory phrases, while determiners otherwise cannot:

(13) a. *Povero lui!*
poor him!

b. *Povero (*il) professore!*
poor the professor!

In Chapter 2 I show further instances of pronouns co-occurring with nominal modifiers, using data from English, to provide strong support for an analysis of English pronouns similar to the N to D raising Cardinaletti (1994) proposes.

1.4.4 Location of phi features (in pronouns or otherwise)

In addition to the debate on the categorical status of pronouns, sub-lexical elements that make up pronominal paradigms are often assumed to correlate with particular syntactic projections (though not necessarily; Elbourne's (2013) account, for example, makes no such assumptions). Since pronouns are a functional category, they are constituted of so-called "bundles of features."[3] This analysis of functional categories as composed of formal features is in part based on analyses of pronouns specifically (cf Muysken 2008:2, 155) and grammatical marking of phi-features is considered to be cross-linguistically very common if not universal.

Broadly cross-linguistically the relevant features for differentiating pronouns are person, number, and gender, grouped together as phi-features. In theories where formal features must be present in the narrow syntax, there have been several proposals about where in the DP structure the phi-features enter the derivation. I briefly review the proposals most directly relevant to gender, which is the primary concern of my analysis.

I start with a brief overview Sauerland's (1998) manuscript in which he proposes that, due to the presuppositional nature of phi-features semantically, syntactically they occur in a phi-P above the DP layer. Sauerland's syntactic evidence for this approach is largely motivated by phi-agreement outside DP; the DP-external position of the phi-features therefore allows phi probes to Agree with the *closest* phi-features (which would otherwise be inaccessible if they were inside DP, in Sauerland's model). This approach is not broadly adopted due to the difficulties of accounting for details of phiP over DP, and I do not pursue it much further, except to note that the exact meaning of what presuppositions are contributed by phi-features (and how) is an issue I take up in great detail in Chapter 4.

---

[3]The earliest citation I have found for the phrase "bundle of features" is Trager (1956), but this is a phrase used by many authors without a lot of discussion on why that phrase in particular is used; I invoke it critically in this chapter and later in Chapter 5 to foreground the question of how features are getting "bundled" in the grammar, among other issues.

In a comprehensive review, Kramer (2016) looks at various analyses of gender features, comparing accounts that incorporate data from languages with grammatical gender as well as social gender marked morphologically. In this account, Kramer argues that gender features likely enter the derivation on *n*, the functional projection responsible for 'typing' roots into nouns (Borer 2005).

Notably, however, grammatical gender and social gender do not always align in individual lexical items; when this occurs, languages have various strategies for reconciling conflicts. In a diachronic study of English gender and pronouns, Curzan (2003) suggests that pronominal gender is biased towards natural or social gender in instances of conflicts, and that this bias was one of the driving forces in the loss of grammatical gender in the history of English. It is standardly (but not universally) assumed that grammatical gender is a reflex of an uninterpretable feature, while social gender is the interpretable counterpart (cf King 2016; Kramer 2016 a.o.).

In order to reconcile the differences between grammatical gender and natural/social gender, especially when dealing with data where the two are in direct conflict, some have proposed that gender features enter the derivation at two different points: once on *n* (grammatical gender) and once on D (social gender, otherwise called "natural" gender or "semantic" gender). This is the approach taken by Armoskaite and Wiltschko (2012). Support for this–which I refer to as the *mixed gender hypothesis* (MGH)–comes from instances where the lexical/grammatical gender of a noun conflicts with the actual referent of the DP. In Russian, DPs show mixed gender morphology where 'high' nominal modifiers (including demonstratives, etc.) can have the gender marking matching social gender, and 'low' nominal modifiers (and the noun) are marked for the grammatical gender.

(14) moj-a nov-aja klassn-yj rukovoditel' vsë pričital-a...
my-F new-F class-M supervisor.M ITER complained-F
'my new (female) class supervisor continually complained (that)...' (Pesetsky 2013:38)

There are several advantages of the MGH that are comparable to mixed pronominal analyses: first, it allows the differentiation between uninterpretable gender features (on *n*)

and interpretable ones (on D) while still potentially providing a pathway for feature sharing/transmission; second, it captures the descriptive insight that grammatical gender features are associated with nounhood, while social gender features are associated with context. One problem with differentiating *n* and D gender is that it does not provide a clear explanation for a diachronic change of language systems from grammatical to social gender over time; the MGH also does not easily account for the apparent expression of gender features on predicative pronouns in English (which I explore in detail in Chapter 2). The analysis of head-raising I give in Chapter 5 should be able to account for the apparent split between grammatical and natural gender through context scanning at D with respect to specific referents.

In a recent extension of the MGH, Sigurðsson (2018) proposes a discourse-sensitive mechanism for the valuation of interpretable (D) gender. Based on data from Icelandic, Sigurðsson again differentiates between gender as it is related to nominal classes (*n* gender, which is uninterpretable and correlates with grammatical gender) and gender as it is determined by the social position of a particular referent (interpretable D gender). Sigurðsson proposes that there are two ways for D gender to get its value: either it can probe within the DP and find the gender feature on *n*, which results in determiner gender matching the *grammatical* gender of a referent; or D can scan discourse context in order to get valued for 'natural' gender. In analogy with context-scanning at phase edges for relative tense features, Sigurðsson proposes that the phase head D contains an edge linker that allows the syntax to scan discourse context and value the D gender feature based on that context. This account for the valuation of D gender explains how gender agreement phenomena can be seen even at great (linguistic) distances, as well as why there exists inter- and intra-language variation in whether pronouns match either grammatical or 'natural' gender features of their antecedents.

This section provided a broad overview of syntactic and semantic accounts of pronouns with special attention to the status of gender, both in how it is reflected in the syntax and how it is interpreted in the semantics. The next section shifts gears signifi-

cantly, focusing on studies of pronouns from a sociolinguistic standpoint.

1.5 Sociolinguistic behavior of pronouns

Because the following chapters present novel sociolinguistic and sociopragmatic data, this section briefly reviews some matters of sociolinguistic interest around pronouns. Chapter 3 begins with a more extensive review of literature investigating the use of singular *they* in English, before showing new data on definite, specific uses of singular *they* in two experiments. In this section I instead turn to cross-linguistic investigations of pronouns in their social context, including honorific pronouns. I then discuss some of the issues of 'gender' as an incoherent concept in linguistics, pointing towards the difficulty of differentiating gendered language and extralinguistic types of gender (such as social behaviors and relationships). Finally, I lay out the sociopragmatic theoretical frameworks which are necessary in order to analyze the English pronoun data that I present and discuss in Chapter 4.

1.5.1 Honorifics and pronouns

In this section I review research on pronouns as they relate to social relationships. Much of this research has focused on honorific pronouns, which include formal/informal pronouns such as the 'T/V' alternation in 2^{nd} person pronouns (in German, Old and Middle English, French, Italian, Spanish) as well as open-class pronouns in Thai and honorific pronouns in Japanese. The analyses of these honorific pronouns are relevant to my analysis of gendered pronouns in English because Chapters 3 and 4 present data that suggests that 'gender' of pronouns is a linguistic reflex of a complex social relationship, much like honorific pronouns are; Chapter 4 provides extensive sociopragmatic data to show the instability of 'gender' as a category (of referents) and argues instead for a more socially dynamic analysis.

1.5.2 Pronouns of Power and Solidarity

In a seminal survey on the use of second person pronouns across European languages, R. Brown, Gilman, et al. (1960) pick out what they call a T/V distinction—that is, *thee*/*you*, *tu*/*vos*, *tu*/*vous*—all originating from a convention in Latin whereby an emperor was addressed by the plural (*vos*) as a signal of deference. Brown and Gilman identify two axes along which the semantic classification of address is decided between speakers: that of power differentials, wherein a superior is addressed by V and an inferior is addressed by T; and that of solidarity, wherein those interlocutors who share some significant social similarity may address each other mutually by T (or, if they are in the upper echelons of society, both mutually by V).

Brown and Gilman characterize both diachronic variation (through analysis of literary texts in various languages) and synchronic variation (through a survey of native speakers of French, German, and Italian). The diachronic trajectory they posit is a shift from the axis of power prevailing (guiding pronominal choice) towards the axis of solidarity, including an intermediary stage of instability that enables some of the synchronic variation they observe. The synchronic variation consists not only of cross-linguistic differences between the three languages, but also individual differences between speakers of these languages. The individual speaker differences correlate with a scale of political or ideological progressivism: the more radical, anti-hegemonic, and progressive a speaker is, the more they are likely to favor T with all interlocutors. Brown and Gilman suggest that this is not arbitrary, but that the ideological anti-hierarchical stance is itself reflected in broader use of pronouns marking solidarity across class, race, sex, nationality, or other social divisions.

The final type of variation Brown and Gilman identify is stylistic—in literary works in English and French, characters utilize *expressive shifts*. A shift to T when V is normally used signals temporary contempt or anger (thus a lowering in status); a shift to V when T is normally used by a pair expresses an elevation in status resulting from admiration or

respect. However, Brown and Gilman point out that these expressive shifts do not survive into the contemporary languages they studied: none of the speakers of French, German, or Italian they surveyed reported any such stylistic moves. Rather, in the contemporary languages (at least these three), they report that any given pair of speakers has established between themselves a generally stable address pattern.

The crucial contribution from Brown and Gilman is not just that different social dimensions are reflected in language use (such as status or solidarity), but that functional elements of language (in this case, pronouns) can be a part of expressing those social reflexes, and that those systems of mapping social relationships to language can change over time. In going forward, I use these insights to explain the patterns found in English third person pronouns.

1.5.3 Pronouns in Discourse Context

Raymond (2016) regards the methodology of R. Brown et al. (1960) as seminal, largely setting the tone for following research that investigates the *inventory* of pronouns in languages and dialects; since Brown and Gilman's data is all self-reported metapragmatically, and all the historical data is found in literary works, it is true that much following work does not focus on the point of expressive or affective shifts. Raymond focuses on these shifts as a way of rethinking what the object of study is in investigations of pronominal choice: he looks not just at what pronouns a speaker has available to them in a given interaction, but when and how speakers leverage those resources to perform specific discursive acts turn by turn. This problematizes some of the basic assumptions of Brown and Gilman and subsequent studies: rather than being static (even relative to a specific interlocutor), identity stance (defined below) and context may shift with every turn in a conversation; every utterance is an opportunity for a speaker to use pronouns to achieve some particular goal.

In a focused analysis of the Spanish spoken across several different dialects, Raymond eschews the self-reporting used by Brown and Gilman in favor of naturally-occurring

conversational data, including both institutional talk (calls to 911, televised political interviews) and casual conversation (between friends or family). As an analytical tool, Raymond differentiates between *identity* ***status*** (the summative total of the aspects of someone's social identity) and *identity* ***stance*** (what aspect of identity someone may invoke at any given time).

The differentiation between identity status and identity stance allows Raymond to account for, as an example, a 911 caller who begins the phone call using *tu* and *voseo*—the non-deferential, casual forms, in Brown and Gilman's terms—but later switches to *Usted* when he is concerned that he may not get the help from the police that he is asking for. In Raymond's terms, the *identity status* of the two speakers remains static (911 call-taker and call-maker), while the *identity stance* shifts (the caller begins as a person asking for service; the 911 operator later takes on a role as a gatekeeper to that service).

Each subsequent transcript closely analyzed by Raymond shows similar signs around shifts: when a speaker changes pronominal form, it is always with a shift in stance to accomplish some particular goal. This is not contradictory with the previous analyses of how the semantics of these pronouns work—rather, Raymond writes:

> . . . while the underlying semantics of *Usted* and *tú* can indeed carry with them the notions of social distance and intimacy, respectively, in a given dialect (Brown & Gilman 1960, Brown & Levinson 1987), the ground-level pragmatic significance of invoking such distance or intimacy is no more automatic or predetermined than the identities of the interactants themselves. Rather, the *interactional relevance of these pronominal options is conditioned by way of the moment-by-moment negotiation of identities* in and through the ongoing talk. (2016:651) [emphasis added]

In positing a continuous moment-by-moment negotiation of stance and how it can invoke multiple relative identities, Raymond provides new tools for considering grammatical resources such as verb agreement (in the case of Spanish) and person-referring expressions

such as pronouns more generally. In Chapters 3 and 4 I carry forward with Raymond's core proposal that any pronoun used by any speaker at any particular turn can encode constant negotiation and re-negotiation of identity and social relationships.

While this section has primarily focused on studies of second person pronouns in Romance (and Spanish specifically), the T/V alternation represents a relatively modest paradigm for honorific pronouns. The next two sections discuss analyses of Japanese and Thai respectively, both of which have a more enriched inventory of honorific markers. The discussion of Japanese has the advantage of discussing syntactic agreement and semantic composition directly, and the section on Thai gives an example of what to expect from languages with open class pronouns from a sociopragmatic perspective. The ultimate goal of these sections is to give context for a richer social analysis of pronouns and how that might be incorporated into formal grammars.

1.5.4 Honorific marking in Japanese

Because it is an open question whether social relationships should be encoded in formal features within the syntax, in this section I briefly discuss previous analyses of Japanese honorific agreement. Japanese, like Spanish, shows morphological marking on the verb that agrees with certain arguments based on their honorific properties. Unlike Spanish, Japanese can also mark full DPs with honorific affixes (not just pronouns), and can show verbal agreement with direct objects as well as subjects. This verbal agreement has been taken as evidence that some feature (HON) marking an honorific exists in the narrow syntax and is accessible for the formal operation Agree. This section also gives an analysis of how the semantics/pragmatics of honorifics can be composed in a generative theory, which I show in Chapter 5 is analogous to pronominal gender.

Syntactic analyses of honorific agreement in Japanese have largely been in two camps: the first, put forth by Niinuma (2003) and later by Boeckx and Niinuma (2004) and Boeckx (2006) have argued that honorific marking on Japanese verbs is a reflex of Agree (the specific formulation of which is dependent on c-command) where v probes its do-

main and agrees with its direct object; opposition to this analysis comes from Bobaljik and Yatsushiro (2006), who reveal technical problems with Boeckx & Niinuma's (Boeckx and Niinuma (2004)) account of agreement. None of these papers suggest how DPs come to be marked for honorific properties in the first place; it is largely assumed that honorific DPs enter the derivation with some formal feature HON(orific) that is interpretable on DP and uninterpretable on *v*.

Semantic analysis of honorific marking in Japanese complicates what would be predicted by the HON feature (interpretable or not); Potts and Kawahara (2004) show that honorific marking does not directly affect propositional semantics but also does not appear to correlate with traditional presuppositions. Potts and Kawahara (2004) instead suggest that honorifics are a type of emotive definite description, of the same type as expressives (e.g. in (15)) below).

(15) That damn dog barked at me again. (Gutzmann and McCready 2014:59)

Expressives like *that damn dog* are not evaluated in terms of truth content—there is no exact truth conditional meaning tied to the *'damn'* descriptor that would render the dog either damn or not-damn. Potts and Kawahara compare qualities of expressives to the qualities of honorifics: nondisplaceability (where the meaning is directly tied to a particular speech-act in context); independence (where the meaning does not project through traditional presupposition "holes" like negation); immediacy (wherein the force of the expressive or honorific has a secondary value independent of the force of the primary speech-act); and descriptive ineffability (where native speakers struggle to formally 'define' expressives or honorifics). Potts and Kawahara instead argue that in order to semantically evaluate emotive definite descriptions (including expressives and honorifics), a second axis of evaluation is necessary, so that the 'meaning' of (15)) is two-tiered.

(15') That damn dog barked at me again. (Gutzmann and McCready 2014:59)
Tier 1: *The dog barked at me again*
Tier 2: *I don't like the dog*

The two-tiered model for evaluating emotive definite descriptions can also be applied to pronouns, and in fact Gutzmann and McCready (2014) propose such an analysis for English pronouns as well. In Chapter 5 I detail how that analysis of gendered pronouns can more accurately capture the real conditions under which speakers choose between various options by incorporating the additional axis of use-conditions into a phase-dependent theory of head raising. Namely, I show that use conditions as Gutzmann and McCready describe them can only be evaluated upon the merge of a phase head that links a structure with the discourse context—so the gender of a particular pronoun can only be evaluated for appropriateness when it is linked with a specific referent (by merging a referential D). This means that, rather than analyzing pronominal switches as aberrant, each referential pronoun, through virtue of its gender features, constitutes a referential description that contributes towards a constantly-updated common understanding between interlocutors. This analysis is more suited to capturing the kinds of social and discursive shifts that are shown in Chapter 4.

1.5.5 Open class pronouns in Thai

In this section I briefly describe some research into Thai pronouns as a basis for comparison to English gendered pronouns, primarily in order to reveal context-sensitivity of pronouns more generally, as well as to directly address whether English pronouns are moving towards an 'open class.' Palakornkul (1975) describes the extensive Thai pronominal system, which includes over thirty 'proper' pronouns in 1st, 2nd, and 3rd person, as well as an extensive inventory of pronominalized nouns including kin and pseudo-kin terms, personal names (in pronominal positions), friendship terms, titles, and loanwords. Palakornkul provides a series of general rules for choosing between pronominal forms ('person-referring expressions'), especially when faced with a potential conflict between competing forms (as when interlocutors have a multidimensional social relationship).

Simpson (1997) investigates the use of person-referring expressions by Thai speakers, noting that individual speakers switch pronouns (even for self-reference) mid-conversation

for discursive reasons such as affect (shock, teasing, sarcasm, emphasis) or discourse moves (changing footing or voice). She also notes that speakers have explicit metapragmatic awareness of the discourse sensitivity around pronoun use, and comment on their own and others' use of different pronouns in context. Simpson therefore proposes "a more dynamic alternative [model] which takes social and contextual categories as the background against which strategic choices are interpreted" (1997:223). In going forward with my analysis of English gendered third-person pronouns, I show parallel behaviors where speakers switch pronouns for discursive reasons (Chapter 4); however, I argue that English has not (yet) reached a stage where pronouns constitute an open or lexical class (and hence why pronouns encoding a wider variety of social relationships are not yet gaining ground). Chapter 2 details this reasoning based on data comparing English pronouns to proper names (a contrast which, in this respect, differs from Thai patterns).

This section and the preceding one are focused on showing cross-linguistic pronominal behaviors (and analyses) that can be used to explain why English pronominal gender can be flexible in discourse context. The next section provides more background on previous accounts of why gender (as a static, binary category) is a problematic foundation in the first place; data from Chapters 3 and 4 will directly support the critiques from Hall and McConnell-Ginet in the next section.

1.5.6 Problematizing gender

One of the biggest problems with discussing the meaning (propositionally or socially) of gendered pronouns in English and other languages comes from the problematic assumptions that underlie much of the work on gendered language. In Section 1, I defined *she* as having a presupposition where the referent is "feminine" – but this is a very subjective and difficult condition to diagnose, and depends upon sociocultural context as well as linguistic context. Sociolinguists have already endeavored to highlight the problems of assuming a binary, static gender property that is simply marked on grammatical elements without any social negotiation.

Hall (2003) shows the failure of binary, static assumptions about gender in language through an analysis of gender morphology in the speech of Hijras speaking Hindi; she shows that the Hijra identity, which is a gender identity that does not map neatly onto Western binary conceptions of *Male* vs *Female*, is at least partially constructed through mixed and dynamic use of gender morphology. Hijras showed expressive/affective shifts within a single conversation, similar to the shifts of Spanish and Thai pronouns in the previous section. Binary conceptualizations of gender are not sufficient for adequately describing and explaining the use of inflectional morphology for expressing mood—the presupposition that a referent or speaker "is" feminine or female suggests that this is an immutable property.

McConnell-Ginet (2013) shows the instability of these assumptions when discussing what she calls "the myth of 'natural' gender" in English, using English pronouns as a case study. Informed by Hall's (2003) data, she shows that English pronoun use is more expressive and mutable than semantic accounts assume. The use of *she* within groups of gay men, presumptive leaps in instances of unknown gender for a referent, and misgendering (an intentional use of the wrong pronoun for a referent) all show that gender is a resource available to English speakers through the use of pronouns, and can be exploited for more than just reference differentiation. Rather than "natural" gender, McConnell-Ginet proposes that English 3rd person pronouns reflect *notional* gender, which encodes not a categorical biological distinction but rather the sociocultural associations around gender groups in a given context.

Ackerman (2019) lays out a three-tiered system for analyzing gendered language: grammatical gender (which is the inflectional marking that groups nouns into classes); conceptual gender (which is predicated upon speakers' knowledge and associations regarding gender); and biosocial gender (which is the "real-world" gender as experienced by subjects in society, and is constituted of aspects of identity, expression, socialization, and sexed embodiment). Under Ackerman's three-tiered system, English pronouns as McConnell-Ginet describes them are controlled by conceptual (not grammatical or bioso-

cial) gender; this makes them a resource freely taken up by speakers for various discursive goals. In Chapter 4 I present sociopragmatic data that demonstrates what these goals may be and how speakers accomplish them through gendered pronoun use (or lack thereof).

In the next two sections I outline some background about the interactional sociolinguistic theories that inform Chapter 4. These theories fall more under the umbrella of pragmatics, and in some ways represent the intersection between semantics and sociolinguistics. Politeness Theory and Stance are both conceptions of social conventions that condition linguistic behaviors in consistent and predictable ways. Chapter 4 is a detailed examination of how these and other pragmatic considerations predict pronominal shifts more adequately than a strict semantic account.

1.5.7 Politeness

Politeness theory is a framework in interactional sociolinguistics aimed at theorizing how interlocutors use speech acts to maintain and build social relationships. In this dissertation I refer primarily to a version of Politeness theory in which interlocutors' needs for respect and positive feelings can be described as their 'face,' and where speech acts can serve to maintain or protect speakers' face or, conversely, threaten or harm it (P. Brown and Levinson 1987; Goffman 1967). In this theory, it is generally assumed that interlocutors foreground important social roles and identity statuses encoded in their language, and cooperate based on shared conventions in order to maximally preserve the face needs of themselves and their conversational partners. Issues of variation or optionality arise when the face needs of one participant are in direct conflict with the needs of another, and these conflicts are resolved with respect to individuals' internal rankings of the relative importance of these needs.

P. Brown and Levinson (1987) expand on the general idea of face in Politeness theory to separate out two separate kinds of needs: *positive face* is an individual's need to be well-thought-of and maintain a positive consistent self-image; whereas *negative face* is an individual's need to be free from obligations or impositions upon their will. Brown

and Levinson's theory of politeness, like Goffman's, focuses primarily on how discourse between interlocutors (alternating *speaker* and *hearer*, or synonymous terms) constructs the participants' self cooperatively. In Chapter 4 I extend Brown and Levinson's (1987) form of Politeness theory to extend to third person referents (who may not be present in a conversation, and are typically analyzed as non-participants); this is a necessary step towards explaining the use of third-person pronouns and reference in socially sensitive ways.

1.5.8 The stance triangle

While Politeness Theory and other pragmatic considerations have previously been applied to analyses of pronouns (especially T/V pronouns and other honorifics), a novel contribution of this dissertation is my analysis of stance in Chapter 4. Based on data from a sociolinguistic study of misgendering (Conrod 2017c) I propose that gendered third person pronouns are one way of taking stance about a referent. In Chapter 4 I show how third person referents can be analyzed as objects of stance, and how gendered pronouns can convey affective or evaluative stance about a referent in particular contexts. In order to model the relationship between a speaker, a hearer, and a third person referent, I refer to DuBois's (2007) stance triangle.

Du Bois (2007) defines *stance* as "*a linguistically articulated form of social action*" (2007:139) which assigns value or positions subjects in relation to referents in discourse. As a component of speech-acts, stance-taking acts are collaboratively constructed by interlocutors in relation to their context and to each other. DuBois identifies three types of stance; examples of each are given below.

- **Evaluation** – an object or referent is assigned value by one or more participants

  (16) *That's good*

- **Positioning** – a type of evaluation where the stance-taker includes themself in a

relationship with the stance object; "the act of situating a social actor with respect to responsibility for stance and for invoking sociocultural value" (2007:143); can be affective (17a) or epistemic (17b)

(17) a. *I hate it*

b. *I don't know*

- **Alignment** – defining stance in relation to an interlocutor or other participant

(18) *I agree (with you).*

In order to place these types of stance in context, DuBois defines roles (named in analogy to thematic roles) with relation to any stance-taking act. A stance taking act must have a Speaker, a Stance Subject (not necessarily the same as the Speaker), the type of stance, and the Stance Object. These roles can be combined into a stance diagram.

(19)

Speaker	Stance Subj.	Type	Stance Object
JAMIE;	I	like (evaluates/positions)	this song

(Du Bois, 2007)

Arguments in the stance diagram can be (verbally) implicit, provided by discourse context. However, the stance diagram given in section 4.5.5 has no place to indicate *alignment*, where Stance Subjects relate their stance to that of an interlocutor. Without alignment, the stance-act cannot include information that would explain utterances like (18). Thus, DuBois gives another dimension to the stance relationship, whereby two subjects can independently evaluate a single object, but also can align themselves with or against the other's evaluation/position.

(20) *I evaluate something, and thereby position myself, and thereby align with you* (Du Bois 2007:163)

This relationship between interlocutors in relationship to both a mutual stance object and each other is DuBois' stance triangle. In Chapter 4 I show how the stance triangle can successfully explain otherwise-mysterious shifts in the use of third person pronouns.

1.6 Conclusion

In this chapter I have reviewed syntactic, semantic, and sociolinguistic accounts of pronouns with the intention of giving background for the following chapters. In Chapter 2 I examine the syntax of English pronouns when they can be modified by nominal modifiers such as relative clauses, articles, and adjectives. Chapter 3 gives a sociolinguistic account of singular *they* in its innovative use, and Chapter 4 shows how the existence of definite, specific singular *they* gives rise to sociopragmatic variation. Chapter 5 proposes an analysis of English pronouns that incorporates head-raising (following Cardinaletti 1994) as a mechanism which controls the sociolinguistic and syntactic patterns that I have shown in the preceding chapters.

Chapter 2

PREDICATIVE PRONOUNS IN ENGLISH

This chapter presents evidence for a new category of pronouns: rather than referential (free) pronouns, or bound pronouns (variables or anaphors), I present evidence for the existence of pronouns that behave as predicates, more similar to nouns than to full DPs. I focus on two types of predicative pronouns–pronominal heads of relative clauses, and depronominalizations– and discuss the syntactic structures necessary for these constructions to exist in English. The first two sections of this chapter show the empirical properties of each construction; the third section compares and contrasts the two constructions.

To briefly introduce the two phenomena I discuss in this chapter: pronominal relative clauses (PRCs) are relative clauses headed by pronouns, as in (1). I have argued in Conrod (2017b; 2018) that these kinds of relative clauses are derived through head raising; this analysis is motivated by reconstruction effects and by the impossibility of split-antecedent relative clauses with pronominal heads. The second phenomenon that I discuss here, depronominalizations, consists of pronouns preceded by a determiner, as in (2).

(1) **He who laughs last**, laughs best.
(2) I was looking for **the other he.**

In the fourth section I propose an underlying syntactic structure that can account for the predicative pronouns, and explain the difference between these and the pronouns that are bound as variables or that are free and referential. In my analysis, pronouns enter the derivation at the functional nominal projection *n* (not N), which allows the maintenance of a distinction between functional categories (pronouns) and lexical categories (common

nouns) while also showing how pronouns can combine with modifiers and determiners.

The theoretical goal of this chapter is to provide empirical evidence for an analysis of pronouns as structurally complex, and to motivate my proposal that English pronouns are base-generated low in the nominal domain and later optionally undergo head movement towards D.

2.1 Pronominal Relative Clauses

Pronominal relative clauses (PRCs) are relative clauses that are headed by pronouns rather than lexical nouns. In this section I first review relevant previous work, then show the syntactic properties of PRCs that pertain to the syntactic properties of pronouns more generally.

2.1.1 Previous Work on PRCs

In his seminal paper on English pronouns, Postal (1966) argues that pronouns are determiners, of the category D rather than N. Under Abney's (1987) DP Hypothesis, (typical) pronouns constitute intransitive D heads with no NP complements, as in (3). Postal's evidence for English pronouns as determiners rests rather crucially on examples like (4) below, which he analyses as examples of pronouns acting instead as transitive determiners:

(3) $[_{DP}\ [_{D}\ \textit{he}\ ]\ \emptyset\ ]$

(4) a. we linguists
    b. you traitors

Much subsequent work on the debate between D pronouns and N pronouns deals with data like (4); a common alternative analysis is that *linguists* is an appositive NP modifying *we*, or else a reduced restrictive modifier (smaller than a RC). A much more thorough review of this debate appears in Chapter 1; for the purposes of this section I

only present Postal's examples that suggest PRCs are sometimes ungrammatical, shown with his judgments in (5) below.

(5) a. we who are opposing Fascism (Postal 1966:45)
b. * he who she married (Postal 1966:47)

However, I will note here that there is an independent reason for the ungrammaticality of (5b): it is insufficiently generic, given the nature of the relative clause. In fact, given sufficient context suggesting genericity, (5b) can be coerced into grammaticality:

(6) When a man and a woman are in love, they may marry. He who she marries must always respect his wife.

In order to formulate a syntactic theory of pronominal relative clauses, it is necessary to fit them into a typology of relative clauses more generally. Citko (2004) delineates a three-way separation of relative clauses into fully headed, headless, and light-headed relative clauses. Fully headed relative clauses are those with a complete nominal head; headless relative clauses are those without any nominal head, otherwise known as free relatives; light-headed relative clauses under this analysis are relative clauses which lack a nominal head but do have some sort of determiner element in the place of the head Citko (2004).

(7) **Fully-headed relative clause** Polish
a. To jest książka, którą wczoraj kupiłam.
this is book-NOM which-ACC yesterday bought-1sg
This is a book that I bought yesterday.
(Citko 2004: 106)

(8) **Headless (free) relative clause**
a. Jan śpiewa co/gdzie/kiedy/jak Maria śpiewa.
Jan sings what/where/when/how Maria sings
John sings what/where/when/how Mary sings.
(Citko 2004: 103)

(9) **Light-headed relative clause**

a. Jan zobaczył tego, kogo Maria zobaczyła.
Jan saw this who Maria saw
John saw the one Mary saw. *Demonstrative head*

b. Zjawi się ktoś, kto ci może pomóc.
appear self somebody who you can help
Somebody who can help you will appear. *Indefinite head*

c. Wszystko, co mogło się zdarzyć się zdarzyło.
everything what could self happen self happened
Everything that could have happened has happened. *Universal head*

(Citko 2004: 98–99)

Citko (2004) compares LHRCs to both headed and free relative clauses and finds differences in both comparisons. This leads Citko (2004) to the conclusion that LHRCs are a unique, third type of relative clause.

In comparing LHRCs to headed relatives, Citko (2004) notes that only *który*, 'which,' is allowed as a WH relativizer in headed relatives; the exact opposite is the case for light-headed relatives in Polish.

(10) **WH pronoun in headed relatives**

a. Zapytam studenta, który pierwszy przyjdzie.
ask-1sg student$_{ACC.MASC.SG}$ which$_{NOM.MASC.SG}$ first comes
I will ask the student who shows up first.

b. *Zapytam studenta, kto pierwszy przyjdzie.
ask-1sg student who first comes-PERF
I will ask the student who comes first. (Citko 2004: 106-107)

(11) **WH pronoun in light-headed relatives**

Potrzebny jest nam ktoś, kto/*który wszystkich zna.
need is us somebody who/which everyone knows

We need someone who knows everyone. (Citko 2004: 107)

Citko (2004) also compares LHRCs to free (or headless) relatives. The primary difference in Polish is a matter of case matching: free relatives require the case assignment

inside and outside of the relative CP to match, while this requirement does not hold for LHRCs.

(12) a. *Wezmę [ktokolwiek$_{NOM}$ pierwszy przyjdzie]$_{ACC}$
take-1sg whoever first comes
I will take whoever shows up first. No mismatch allowed in FRs

b. Wezmę tego$_{ACC}$, kto$_{NOM}$ pierwszy przyjdzie.
take-1sg DEM who first comes
I will take the one who shows up first. Mismatch OK
(Citko 2004: 104)

This is also true for English, made apparent only by the case marking on pronouns. The grammaticality contrasts in English free relative clauses are subtle, due to other issues with *whom*, but FRs are less tolerant of case mismatches than PRCs. Still, in (13) below, the non-syncretic form *whomever* is degraded in instances of case-mismatches, while no such contrast exists for (morphologically distinct) case-mismatches in PRCs (14). I conclude that the case assigned inside the relative clause for English PRCs does not need to match the case assigned in the matrix clause.

(13) **Free Relative Clauses**
a. People respect [whoever$_{NOM}$/*?whomever stands for freedom]$_{ACC}$
b. [Whoever/*?whomever$_{ACC}$ people respect]$_{NOM}$ will win the election.

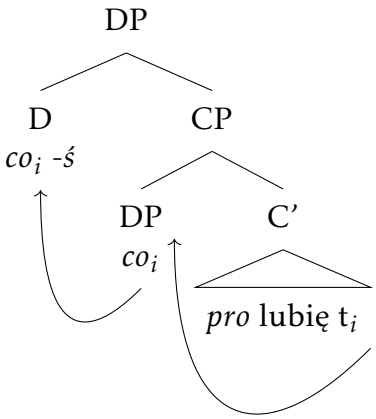
(14) **Pronominal Relative Clauses**
a. Everyone respects [him who t_{NOM} stands for something]$_{ACC}$.
b. [He who everyone respects t_{ACC}]$_{NOM}$ will win the election.

Citko (2004) concludes based on these two comparisons that Polish LHRCs have a different structure from either headed or headless relative clauses. In addition, Citko (2004) notes a morphological similarity in Polish between WH words and the allowable light heads: demonstratives, negative indefinite determiners, indefinite determiners, and universal quantifiers. Thus, Citko (2004) analyses the light heads of LHRCs in Polish as D heads which are composed by means of combinatorial movement, in which the D head

of the relative clause begins the derivation within the relative clause, moves to Spec CP, then moves further to combine with an externally Merged D:

(15) a. Czytam coś co lubię
read-1sg something what like-1sg
I am reading something I like.

b.



Citko's (2004) treatment of LHRCs does not specifically cover PRCs; Polish does not in fact allow PRCs, only headed by non-pronominal D-like elements (Barbara Citko, p.c.)[1]. However, English does allow the type of non-pronominal LHRCs seen in Polish. As they are headed primarily by quantifiers, I will class these quantificational relative clauses (QRCs) as a separate subtype of LHRCs. (16) below show a comparison of an English QRC with an English PRC.

(16) a. he who will not die
b. many who will not die

QRCs are more directly analogous to Polish LHRCs and other LHRCs cross-linguistically; however in English QRCs rely on either noun phrase ellipsis or null nouns to license re-

[1]Citko adds that this construction may appear infrequently in Polish, but only in Biblical quotes; I have observed generally that translations of Biblical writing tend to introduce PRCs to languages that do not have them natively cross-linguistically. It may therefore be a construction borrowed from Greek, which does apparently have PRCs natively.

strictive readings of the relative clause (i.e., *"Many who will...* ≈ *"Many PEOPLE who will...*). This suggests that the quantificational heads of QRCs (and demonstratives—*those who will...*) are likely structurally high in DP but include lower, silent nominal complements. Furthermore, QRCs need not be generic and do not have significant semantic departures from full relative clauses. Thus, while QRCs are of interest when investigating non-pronominal DP structure and nominal ellipsis, they will not be a useful point of comparison against PRCs and other structures where pronouns are modified. The remainder of this chapter focuses primarily on PRCs, and the analysis given here will not be necessarily applicable to QRCs or other apparent nominal ellipsis phenomena.

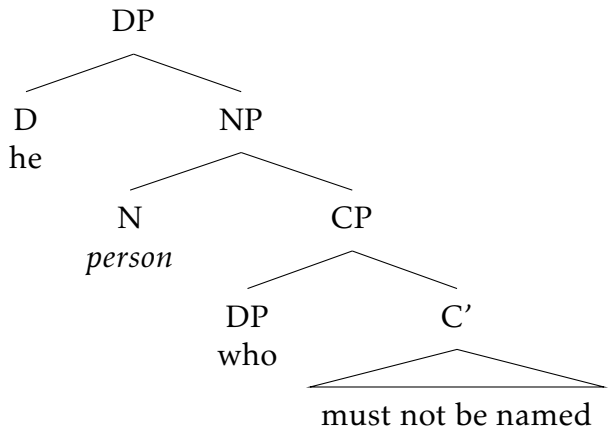
Elbourne (2013) analyzes pronominal relative clauses (what he and Zobel call 'Voldemort phrases') without making recourse to other types of LHRCs, such as those headed by demonstratives (as in Polish) or other D-like elements (e.g. *few who are brave*). Despite this, Elbourne's (2013) analysis of PRCs is in some ways similar to Citko's (2004) analysis of LHRCs. The differences are in the nominal layer and in the use of movement in the derivation. Elbourne's (2013) syntactic analysis of PRCs involves a structure in which the pronominal head is a definite determiner. Citko (2004) does not use the null nominal layer, as seen in Elbourne's (2013) proposed structure, and Elbourne makes no case for a raising or matching analysis in that work. Citko (2004) uses a raising structure which, specifically for Polish determiners, involves a partial copying operation from Spec CP to D as shown above.

The analysis of relative clauses in Elbourne (2013) is agnostic as to whether the pronominal head originates within the relative clause or merges externally, but Elbourne (2013) does include a nominal layer in the structure that constitutes a null element meaning something like 'person.' It is not clear whether this is an element that is present and elided later, or whether it is a null element that is never pronounceable in the first place; nor is it clear what the consequences of this ambiguity would be for Elbourne's analysis.

(17) [[he [*person* [who must not be named]]] (Elbourne (2013): 207)

Elbourne argues for an analysis of pronouns as definite articles occupying the D head, and suggests the following structure for PRCs:

(18)



Furthermore, Elbourne (2013) does not propose a structural difference between restrictive and non-restrictive relative clauses; this is a distinction that others (Partee 1975:231 and many following) maintain through a difference in the height of attachment. Standard treatments of relative clauses generally adjoin to the NP level for restrictive relative clauses, and to the DP level for nonrestrictive ones. Elbourne and Citko both conceive of the 'head' of a LHRC as a D – for Citko, this is a D that combines directly with a CP, while for Elbourne, pronouns are D heads that select a null NP complement.

There are two main differences between Citko's (2004) and Elbourne's (2013) analysis: the first is the inclusion of a nominal layer in Elbourne (2013); the second is the movement of a D element in Citko (2004), about which Elbourne (2013) remains agnostic. The two analyses both place the light head in D: Elbourne's pronominal D is base-generated there, while Citko's demonstrative D moves there through a copying operation.[2] Zobel (2015) analyzes the semantic denotation of restrictive PRCs, demonstrating that they de-

[2]Citko's move-and-copy operation is specifically designed for the Polish data upon which her analysis is based, since there is apparent morphological similarity between the WH-words and the D that selects the relative clause. Because English does not show the same morphological similarity, and because there is evidence for an intervening nominal layer in English PRCs, I will not utilize the move-and-copy mechanism.

note generic statements (in the relative CP) generic kinds (the pronominal head). The insight of this semantic derivation captures what makes pronouns—which are usually definite, and which otherwise would be referential—allowable as heads of restrictive relative clauses. In order to head a restrictive relative clause, a pronoun must not be referential, must not have an antecedent, and must not be interpreted as specific. Zobel (2015) derives a semantic structure of PRCs that analyzes them as generic kinds: that is, *he* is interpreted as something more like *the sort of man*. The analysis is compatible with a general analysis of generic kinds where a definite DP does not refer to an entity, but rather attributes a property to a kind of entity, e.g. *The tiger lives in the jungle*. This allows us to maintain the definiteness of pronouns while relieving them of direct reference—a point which will be important for my analysis in Section 4.

PRCs are also, on the surface, easily mistaken for a similar construction which Donati and Cecchetto (2011) identify in Romance languages; however, these so-called pseudorelatives have some properties that English PRCs do not have. The pseudorelatives in Italian that Donati and Cecchetto (2011) discuss are, on the surface, very similar to PRCs in English: they appear to be essentially restrictive relative clauses headed by pronouns.

(19) Ho incontrato lui che baciava Maria.
(I) have.met him that kissed Maria
I met him while he was kissing Maria.

(Donati and Cecchetto 2011: 548)

However, Donati and Cecchetto (2011) identify a major characteristic of Italian pseudorelatives that English PRCs do not share: there is a subject/object asymmetry with respect to allowable relative-clause-internal roles the head can play.

(20) *Ho incontrato lui che Maria baciava.
(I) have.met him that Maria kissed
I met him while Maria was kissing him.

(Donati and Cecchetto 2011: 548)

This is crucially not true in English, as noted above in case-mismatch examples. Indeed, regardless of the case of the head, English pronominal heads may be either subjects or objects within the relative clause.

(21) a. He who loves God is holy. Subj.
b. He who God loves is blessed. Obj.
(22) a. He who kisses Maria is a lucky man. Subj.
b. He who Maria kisses is a lucky man. Obj.

Donati and Cecchetto (2011) conclude that pseudorelatives in Romance languages are therefore not instances of true relative clauses, but rather that the pronominal head undergoes a different type of movement from relative head-raising[3]. The only allowable relativizer in Italian pseudorelatives is *che*, a complementizer; WH-pronouns do not appear as relativizers.

Once again, this does not hold for English PRCs: WH-pronouns are in fact preferred over complementizers for relativization in English PRCs.

(23) He who/?that kisses Maria is a lucky man.
(24) She who/?that leads must know what she is doing.

Thus, while Donati and Cecchetto (2011) suggest that Romance pseudorelatives do not involve WH-movement as a first step, I will maintain WH-movement in my analysis, and generally work with the understanding that pseudorelatives in Romance languages are structurally non-identical to English PRCs, despite their surface similarities.

In this section, I have reviewed Postal's (1966) proposal that pronouns are determiners, Citko's (2004) analysis of light-headed relative clauses and their similarities to PRCs; Elbourne's (2013) and Zobel's (2015) account of PRC; and demonstrated that PRCs are

[3]It is not clear what different type of movement this is from the analysis in Donati and Cecchetto (2011); just that it is distinct from true relative clause head-raising.

not comparable to Italian pseudorelatives. In the next section I will detail the properties of PRCs.

2.1.2 Properties of PRCs

I analyze PRCs as a type of light-headed relative clause in which a pronoun begins in a nominal layer (nP) within the relative clause, then moves to the traditional relative-head position above the relative CP. PRCs denote generic kinds and are not referential, which motivates my proposal that the pronouns heading PRCs do not, as Elbourne suggests, occupy D as a definite determiner. Instead, I aim to capture Zobel's analysis of PRCs as generic kinds by leaving the pronominal head in the NP projection, and combining it with an external generic determiner. In this section I describe the propertiess of PRCs, including examples from a corpus study establishing their grammaticality and productivity.

Despite claims to the contrary (as in Postal 1966), PRCs are possible in English, but tend to be associated with a certain literary register:

(25) He who does not study history is doomed to repeat it.
(26) We who are about to die salute you.

The association with the literary register is not, however, necessary for a PRC to be grammatical. A Twitter corpus study Conrod, Tatman, and Koncel-Kedziorski (2016) produced many attested examples that were clearly non-literary in register:

(27) We who #FeelTheBern feel the same about you! [twi.7511]

Conrod et al. (2016) concluded, based on the robustness of novel and non-literary tweets including restrictive PRCs, that the construction is productive in contemporary English. Conrod (2017b) is a syntactic analysis of PRCs that compares them with relative clauses headed by quantificational elements (Quantificational Relative Clauses, or QRCs,

compared to PRCs in (16). In Conrod (2017b) I found that nominal modifiers to the heads of PRCs and QRCs behaved differently: the semantic effect of nominal modifiers like *first* reconstruct into embedded relative clauses that were headed by pronouns, but not those headed by quantifiers (16b); below I show the relevant reconstruction facts that motivate a raising analysis for PRCs.

Bhatt (2002) uses adjectival modification to detect reconstruction of relative heads into embedded relative clauses. (28) below shows his application of these diagnostics; in cases where the scope of the adjectival modifier can be low, he takes this as evidence that the relative head (and its modifier) originated in that low, RC-internal position.

(28) The first book that John said Tolstoy had written was *Anna Karenina*.
   a. Yes, that's true, John talked about *Anna Karenina* first.
   High reading: *In 1990, John said that Tolstoy had written Anna Karenina; in 1991, John said that Tolstoy had written War and Peace.*
   *(I.e., order of **saying** matters, order of **writing** is irrelevant.)*
   b. No, that's not true, Tolstoy wrote *War and Peace* first.
   Low reading: *According to John, Tolstoy wrote Anna Karenina first—which is false.*
   (I.e. order of writing matters, order of saying is irrelevant.)
   (adapted from Bhatt 2002:57)

These high and low readings are slightly more difficult to apply to PRCs, since the generic readings must be maintained for PRCs without external determiners to be felicitous.[4] In (29) I use *first* as a modifier modeled after those in (28); likewise, the superlative modifier must be licensed by a definite determiner.

(29) The first she that Adam said he had married was Eve.
   a. Yes, that's true, Adam said that before later admitting to his history with Lilith.

[4]The example I include here includes a determiner rather than the pure generic reading of bare PRCs; this is because any adjectival modifier of generic PRCs that is useful for differentiating readings will render the PRC non-generic.

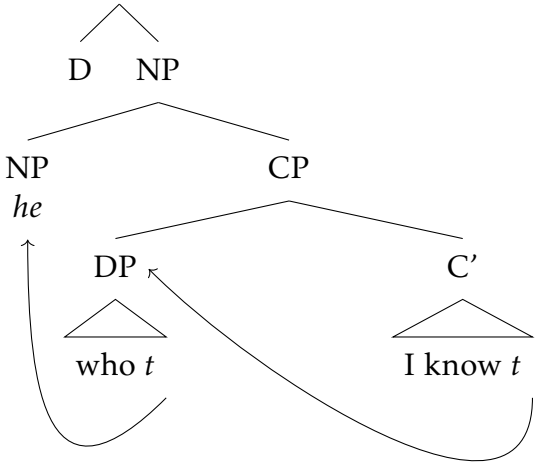
High reading: *Adam said earlier that Eve was his wife, then later said that Lilith was his wife.*

b. That's not true, he was married to Lilith first!
Low reading: *According to Adam, Eve was the person he married first—which (depending on your beliefs) is actually false, regardless of whatever Adam says about it.*

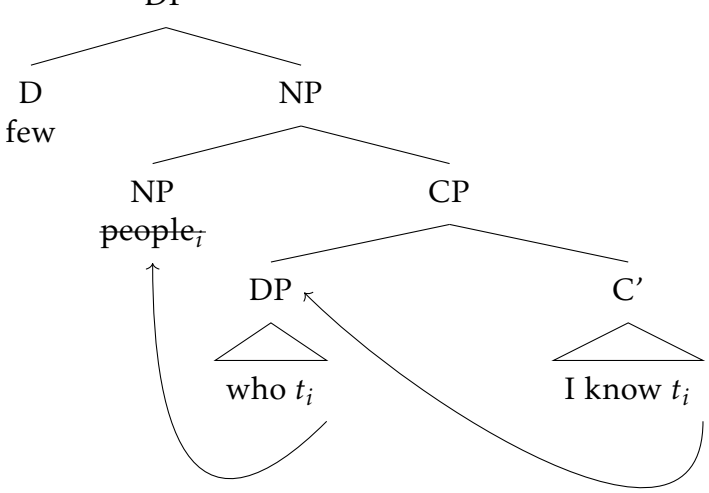
As with (28), the high and low readings in (29) give different truth conditions: in a high reading, the adjectival modifier applies to (the direct object of) *said*, whereas in the low reading *first* modifies the complement of *married*. The availability of the low reading in (29) suggests that the reconstruction facts for PRCs are similar to those of fully-headed relative clauses.

Based on this evidence, I proposed two different analyses for PRCs and QRCs: pronominal heads always undergo head-raising to N (or *n*), while quantificational heads sit in D and are merged externally to the relative clause. When reconstruction effects do obtain for QRCs, the moving element is not the quantificational "head," but rather a null *pro* that sits in a nominal position; the other possible underlying structure for QRCs involves nominal ellipsis, where an N is deleted under identity with sufficient discourse antecedence. (30) shows the structures I proposed in Conrod (2017b):

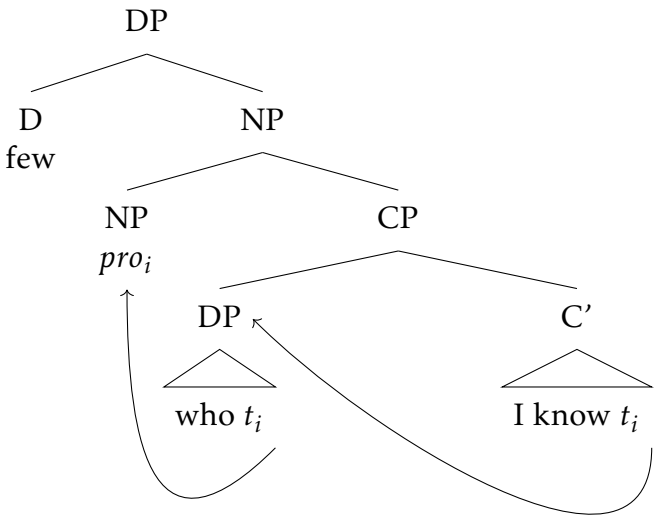
(30) a. PRC



b. QRC: deleted N



c. QRC: null *pro*



I leave aside the quantificational relative clauses for the purposes of this dissertation except to say that they contrast with PRC in the position of the head: there is ample evidence that the heads of quantificational RCs do base-generate in D, while that evidence is absent for PRCs.

Also against Zobel's claims, I show that the heads of PRCs are not restricted to *he*, but can in fact be any pronoun except for *it*. The examples in (31) below are all taken from the corpus study in Conrod et al. (2016).

(31) a. But value **he** who shows you respect, honesty & trust. [twi.310]

b. A preacher that fears the powers that are contemporary, and dismisses the power of **him** who is eternally in power, is not fit to lead people [twi.4673]

c. **She** who leads rules, so play nice or I won't let you [twi.1455]

d. Every moment is a golden one for **him/her** who has the vision to recognize it as such. [twi.5951]

e. **they** who control the pumpkin spice control the universe [twi.2610]

f. It's a funny old game ain't it but **them** who take part in it wouldn't change it [twi.6439]

g. And **we** who know and realize this should always be willing and eager to save others and not condemn them [twi.7637]

h. u can't tease **us** who weren't there with a new song and not let us hear!!!! [twi.9453]

i. **you** who make me smile, you are what makes my heart [twi.8765]

Note that the pronominal heads in (31) include not only third person pronouns, but also first and second person pronouns. While this dissertation is primarily concerned with third person pronouns, in Chapter 5 I suggest that person features are linked to, but not to be conflated with, reference; in the case of $1^{st}/2^{nd}$ pronouns, the person features when divorced from direct reference can be construed predicatively in a way that is closely analogous with gendered or honorific pronouns.

There is a confluence of features that motivates analyzing pronominal heads of PRCs as predicative and nominal: first, the inability of the pronominal heads to refer to any particular referent; and second, the possibility of a restrictive reading of the relative clause even when the pronominal head is singular.

Semantically, the heads of PRCs are obligatorily generic, and the PRCs as a whole denote generic kinds (Zobel 2015). Heads of PRCs demonstrably resist reference, either through syntactic antecedence or pragmatic prominence:

(32) a. *Ali$_i$ is going to the grocery store. He$_i$ who has many things to buy should bring many bags.

b. **I'm pointing at my friend Lydia$_i$ standing next to me.* You should respect she$_i$ who respects you.

In (32) the coindexed pronouns cannot get a restrictive reading when heading relative clauses. (The * in (32)) indicates the unavailability of this reading.) In order to be heads of restrictive relative clauses, pronominal heads must be restrictable–meaning they cannot be referential. The generic reading of pronouns is necessary for restrictiveness.

Zobel (2015) argues that all relative clauses headed by non-*he* pronominal heads are nonrestrictive. Some PRCs that Conrod et al. (2016) found are ambiguous between restrictive and non-restrictive readings, but the fact that it is possible to interpret the relative clause restrictively is an important point. Relative clauses can be restrictive when

headed by any pronominal head (excepting *it*), including pronouns which would normally have indexical readings:

(33) a. He who angers you controls you.
There is some type of (male) person. There is a general sub-type of that type, which is people who make you (general) angry.

b. You who give to the poor are the real heroes.
There is some type of person (who are the general audience of this statement). There is a general sub-type of that type, which is people who give to the poor.

c. We who are about to die salute you.
There is some type of person (of which I am an example). There is a general sub-type of that type, which is people who are about to die.

Singular pronouns are interpretable as restrictive only if they have a generic reading. They may not have syntactic antecedents or refer to a specific individual in the discourse.

The uses of *he*, *she*, and *they* are all compatible with Zobel's (2015) semantic analysis: so long as they are generic and not referential, there is no reason why these third-person pronouns should not be able to be part of generic-kind statements.

The fact that PRCs are indeed restrictive is strong motivation for an analysis where the pronoun sits in a lower position than D. Wiltschko (2012) analyzes different types of relative clauses, including restrictive and non-restrictive relative clauses, as being adjoined to different levels of the nominal projection. In order for a restrictive reading to be available, the RC must be adjoined to NP, not DP. In Section 4 I show a syntactic structure and a related semantic denotation that can account for the reconstruction in PRCs (and not in QRCs), the ability for various pronouns to head PRCs, and the requirements of genericity/non-referentiality.

2.2 Depronominalizations

In this section I discuss depronominalizations–constructions where pronouns are 'demoted' to act more like common nouns through combination with external determiners,

as with the indefinite (34) and definite articles (35) below. I first briefly review previous work that mentions depronominalizations, then show their properties.

(34) I need a she who can trust
(35) I meant the other she

While the use of *a she* in (34) appears to be connected to the appearance of a relative clause, external determiners can also co-occur with pronouns outside of a PRC; *the other she* in (35) also shows the depronominalized use of *she*, without a relative clause. Depronominalizations frequently, but not exclusively, include at least a determiner and some other nominal modifier (such as a RC or adjective). This section investigates what conditions allow a determiner to appear before a pronoun in English.

2.2.1 Previous Research on Depronominalizations

The literature on English pronouns occasionally discusses depronominalizations, usually as an aside. I will discuss what has been said about the construction generally, and review one proposal that argues that these constructions are in fact a type of category-switching.

Cardinaletti (1994) responded to Postal's (1966) analysis of English pronouns as Ds, which also includes data from Italian. She identifies two different types of pronouns in Italian, for which she proposes two different structures: 'strong' pronouns, are merged at N and combine with D to form a full DP, while 'weak' (clitic) pronouns are simply D heads with no complements.

Cardinaletti shows several empirical asymmetries between 'strong' and clitic pronouns. Strong pronouns are not cliticized, are free morphemes, occur in the same argument positions as full DPs (so post-verbally in Italian), and don't undergo clitic movement. Strong pronouns also are able to appear in A' left-dislocated positions and in isolation. Strong pronouns in Romance generally do not appear with determiners, and when they appear with modifiers strong pronouns precede the modifier. (36) shows the distribution of strong pronouns with respect to the position of the verb (Cardinaletti 1994).

(36) a. Conosco lui
[I] know him.
b. * Lui conosco.

Weak pronouns, on the other hand, are phonologically reduced and dependent; they don't occur in argument/DP positions but instead undergo clitic movement and appear pre-verbally; they cannot occupy left-dislocated (A') positions, nor appear in isolation. (37)) shows a summary of these properties, which constitute apparently complementary distribution with strong pronouns (Cardinaletti 1994).

(37) a. * Conosco lo.
[I] know him.
b. Lo conosco.

Based on the properties shown by Cardinaletti (1994) English pronouns are all strong pronouns; the head-raising analysis that Cardinaletti suggests for this type will be the basis for my proposal in Chapter 5.

In their analysis of the featural geometry of English determiners, including personal pronouns, Cowper and Hall (2009) observe that personal pronouns can be used predicatively and be modified or appear in compounds, essentially behaving like nouns. Their examples of these predicative pronouns are of the subtype I call depronominalizations. Their examples are similar to the ones I present in the next section:

(38) a. a **she**-wolf
b. a **he**-man
c. Is that a **he** or a **she**? Neither; it's an **it**.[5]
d. "Lady, you are the cruell'st **she** alive,/ If you will lead these graces to the grave/ And leave the world no copy." (William Shakespeare, Twelfth Night I.v.241–243)

[5]The fact that *it* can appear in a depronominalization but cannot head a PRC may be related to the fact that relative heads are focused, and *it* resists being focused. If there is another reason for this asymmetry I haven't thought of it.

e. "The **theys** are not individual **hes** and **shes** with votes." (Jenkins 1973: 1015)
f. It just looked absolutely **us** somehow.
g. " 'Roses are worth more dried than alive'—such a **you** thing to say./ O! how I adore you when you reinvent a rosy cliché." (The Tragically Hip, "Impossibilium")
h. "You say to **me**-wards, your affection's strong;/ Pray love me little, so you love me long." (Robert Herrick, "Love me little, love me long")
i. Mini-**Me**
j. the **Me**-Decade, the **We**-Decade

(Cowper and Hall 2009: 12)

Cowper and Hall propose that these predicative depronominalizations are inserted in N, essentially used as nouns, and do not spell out grammatical features in higher functional projections in the */phi*P (which, in their system, constitutes personal pronouns in English generally). Instead, they propose that the semantic denotation of these depronominalizations is de-grammaticalized, so that *she* spells out the semantic property *female*. In the proposal I make in Section 2.4 I will expand on this, showing that the effect of de-grammaticalization can in fact be derived from the underlying syntactic structure of pronouns (and DPs more generally). Melchin (2015) deals with English depronominalizations. He argues that English pronouns with articles are examples of category-switching, providing evidence that English can have pronouns in an N position. He compares languages with D-pronouns like English, German, and Mandarin with languages with N-pronouns like Japanese; in doing so, he suggests not only that Mandarin pronouns *occupy* D, but that they cannot have undergone head-movement to get there:

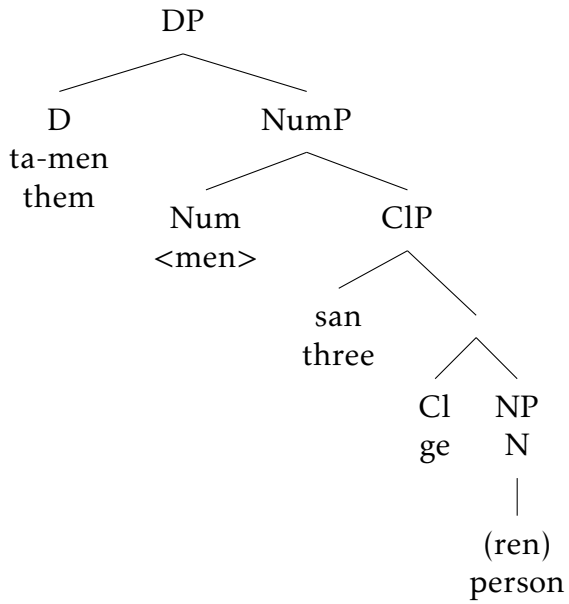
(39) Wo dui **ta-men** san-ge (ren) tebie hao
I to them three-CL (person) especially nice
(lit.) 'I am especially nice to them three.'

(40) nan-guo de **ta-men** mei-you qu can-jia hun-li
sad DE they didn't go participate wedding
Sad them didn't participate in the wedding

(41) * nan-guo de **ta-men** san-ge mei-you qu can-jia hun-li
sad DE they three-Cl didn't go participate wedding

Sad them three didn't participate in the wedding

(42)



(Melchin 2015:5)

Under the structure Melchin shows in (42), head movement from N to D is blocked by the intervening Num head, which is itself what motivates analysis of Mandarin pronouns as base-generated in D.

He instead suggests that while Mandarin (and English) pronouns are base-generated as D heads, they can be coerced to appear in N as a type of category-switching. He notes that this type of coercion can apply to a variety of different categories in English, where function words can be used as (apparently) common nouns in certain circumstances:

(43) This dictionary is a must. (Melchin 2015:8)

Melchin calls this process of coercion in English *"idiomatic and non-productive"*; my work in attestations of PRCs and depronominalizations in corpora call this generalization into question (Conrod, 2018b; Conrod et al., 2016).

My proposal (in this chapter and expanded throughout this dissertation) does not have pronouns base-generated as (lexical) N heads, however. Instead, I will propose that

pronouns enter the derivation at n. If this is the case, then it is possible to rescue the head-raising analysis for Mandarin: instead of raising from N (which would be blocked by Cl anyways), NumP is dominated by a *n* projection, which is where pronouns like ta would enter the structure. If this is the case, then *men* can raise as far as the *n* head *ta* without actually raising to D.

The advantage of my alternative proposal is that it correctly predicts the possibility of certain restrictive modifiers to pronouns in Mandarin which would otherwise look non-restrictive.[6]

(44) ni-de ta
your him/her
(idiomatically,) your person / your boyfriend/girlfriend

(45) ?* ni-de ta-men
your them
*your those people

(46) * ni-de ta-men san-ge (ren)
your them three people
*your those three (people)

Melchin (2015) predicts that Mandarin pronouns, being strictly D-pronouns, should not be modified except through coercion; my analysis however predicts that even apparent D pronouns should be modifiable if they are blocked from successfully head-raising from n to D. If it were the case that *ta* was base-generated in D, then (44) should be equally ungrammatical as (41); even more interestingly, an informant suggests that there is a construal where (46) is grammatical (with or without ren): *" [(46)] may be ok if you and I are each in charge of some people, and ni-de means [ni guan de; 'who you supervise], as in [wo ba ni-de ta-men san-ge (ren) zhao wo dao le; 'I found those three people who you supervise]. I accept (45) as neither polyamorous (44) nor uncounted (46)."* Additionally, this informant reports, *"[ni (guan) de] feels like a restrictive relative clause—you might be in charge of three*

[6] I thank Edith Aldridge for the suggestion to explore this, and I am very grateful to Ang Li, Naomi Lee, Celia Liu, and Chung-chieh Shan for judgments and discussion on these data.

people and I two other—but if you're actually in charge of four people then the sentence feels infelicitous."[7] The fact that (44) is accepted by most speakers - and that (45) and (45) are only acceptable when interpreted as modified by reduced relative clauses - supports my analysis of Mandarin depronominalizations as pronouns in a lower, nominal layer like nP.

The fact that Mandarin pronouns can be modified by restrictive relative clauses is in fact evidence that the pronouns start the derivation structurally lower than D; it is not necessary to say that they are all the way down in N, however, as an intermediate functional projection in the nominal domain can more readily explain the differences between referential and restrictable pronouns. Moreover, this allows us to maintain the insights from pronouns that act mostly like D but not always.

Melchin does not propose an exact theoretical mechanism for how coerced category-switching; my account in Section 2.4 and moreso in Chapter 5 will suggest that intra-domain category switching is indeed derived by head-movement, and that extra-domain category switching may be related to the way that categorical heads combine with (category-free) roots.

2.2.2 Properities of Depronominalizations

Depronominalizations are attested and productive; my previous work on a Twitter corpus suggests that they are sociopragmatically strongly gendered, and often used to contest or assert the gender of a person (Conrod 2018b).

Depronominalizations initially appeared in a different data set, where I was not looking for them. In investigating the pronouns and names used to discuss a publicly well-known transgender woman, Chelsea Manning, speakers used depronominalizations to explicitly comment on the gendered pronouns used by others.[8]

---

[7] Commentary from Chung-chieh Shan, to whom I am very grateful for the discussion!

[8] I more thoroughly discuss the data I collected on Chelsea Manning in Chapter 4, where I connect pronoun use and misgendering with stance.

(47) Please stop referring to Chelsea Manning as "him", **she is a "her"**.

(48) Bradley Manning **is a HE**, regardless of whether he likes it or not.

In this chapter I will propose a complex internal structure for pronouns cross-linguistically that will account for both their N and D-like behavior, and which can be easily parameterized to account for cross-linguistic variation without suggesting that pronouns are categorically differentiated between languages.

Depronominalizations in English can form around almost any pronoun; *he* and *she* are quite common, and as I have shown are closely linked to discussions of gender (Conrod, 2018b). First and second person pronouns, such as *you*, are also attested in depronominalizations; *it* seems to be the only pronoun that is relatively rare, in this construction. Below are examples from a corpus of Twitter data showing the productivity of depronominalizations involving *he* and *she* Conrod (2018b).

(49) a. My person is **a he**. He'll always be remembered and I daresay till the day I die unless amnesia kicks in then forgive me. [2370]

b. he hid. if he is **a he** and if a cockroach has a gender. he hid. i know he did. i cant find that cockroach now but i know he's ready 2 attack [5094]

c. you fucking sexist hippocryte! nothing other than the tutu suggests it's **a she** or it needs a lift. and I like it! [7074]

d. He should learn to put his lipstick on better if he is going to be **a she**. [6880]

Semantically, depronominalizations are similar to pronouns in PRCs in that they cannot be referential. Instead, they act as nominal predicates, where the semantic content is quite general and always related to the phi-features of the pronoun (such as gender and animacy). Depronominalized pronouns can combine with a number of different determiners, including indefinite articles and demonstratives; they can also be modified by adjectives, and can be pluralized.

(50) a. * Laura$_i$ is our professor. ... **The she**$_i$ will write the final exam.
Can't be referential

b. What type of person do you want to meet? ... I want to meet **a she**$_x$**.**
Refers to type by gender features

Depronominalizations tend not to appear with the definite determiner alone, which may be an effect of redundancy. I will propose later in this chapter, following Elbourne (2013), that (referential) pronouns are made from a nominal predicate combined with a definite determiner, which implies that pronouncing the non-combined version is less efficient. When depronominalized pronouns are modified by an adjective or a focus marker, however, they can combine with the definite determiner perfectly well if given sufficient context.

(51) a. ? I want to meet the she.
b. *Speaker A:* I have a she and a he you could date, since you said you were open to experimenting
Speaker B: I want to meet the she.
c. I want to meet the other she.
d. Anna is the only she I want to meet.

Conrod (2018b) discusses primarily *he* and *she* in depronominalized contexts, connecting their use to metapragmatic comments about the gender of pronouns and referents. For the rest of this chapter I will focus on these depronominalizations, using attested examples from that corpus whenever possible.

2.3 The Structure of Predicative Pronouns

In this section I propose a syntactic structure and semantic denotation for predicative pronouns that explains both the phenomena of PRCs and depronominalizations. The first subsection will address the syntax, focusing first on PRCs and then showing how such an analysis can also account for depronominalizations. The second subsection will focus on the semantics of predicative pronouns, again first showing the full denotation for PRCs and then applying it to depronominalizations. Finally, I will explore a similar proposal made for proper names to show how the similarities are in fact theoretically

desirable and contribute to the parsimony of a general theory of pronouns, and of lexical and functional categories in the nominal domain.

2.3.1 Syntax of Predicative Pronouns

In Conrod (2017b) I adopted a raising analysis of relative clauses of the type that Bhatt (2002) proposes, in which a head (in this case the pronominal head of the PRC) head is merged within the relative clause, undergoes A' movement to the Spec,CP of the relative clause, then projects a NP layer outside the relative CP.

In Section 2.1.2, the example in (29) showed that pronominal heads of PRCs show reconstruction effects that suggest they started the derivation within the RC, then moved to their surface position. However, not all heads of LHRCs show these reconstruction effects. QRCs in English are very similar to PRCs on the surface, but the demonstrative and quantifier heads do not reconstruct.

(52) Few who Bernie said were true socialists attended the Antifa demonstration last Tuesday ...

a. ...in fact, almost no one was there met Bernie's qualification for a "true" socialist, although the protestors claimed otherwise.
High reading: *Of the people who attended the demonstration, Bernie says not many of them were true socialists. Whether they ARE socialists or not is irrelevant, Bernie's opinion is at issue here.*

b. # ...in fact, almost no one there was a socialist of any kind—many of them were Alt-right protestors instead.
Low reading: *Of the people who attended the demonstration, many of them were not socialists at all. Bernie's opinion is not the important thing, the important thing is whether the protestors actually ARE socialists or not.*

As shown in (52) above, the 'low' reading as it is given in (52b) is not available when the head of a LHRC is not a pronoun—what this suggests is that the heads of QRCs are Ds that are base-generated outside the RC, more like the Polish data that Citko shows (2004).

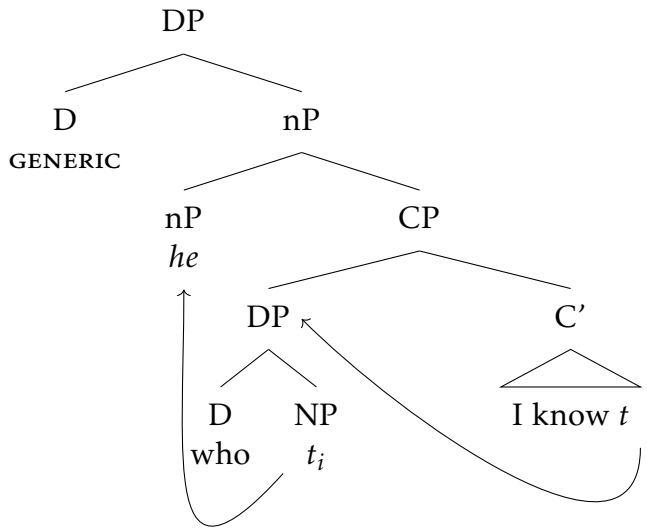
The question remains, then, as to why pronominal heads reconstruct? I here build on Elbourne's empty noun (Elbourne 2013) and claim that, against traditional analyses like Postal (1966), pronouns in English do not consist purely of a D, but rather must be composed of *both* a nominal layer and a determiner. This instead follows Déchaine and Wiltschko's (2002) proposal that in order to receive a non-referential reading a proform must be less than a full pro-DP.

It is the nominal portion of the pronoun (nP) that is base-merged in the low position within the relative clause and undergoes head-raising typical of a fully headed relative clause head; the nominal element projects an NP layer (Bhatt 2002), then combines with a null determiner that denotes the generic reading suggested by Zobel (2015).

Thus, my derivation for a head-raising analysis of PRCs is similar to that in Bhatt (2002), with the pronominal head essentially undergoing the same movements that a lexical relative head would. The important point of my analysis is not the particular of *how* the pronominal head moves to its surface position, but rather *that* it moves; the reconstruction effects obtain for pronominal heads in the same way that they do for lexical heads. Any head-raising analysis of relative clauses should produce the same solution for PRCs: namely, that the pronoun starts in a low, nominal position, and does not at any point behave as a determiner[9]. Below I apply Bhatt's 2002 mechanism of head-raising to a pronominal head to show that, by merging the predicative pronoun in *n*, essentially the same type of relative clause head-raising can occur for PRCs that occurs for any other type of relative clause. The analysis of PRCs that I give here is not dependent on Bhatt's specific method of raising, and should work equally well with any other account of head-raising in relative clauses.

[9]The syntactic structures below show a somewhat abbreviated articulation of the functional hierarchy I assume in the nominal domain for the sake of directing attention towards the relative clause. In Chapter 5 I will extensively show the functional projections in DP that make up pronouns.

(53)

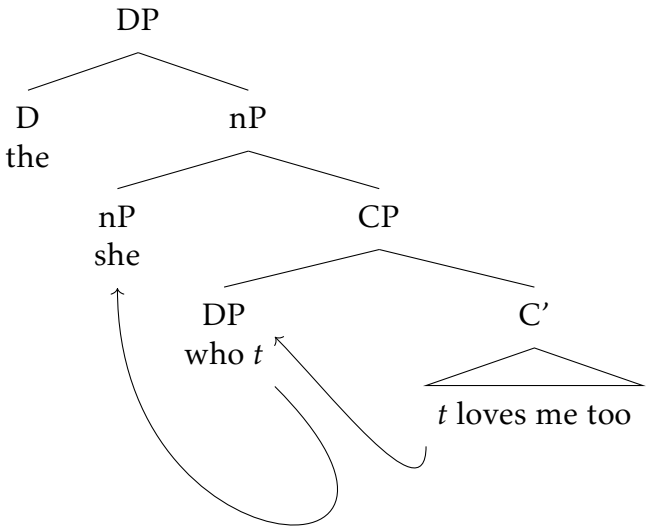


That pronouns are capable of occupying a position structurally lower than D is demonstrated by the existence of sentences in which determiners precede pronouns heading PRCs:

(54) a. A good relationship doesn't need promise, terms and conditions. It just needs a SHE who can TRUST and a HE who can be LOYAL! [twi.1942]

b. the she who loves me too. [twi.1072]

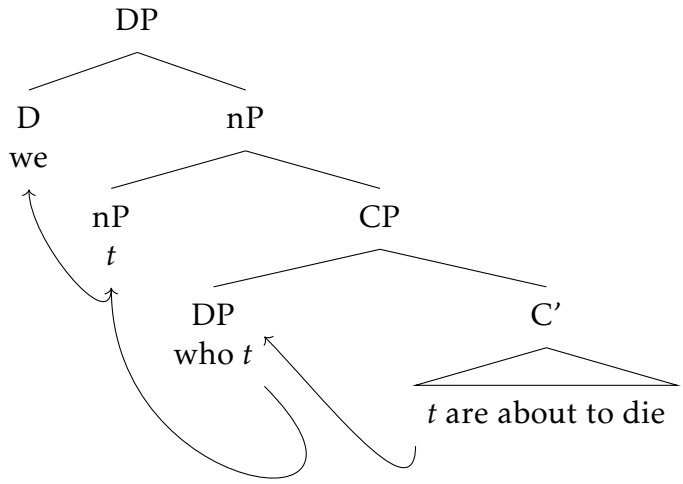
c.



The difference between generic PRCs and the ones in (54) is what determiner is merged after the pronoun raises: if a GENERIC determiner is merged, the usual generic reading of PRCs results; when another determiner such as *the* or *a* is merged, then an (in)definite nominal reading results. In these cases as in (54), *she* still means *woman*, but without the generic kind that Zobel discusses Zobel (2015).

Pronouns originating in *n* and later raising potentially allows first/second pronouns to head PRCs without a generic reading—rather than the depronominalized *she* in (54), a (restrictable) second person pronoun can optionally still move to D if it isn't generic.

(55)

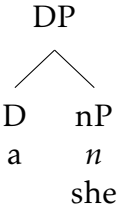


In (55), the pronominal head *we* is plural and restrictable but still starts in a low nominal position before raising to become referential. (Chapter 5 discusses the connection between referentiality and movement to D further.)

This also neatly accounts for depronominalizations: just as in PRCs, the pronoun is initially merged in a low position in the nominal spine, and can combine with an external determiner and even adjectival modifiers independently—which means that the same analysis can be extended to depronominalizations *without* RCs attached (below).

(56) a. I want to meet a she, not a he

b. Oops, I meant the other she

c.


The extension of the analysis of PRCs for depronominalizations shows that 'low' pronouns can be generalized as pronouns that are modified and non-referential—which accounts for the generic readings of PRC heads as well as the predicate reading of depronominalizations. The difference between PRCs and depronominalizations lies in two things: first, PRCs of course have restrictive relative clauses adjoined to the low (n) pronoun while depronominalizations do not; and second, depronominalizations may be introduced by various other external Ds in place of the generic D that licenses PRCs. In the next section, I will discuss what the semantic component of pronouns would need to look like in order for either PRCs or depronominalizations to be possible and meaningful.

2.3.2 Semantics of Predicative Pronouns

Following Zobel (2015) I have analyzed PRCs as denoting 'generic kinds.' In this section I propose a formal analysis of generic kinds that can incorporate Predicate Modification as the semantic mechanism of restrictive RCs (Kratzer and Heim 1998); in doing so I adapt an analysis of pronouns (partially following Elbourne (2013) and Déchaine and Wiltschko (2002)) in which the pronominal heads of PRCs sit lower in the nominal domain (and can thus be modified by nominal modifiers).

I work generally within an adaptation of the system put forth by Heim and Kratzer Kratzer and Heim (1998) in which relative clauses combine with nominal predicates through predicate modification, which essentially captures the intersection of sets.

(57) a. **Predicate Modification** (PM): If α is a branching node, $\{\beta, \gamma\}$ is the set of α's daughters, and $[[\beta]]$ and $[[\gamma]]$ are both in $D_{<e,t>}$, then $[[\alpha]] = \lambda x \in D_e.[[\beta]](x) =$

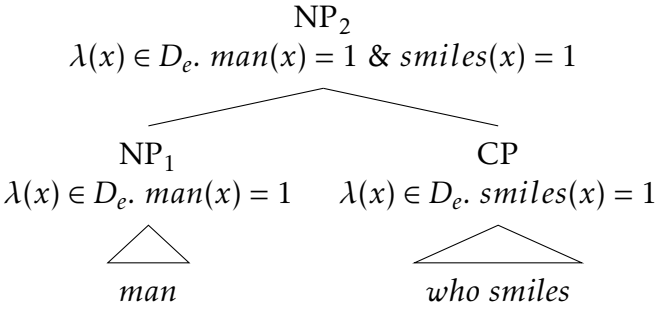
1 and $[[\gamma]](x) = 1$
(Kratzer and Heim 1998)

b. *Paraphrase of PM*: if α is a branching node with daughters $\{\beta, \gamma\}$ then both β and γ have to be true of an entity x for α to be true of an entity x.

Applying PM to relative clauses takes the head noun as the first predicate and the relative CP as the second; for an entity to be a *man who smiles*, that entity must both be a man and must smile:

(58) a. $[[man\ who\ smiles]] = \lambda(x) \in D_e.\ man(x) = 1\ \&\ smiles(x) = 1$

b.



Under Elbourne's analysis of pronouns, predicate modification (PM) should not be able to combine a relative clause with a pronoun's gender features (Elbourne 2013). Elbourne himself posits a null nominal *person* selected by the pronominal heads of his 'Voldemort phrases'; however, the pronoun itself is a (gendered) definite determiner that sits in D. Under his proposed semantics for pronouns, these relative clauses should not be able to combine with the gender features, since the nominal predicate *person* is ungendered and since he analyzes the actual, gendered pronominal heads as definite D heads denoting unique individuals (by means of ι) bound by situation variables (s) as in (60), essentially the same as his denotation for the definite determiner (59).

(59) $[[the]] = \lambda f_{<e,st>}.\lambda s : s \in D_s\ \&\ \exists! x f(x)(s) = 1.\iota x f(x)(s) = 1$
Paraphrase: within a situation s there is exactly one entity x such that x is true for the function (predicate with which the determiner combines)

(60) $[[he]] = \lambda f_{<e,st>}.\lambda s : s \in D_s \ \& \ \exists! x f(x)(s) = 1.\iota x f(x)(s) = 1 \ \& \ \textsc{masc}(x) = 1$
Paraphrase: within a situation s there is exactly one entity x such that x is true for the function (predicate with which the determiner combines) AND x is masculine

The denotation Elbourne give in (60) is for a referential pronoun which, in order to refer, must be definite. There can only be one unique entity that, when *the* is combined with whatever nominal predicate it combines with (a normal NP), is the unique referent that yields truth and situational appropriateness. NPs, being predicates, are of a type <e,t>; thus, in Elbourne's view of pronouns as definite descriptions, *he* essentially acts like "*the* plus masc" where masc is the predicated property (also of type <e,t> like other nominal predicates) of being masculine. Essentially, if *he* in Elbourne's view is itself a definite determiner D, it combines with a null NP predicate and ends up meaning something like *the man*.[10] When Elbourne combines the restrictive RC at the nominal level, he may obtain the generic and gendered reading only by the same means that definite NPs obtain a gendered reading–which means that the content of the RC does not actually directly combine with the gender via PM. Elbourne's analysis of pronouns captures one aspect that I have argued for—that a nominal predicate must be contained somewhere inside the pronominal DP in order to both refer uniquely and denote a property of its referent; his denotation however is not exactly compatible with either my or Zobel's analysis of PRCs, where the pronominal heads are crucially *not* unique or referential. Thus I propose a very minor adjustment to Elbourne's proposal, informed by the insights from Déchaine and Wiltschko (2002). Déchaine and Wiltschko (2002) claim that pronouns may project differently-sized layers of nominal structure which may add up to less than a full DP. I here notate my pronouns as nPs but they are essentially equatable with ϕPs under the system used in Déchaine and Wiltschko (2002). Effectively, this means that for referential pronouns, Elbourne's definite determiner is still *there*, but for predicative pronouns of the types I have shown, the masc predicate is in fact itself the pronoun, which combines with a different determiner (not a definite article).

[10]This section is one of many that benefitted greatly in discussion with Edwin Howard about semantics.

Restated briefly, my analysis swaps what Elbourne calls null (the nominal predicate) and what he places in the D position. Elbourne's structure for pronouns (as he roughly sketches it in (61a) and mine in (61b) otherwise have the same components. The advantage of mine is that, by having the pronoun enter the derivation in place of a nominal predicate, different types of pronouns may combine with different types of determiners – which also conveniently predicts the possibility of depronominalizations and PRCs.

(61) a. $[he[_{NP}person]]$
b. $[the[_{nP}he]]$

Since the pronominal heads of PRCs are obligatorily non-referential (and cannot be coindexed with anything), I will not treat them as entity-denoting expressions. Instead, I take them to be nPs which, much like other nominal elements, denote predicates. In this case, I do follow Elbourne in my analysis of an essentially empty noun: the nP is intransitive, and does not combine with a root. Instead, the n head hosts ϕ-features. My denotation for this reduced pronominal element is a predicate, so I will be referring to these as predicative pronouns. That denotation is below in (62):

(62) $[[he]] = \lambda x \in D_e$. $\textsc{masc}(x) = 1$
Paraphrase: *an entity who is masculine*

Notice that the denotation for predicative pronouns that I give in (62) is quite similar to that of a noun. This captures the "predicative" property of pronouns that are not fully referential. This denotation captures the generic noun-y reading for depronominalizations like *a he*, and is equally crucial for the correct semantic derivation of PRCs. However, it is important to maintain the distinction between pronouns, which have many qualities of functional categories, and nouns, which are an open class and lexical category. I will capture this difference by proposing not that pronouns are N (lexical) or D (determiner), but rather that they are the intermediate nominalizing functional projection *n*,

which is a 'typing' projection that can combine with category-free roots to constitute a noun (or nominalized element).

In Section 2.4 I will further discuss what it means for an entity to 'be masculine,' as well as giving a proposed semantic contribution from the *n* head that will allow me to differentiate between lexical nouns and pronouns while maintaining a predicative syntax and semantics for 'low' pronouns.

The final question to be addressed is the question of what, if anything, sits in the D head of PRCs. Since they can combine with any other part of a matrix clause it is reasonable to assume that PRCs are indeed full DPs; I depart from Déchaine and Wiltschko (2002) in maintaining selectional consistency.

I will continue to follow Zobel's convincing argument that PRCs denote generic kinds (2015); thus, the question remaining is how generic kind-referring DPs fit into a broader semantic theory. Carlson and Pelletier (1995) proposes two standard options: the first is an analysis of kind-referring DPs as (covert) proper names, and the second (which I adopt) is an analysis of genericity as quantificational. It is interesting to note that, despite otherwise patterning very cleanly with other kind-referring DPs, PRCs do not allow the 'so-called' construction (which other kind-referring DPs do). (63)-(64) below show this contrast; notably, the *non*-restrictive pronominal relative clause in (64)b) does allow the construction - which goes further to establish a clear difference between restrictive and non-restrictive PRCs.

(63) a. The Incredible Hulk is so-called because of his size.
b. Ligers are so-called because they are the offspring of lions and tigers.

(64) a. *? He who laughs last is so-called because he doesn't understand jokes.
b. He who must not be named is so-called because wizards fear his name will summon him.[11]

[11] This reading is valid only in an interpretation where *He who must not be named* refers specifically to the evil wizard Voldemort from the *Harry Potter* novels.

The *so-called* construction that supports naming predicates shows that, while non-restrictive PRCs may be naming predicates if and only if they are associated with a specific individual (and so operate as a kind of psuedonym), restrictive PRCs are clearly not (covert) proper names. It is based on this contrast that I instead adopt a quantificational view of generic kind-referring DPs, which aligns quite nicely with Zobel's analysis. I assume the existence of a (null) quantifier, called GEN, which may co-occur with certain determiners. The formal reflex of GEN is an operator called Gen which takes as arguments an entity (x) and a situation (s). Zobel's final denotation is shown below in (66).

(65) He who pays the piper calls the tune.

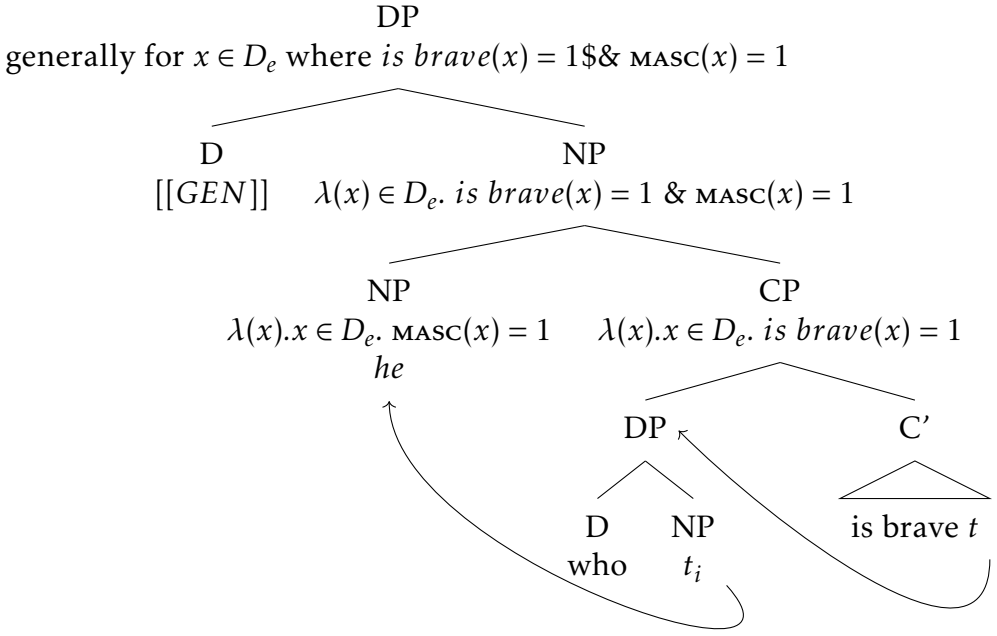
(66) = GEN y[member-of(y, ι X[PERSON-WHO-PAYS-THE-PIPER(X)] ∧ C(y)] → [calls-the-tune(y)]
Paraphrase: for all individuals y (generally) if y instantiates PAYS-THE-PIPER kind and C(y) then y calls the tune
(Zobel 2015:42a)

In employing such a quantifier, I also leave space in the structure for predicative nPs to combine with other types of Ds such as real determiners– something I show in (54) to be attested and grammatical (Conrod et al. 2016). As an abbreviation, I show here structures with GEN in the D position; however it is compatible with other determiners (since DPs with *a* or *the* can also be generic). This D (with the GEN and with or without any other determiner element) can combine with a predicative pronoun nP, which itself can combine with a restrictive relative clause through PM.

Combining the denotation of predicative pronouns, including their ϕ-features, now allows us to expand Zobel's analysis to account for gendered readings and for pronominal heads other than *he*. Thus the structure below shows the denotation at each Merge, after which the Gen operator can apply, yielding the end point at Zobel's denotation (2015).

(67) a. [[*he who is brave*]] =

b.



≈ people who are brave and male

In the structures that I have shown here, I am claiming that pronouns can be predicative, can be modified by things that modify nouns, and therefore must occupy a noun-like position (and have a noun-like semantic role) at least some of the time. The syntactic and semantic structures I present are designed to reflect this claim and appropriately predict all the facts.[12]

In the next section, I will return briefly to the similarities between my proposal for pronouns and Matushansky's (2015) proposal for proper nouns, and explain why those similarities are in fact well-motivated and expand our understanding of pronouns and proper names more generally.

---

[12]The analysis I have given in this chapter is in some ways in conflict with many of the previous arguments for pronouns to be categorized as (and start the derivation in) D. One of the main arguments for this, as pointed out by Postal (1966), is the existence of phrases like *we linguists*. However, the availability of *restrictive* RCs that modify pronouns opens the possibility that pronouns can be modified by *reduced* restrictive relative clauses. I have only seen replies to Postal that consider the possibility that the noun in *we linguists* is an appositive (reduced) RC, which would not be restrictive; that yields various problems with empirical evidence of course.

2.4 Names and Pronouns

It seems intuitively true that names and pronouns should be two sides of the same coin: both are almost always used to refer, often directly and without further qualification, to entities under discussion in the discourse. They are syntactically similar, too, in that they typically resist modification by nominal modifiers or determiners. In my work here and in Matushansky (2015), however, we have seen that pronouns and names are not always immune to modification - and that when they are modified, they exhibit different properties than in the unmodified uses.

I propose here that the reason for these similarities in empirical properties is because of an underlying structural similarity: both names and pronouns are initially merged low in the nominal domain, and must raise up to D if they are to be referential. This head-raising account has been proposed by Matushansky (2015) for proper names, and by Cardinaletti (1994) for Italian pronouns. The English data that I have presented in this chapter expand that account to include English pronouns.

Beyond the apparent similarities, my account also straightforwardly allows us to derive the crucial difference between names and pronouns: names are lexical, and pronouns are functional. In this section I will review the general properties that differentiate lexical and functional categories, and discuss a corpus study comparing names and pronouns referring to a particular referent. I will then show how the semantic and syntactic structures I give for pronouns in this chapter, combined with (an adaption of) Matushansky's (2015) proposal for proper names, correctly predicts these differences. Ultimately, my conception of what differentiates functional and lexical categories will be reduced to a structural definition that can be determined based on which syntactic elements are present or absent.

I will focus primarily on two properties of functional categories differentiating them from lexical ones, both observed by Abney (1987). First, functional elements are closed classes; second, functional elements lack descriptive content. These two properties can

clearly differentiate names from pronouns.

First, 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. 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 versus 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 rather; it allows the speaker to propose that certain entities are male).[13]

In Conrod (2017a), a corpus study of names and pronouns used to refer to a transgender referent, I showed that people more readily adopted the referent's name than their preferred pronouns. By focusing on one particular transgender referent (Chelsea Manning, a 2018 candidate for Senate), the data I collected constituted an apparent-time 'snapshot' of what may have potentially been the amalgamation of many speakers' changes in apparent time.[14] This is only possible because Chelsea Manning attained public notoriety before undergoing her gender transition. Speakers all (likely) started out

[13]Chapter 4 delves much further into why I am casually referring to the gender of a pronoun as propositional; put simply, speakers can use gendered pronouns to introduce, contest, imply, or otherwise do pragmatic work with gender as a social relationship.

[14]The Apparent Time Hypothesis states that change over time can be inferred from differences in synchronic data; I discuss this concept more deeply in Section 3.1.1 in Chapter 3.

referring to Manning as *Bradley* and *he* before her gender transition, and many have since switched to calling her *Chelsea* and *she* over time; however, crucially, not all speakers have switched over, and there is an asymmetry between those using her feminine name versus pronouns. If pronouns and names are functional and lexical (respectively), then we would expect the course of this change to occur at different rates (Muysken, 2008).[15]

If it were the case that speakers generally move from using all male forms of reference (including both names and pronouns) to all female forms, then there would not be a disparity between the uses of first name and pronoun. The predicted values shown in Table 2.1 below show no disparity: these token counts are predicted based upon the assumption that either people think Manning is a woman, and use *she* and *Chelsea*, or they think she is a man, and use *he* and *Bradley*. Predicted values in Table 2.1 are proportional to the sample of tweets collected overall (104).[16]

	She, her	He, him
Chelsea	51	0
Bradley	0	51

Table 2.1: Predicted values of pronoun and name distribution

The actual results, however, do show a discrepancy: there were more instances of people using Manning's new name than those using her new pronouns. 2.2 below shows the actual token counts of tweets using each combination of name and pronoun.

[15]This study is predicated on the assumption that pronouns and names are equally associated with in the minds of the speakers in this corpus; this may or may not actually be the case, but it is safe to assume that the names Chelsea and Bradley are at least strongly gendered based on the observation that many of the speakers overtly commented on Manning's transgender status—and many moreso commented who used the name Bradley. Chapter 4 deals much more thoroughly with the question of actual gender associations.

[16] These predicted values are also based upon the assumption that it is equally likely that any given speaker will attribute each gender to Manning, which also turns out not to be true; if, however, we maintained the proportion of (apparent) gender perception, we would still expect 36 tweets using *Bradley+he* and 0 tweets using *Chelsea+he*.

	She, her	He, him
Chelsea	67	16*
Bradley	0	20

Table 2.2: Actual values of pronoun and name distribution

The discrepancy (such that people did not proportionally use female pronouns based on their use of the female name) resulted in tweets where there was a mismatch in the same sentence, as in (68) below.

(68) Chelsea Manning can change her name legally but he is still a man
(Conrod 2017a)

Along with the more traditional differences between lexical and functional categories discussed above, this asymmetry suggests that speakers' use of pronouns 'lags behind' their use of names in instances where they (might) change over time. I therefore take the conclusion of Conrod (2017a) to support the categorization of names as lexical and pronouns as functional. Below, I will briefly show how Matushansky's (2015) proposal can be adapted to be compatible with the one I have made in this chapter.

Matushansky takes the semantic composition of proper names to be complex, consisting of a referring element and a predicative element; so far, this aligns very closely with my proposal. Her semantic denotation for the predicative element of names is necessarily slightly more complex than pronouns. She denotes the predicative portion of names as a *is-called* predicate (R), shown in (69) below. In Matushansky's formal denotation, the *is-called* predicate R takes one argument which is the phonological content of the name itself, and another argument which is contextual type of naming convention under discussion (e.g. *baptized*, *nicknamed*, *known as*); she calls the most basic, unmarked naming predicate R_0, where the naming predicate argument of the R function is assumed to be whatever convention is most relevant in the discourse context.

(69) $[[Magritte]] = \lambda x \in D_e \,.\, R_{<e,<n,t>>} \,.\, R(x)(\text{"}Magritte\text{"})$
where n is a sort of the type e (a phonological string or some other PF output representation)

(Matushansky 2015:340)

(70) $[[\ \sout{\text{the}}\ Magritte]] = \iota x \,.\, R_0(x)(\text{"}Magritte\text{"})$

(Matushansky 2015:340)

Due to the apparent arbitrary nature of names, it cannot be said that a person named *Brent* has any inherent property of *brenthood*; instead, what is being predicated is the social convention of reference to the person under discussion. Thus, the denotation neatly explains the possible confusion when that *is-called* predicate picks out a set of more than one person:

(71) the two Brents that live in Seattle
(72) Both Brents walked into the room.
(73) Which Brent do you know better?

Previously, I have referred to pronouns denoting (predicative) properties of referents; this has been a shortcut, but it is now time to dispense with it. Along with the apparent difficulties of analyzing the data from Conrod (2017a), Chapter 4 will introduce much more significant problems for this interpretation of gender. For now, I suggest that the arbitrariness and social agreement involved in names is equally involved in pronouns (which will allow me to draw a much more direct parallel between gendered pronouns and honorific pronouns).

In the following chapters of this dissertation, I will comprehensively show that pronouns, too, denote an *is-called* predicate, which invokes not an inherent property of humans but rather a social convention of calling, referring, naming. This has the added advantage of putting pronouns on a clearer analogy with proper names. Thus, from this point forward for the rest of the dissertation, I am discarding the (predicative) meaning of *he* in (74) below, and adopting instead the paraphrase in (75):

(74) HE = A MALE PERSON

(75) HE = A PERSON WHOM WE AGREE IS CALLED HE[17]

The use of gender features socially may in some ways appear to be a departure from an analysis of pronouns as a functional category (n rather than N); however I draw an analogy here to the use of formal features for honorific/pragmatic purposes in pronouns cross-linguistically; for example, the pragmatic difference between *Usted* and *tú* are not readily reduced to strictly formal features, yet formal features do still contribute to the difference, which manifests most readily in the different verb agreement:

(76) a. Usted **es** ...
You.FORMAL are.3SG ...

b. Tú **eres** ...
You.INFORMAL are.2SG ...

Interpreting "natural" gender as conveying social meaning through the use of formal features abstracts away from maleness as an inherent property of certain people - which much more readily explains sentences like (47). Instead, I analyze gendered pronouns as a grammaticalized reflex of complicated, sometimes-messy social relationships. Much the way that a person can go by different names in different social circles (*Theodore* to my students, *Teddy* to my friends), so too can a person go by different pronouns in different social circles. While this is more rare, it is perfectly well-attested; in Chapter 4 I will discuss situations like this extensively. For the purposes of this chapter, it suffices to call pronouns a kind of 'light noun' spelling out formal features with no root, analogous to light verbs; this adds explanatory power to the presence of n in the functional hierarchy of the nominal domain at all. By linking n to the semantic/syntactic contribution of naming conventions, this analysis of pronouns essentially categorizes pronouns as nameless names; without the 'name' itself, only the formal features remain.

[17] "WE AGREE" is an oversimplification, but I include it here only to keep in mind that 'being called' anything is a social relationship that is contextually bound. Chapter 4 discusses this much more thoroughly.

This analysis of the functional head n may be extended further to lexical nouns: what *n* provides is the semantic operation of invoking a convention of naming; what it combines with is the thing that is so named. In the case of lexical nouns the naming predicate abstracts away from things having properties of BEING a certain object (a thing "is" a pipe) and instead denotes the quality of being CALLED a certain name (a thing is "called" a pipe).

Thus, I expand Matushansky's denotation in (77) to cover all lexical nouns, shown in (78), to capture the notion that things (entities in the semantic model, or objects in the world if you're daring) are CALLED names; I maintain R_0 as the default calling function.

(77) $[[\ \text{~~the~~}\ Magritte]] = \iota x\ .\ R_0(x)(\text{"}Magritte\text{"})$

(Matushansky 2015:340)

(78) $[[pipe]] = \lambda x\ .\ R_0(x)(\text{"}pipe\text{"})$

The denotation shown in (78) is a slight expansion of a typical predication of a lexical noun, by inclusion of the R_0 predicate (rather than *pipe* itself being a lambda function that takes an entity *x*). For some entity *x* in the semantic model, *pipe(x)* is true if it is conventional to call that thing a "pipe."

The separation between an entity and the signifier which invokes it is well explored in semiotics and philosophy (cf De Saussure 2011); the goal of my proposal here is to formalize this relationship within the semantic denotation of nouns and pronouns directly. René Magritte's Surrealist work, *La trahison des images*[18] is an ironic commentary on the gulf between signifier (a painting) and signified (a pipe); the act of representation is part of the core function of language.

In invoking, naming, describing, or depicting any real-world phenomenon, the sign necessarily exists at a removed level of abstraction from that which it signifies. I here attempt to codify the relationship of naming and abstraction into the syntax as a way of denaturalizing the otherwise-easily-conflated concepts of "a thing" and "a name for a

[18] Available through the LA County Museum of Art: https://collections.lacma.org/node/239578

thing." The crux of my proposal here is that this operation exists in pronouns just as well as in fully lexical nouns; the content of the pronoun itself is a radically distilled essence of how an entity can be called, named, or invoked.

Under this proposal, the function of *n* in the syntax is that act of invocation which refers to a convention of naming; this allows us to pinpoint what *n* and the root ($\sqrt{}$) each individually contribute. In cases where the calling convention R_0 takes as one of its arguments some string of phonological representation and/or semantic sense, lexical and proper nouns are the result; however, above I suggested that R_0 is also present in pronouns – apparent especially in predicative pronouns.

Adopting the *is-called* denotation for predicative pronouns dissolves one of the few remaining differences between the underlying structure and meaning of names and pronouns. The only remaining difference, I will argue, is the presence of a root.

The root ($\sqrt{}$) is a container for the linguistic representation, while *n* is a syntactic/semantic operator that transforms the raw representation into a properly syntactic object (SO); under this theory roots have no syntactic category, and instead are categorized by the typing operator that they combine with in the first merge (Borer 2005; Marantz 1999, 2000). This layer of abstraction between phonological representations and syntactic objects is in fact desirable: it allows for a great deal of flexibility in assigning syntactic categories to words, and reduces the amount of information that the lexical entries of roots must encode.

What this allows is a structural differentiation in the syntax that is irrespective of category: functional categories are those 'words' formed entirely by syntactic heads, without incorporating a $\sqrt{}$ at first merge. For pro-forms, this means that the entire domain (nominal, verbal, etc.) would be present, but a $\sqrt{}$ would not; thus, pronouns are nominal structure with no actual representational content in their nucleus. The LF denotations provided by *n* and all the higher syntactic heads on the spine are still intact, however, so semantically what is being denoted is "a thing which is called" (provided by *n*, with number and definiteness and quantification provided in the usual way) but without the

name by which the thing is called.

What, then, differentiates pronouns from each other? English pronouns (and certainly this should be expandable cross-linguistically) are differentiated by case and ϕ-features; thus, those differentiating elements must not be provided only by the $\sqrt{}$, but must also enter the derivation on functional heads of the nominal spine. For now I abstract slightly away from the individual features that compose ϕ (I expand on this in Chapter 5). It should suffice to say that ϕ and case are marked on functional heads on the nominal spine, which includes n, NUM, and D.

Because I have proposed here that n's LF contribution is the R_0 predicate of conventional being-called-something, and because I am proposing that n is still present in pronouns, thus it follows that what a pronoun with ϕ-features such as *he* "means" is "a thing which is called [the pronoun that we use for male humans]."[19]

For another way of abstracting this: the calling-convention being invoked at the point of n is composed of social conventions related to referring to people; in the case of English pronouns, these social conventions are an invocation of social gender. While Chapters 3-5 much more closely inspect what I mean by social gender, for the purposes of this analysis I take gender features as essentially different classes of honorific features. The features Masc and Fem are not reflections of pure and static classifications populated by human referents, but rather these are conventional ways of invoking a type of social relationship that is built upon the social conception of gender.

This is what I have proposed in (75). The formal denotation for a free, referential pronoun and its paraphrase are given below in (79); the denotation for a predicative pronoun is given in (80). The denotations are based upon a conventionalized and abstracted notion

[19]The morphosyntactic features that differentiate pronominal forms by ϕ-feature makeup are deliberately not explicitly stated in this chapter, because in Chapter 4 I will be giving a great deal of evidence to show that formal treatments of gender are going to be inherently problematic. In Chapter 5 I will show how ϕ-features (particularly gender) can be incorporated into the syntax and semantics despite those problems. For now, I will direct the reader towards the morphosyntactic account given by Bjorkman (2017) wherein gender features are optional and non-contrastive. This is motivated by the inclusion of singular *they* in the paradigm of third person pronouns.

of the Masc gender feature (as well as the pronoun's third person and singular properties, which I discuss in detail in Chapter 5).

(79) $[[\ [_{DP}\ he\ [_{nP}\ \sout{he}]\ ]\ ]] = \iota x\ .\ R_0(x)(\text{"}he\text{"})$
(referential, non-predicative) "he" ≈ the person whom we have agreed to call "he"

(80) $[[\ [_{DP}\ a\ [_{nP}\ he]\ ]\ ]] = \lambda x\ .\ R_0(x)(\text{"}he\text{"})$
(non-referential, predicative) "he" ≈ some person whom we have agreed to call "he"

The syntactic difference between the referential pronoun in (79) and the predicative pronoun in (80) is the success or failure of head movement. In (79), the pronoun has moved to the highest projection (D) in the nominal domain, which allows the LF to take it as an entity rather than a predicate; in (80) that head-movement is blocked by the presence of an external indefinite determiner; the pronoun is instead interpreted in its low position, necessarily as a function rather than an entity. It is this difference that explains what other authors have called category-switching (Cowper and Hall 2009; Melchin 2015; Pesetsky 1978 a.o.) – when pronouns act apparently noun-like, it is because they remain in a low, nominal position; when they act determiner-like, it is because they have raised to a determiner position. This, in addition to the agnosticism baked into the category-free root analysis I have used here, accounts for any apparent category-switching.[20]

This proposal captures many different properties of pronouns which I have discussed in this chapter: the syntactic structure and semantic meaning of predicative pronouns obtains; the definiteness observed by Elbourne (2013) as well as the genericity observed by Zobel (2015) are maintained; a clear parallel is drawn between pronouns and proper names (as analyzed by Matushansky (2015)); and the social conventions at play in gendered pronouns are given a clear syntactic/semantic mechanism in the formal theory. Additionally, in this chapter I have made a significant proposal for what *n* contributes, syntactically and semantically; as well as making a broad proposal for how functional

[20] It is important to note that "noun-like" behavior mentioned here does not necessitate that pronouns be lexical Ns – rather, all "noun-like" behavior is compatible with the nP analysis I have given here.

and lexical categories can be differentiated through pre-existing primitives of Minimalist theory.

In the following two chapters I will explore some of the social intricacies of gendered pronouns hinted at in this chapter, both from a view of change over time and from the perspective of synchronic variation that occurs in the English pronominal system. The final chapter will tie together these sites and patterns of variation with the proposals I have made about predicative pronouns in this chapter, and more thoroughly explore the mechanism and consequences of head movement.

Chapter 3

CHANGES IN SINGULAR THEY

In this chapter I turn my attention from the syntactic modification of pronouns to focus on one particular pronoun, *they*, in its singular use. This chapter will constitute a variationist sociolinguistic analysis of the definite, specific use of singular *they* (henceforth **dsT**) and argue for the existence of a currently-ongoing language change in the grammar of the English pronominal system. In this chapter I use apparent time to show trends that suggest an ongoing change; in the next chapter I discuss specific sociopragmatic alternations that suggest the nature of the underlying structure of dsT morphosyntactically. In my final analysis chapter I propose a syntactic analysis of pronouns in English that can straightforwardly account for how the language change is progressing in its current direction, including what role the syntax plays in evaluating sociopragmatic appropriateness of coreference.

The first section of this chapter summarizes the history of broader uses of singular *they*, including with indefinites and with generic definites. I will then focus on definite, specific *they* (dsT) as the newest instantiation in a hierarchy of grammatical possibilities. The second section will lay out the predictions of acceptability of singular *they* in a hierarchy, including predictions about its production in free-flowing speech and predictions about its acceptability in grammaticality judgment tasks. The third and fourth sections explain the methods and results of two experimental investigations into singular *they*: the first experiment involves an apparent-time analysis of a corpus of spoken pronouns, and the second experiment is an online survey of acceptability of different forms of singular *they* compared to other pronouns. In the fifth section I discuss the results of the experiments presented, and lay sociolinguistic theoretical groundwork that will help to

constrain the pragmatic and syntactic theory of the chapters that follow this one.

3.1 The gradient acceptability of singular they

Bjorkman (2017) reports varying levels of acceptability of singular *they* with different antecedents. The examples below are modeled off of Bjorkman's, but with some adjustment in their categorization which will be more useful for my own investigation.

(1) **Someone** ran out of the classroom, but **they** didn't bring **their** backpack.
Generic, indefinite antecedent

(2) **The ideal student** completes the reading before class, but not if **they** have a family emergency.

(3) **The math teacher at my school** is very talented, but **they** often forget to grade exams on time.
Specific, definite (ungendered/distal[1]) antecedent

(4) **Jared** is great at laundry and mopping, but **they** always forget to wash **their** dishes.
Specific, definite (gendered?) name antecedent

In Bjorkman's 2017 paper she proposes that sentences like (4) are ungrammatical; this cannot be true for all speakers of English, as (4) is an actual utterance which I collected under naturalistic conditions. However, Bjorkman does make the point that there seems to be a cline in acceptability between (1), which almost everyone except for the most stringent prescriptivists accepts, and (4), which even prominent champions of descriptivism claim is unacceptable (more on this later in the chapter). Definiteness, which separates

[1]Chapter 4 is a closer investigation of what exactly a 'distal' pronoun is in English, and why (3) and (4) should be different at all (considering that they are both referring to specific individuals)—that will be mostly a pragmatic analysis, and does not figure into this chapter's basic investigation of interspeaker variability.

(1) and (2), is only one aspect of what determines acceptability of singular *they* with certain antecedents. Specificity differentiates (2) and (3), and it does seem to be the case that not all speakers who accept (2) necessarily accept (3).

Furthermore, apparent gender, especially as it appears on proper names, seems to draw a clear line in the sand between speakers who accept (4) and those who do not; Bjorkman's own and reported judgments do apparently vary based on the apparent gender of the proper name (e.g. *James*, considered masculine, vs *Hayden*, considered gender-neutral). Furthermore, Bjorkman notes the apparent problem with even non-proper definite NPs, such as *my wife*. Thus, in my acceptability study in this chapter I include names of varying apparent genders—this is a test of Bjorkman's hypothesis that names have syntactic gender features. The results of these data show that there is largely no difference between the acceptability of differently gendered names for speakers who don't accept proper names with singular *they* overall.

Looking at the entire scale from (1)-(4), the antecedents are (impressionistically, noted by Bjorkman as well as Konnelly and Cowper 2017) arranged from the least specific and ungendered to the most specific (and gendered). The first step of this scale seems to be a fairly established one: generic indefinite uses of singular *they* are attested as early as the 15th century (Curzan 2003), and definite generic uses like (2) are the subject of much of the linguistic research I discuss here, of which a fairly large portion was conducted in the 1960s and 70s. The uses like (3) have been shown in some newer studies (as in the 1990s and later) to be either variably acceptable or not widely acceptable at all. Uses like (4) have been claimed to be ungrammatical as recently as Bjorkman's 2017 paper, though forthcoming work contests that assertion (Konnelly and Cowper 2017, Conrod 2018a). These claims by linguists are not at all spurious, and the change in conclusions of researchers over time suggests that there is something still in the works, which has not been completely described yet. This chapter is an early attempt to do so.

I first review earlier studies of singular *they*, many of which start with indefinite and generic uses like (1) and (2). I then move on to definite uses like (3), and finally, definite

specific uses like (4). The eventual goal of this chapter is to investigate interspeaker variation in uses like (3) and (4).

The research on singular uses of *they* is largely focused on two points: asserting its existence, and associating its use with certain syntactic types of antecedents. Hughes and Casey (1986) and others (e.g. Hyde 1984) focus on pronoun production in the context of gender-ambiguous referents in paper-and-pencil tasks. Similar studies, e.g. Moulton, Robinson, and Elias (1978) and Gastil (1990), have also looked at linking gender and pronoun use, but these are more focused on the perception of gender as conditioned by pronoun use—in both, the experimenters used different pronouns to prompt different gendered interpretations. In this section I first review some literature that has looked at generic, indefinite uses of singular *they*, before moving to definite generic uses, and finally definite specific uses.

The indefinite use of singular *they* has the longest history of attestation in English, with generic uses dating back to the 15th century (Curzan 2003). Curzan traces a history of prescriptive grammarians proposing generic use of *he* as a preferable alternative to singular *they* (2003: 59). A 1794 grammar by Lindley Murray posits *he* as a potential 'solution' to the 'problem' of singular uses of *they*. However, this generic use of *he* did not necessarily originate from prescriptive grammars like Murray's. Newman (1997) finds generic *he* and *they* in variation with each other dating back to at least Chaucer; rather, Newman reports that the prescriptive push for generic *he* over *they* is an artefact of the 18th century (Newman 1997: 21). Curzan notes that, despite retaining the morphological markings of grammatical gender, Old English pronouns apparently varied between agreeing with grammatical gender and semantic gender of their antecedents (including when the grammatical and semantic gender were not the same) (2003: 70)—when looking, then, at the pronouns used with indefinite antecedents, Curzan finds that *they*, *he*, *she*, and *he [and/or] she* are all attested in the OE period. The prescriptive grammarian commentary from the eighteenth and nineteenth century therefore appears to be an attempt to reign in already-existing variation, rather than to replace one settled form with

another. While some grammarians explicitly advocated for generic *he* on the grounds of a hierarchy of the sexes (Thomas Harvey in 1878), others based their arguments purely in terms of number agreement (Goold Brown in 1828), and others (the most descriptive) simply acknowledge that generic singular *they* is common in colloquial English but inappropriate for formal writing (J.N. Hooks and E.G. Mathews in 1956). The long history of prescriptivism around generic singular pronominal use that Curzan details leads quite naturally towards the efforts of descriptive linguists in establishing the extent to which singular *they* is actually used, acceptable, or even preferred. The next several experiments reviewed here investigate this matter, both from the standpoint of production and acceptance.

In Hughes and Casey (1986) and Hyde (1984), the experimenters were interested in the subjects' production of pronouns. Both experiments used fill-in-the-blank tasks that asked participants to produce pronouns that coreferred with some earlier agentive noun in the sentence, e.g. *teacher*.

Hughes and Casey (1986) reviewed the way speakers use pronouns to refer to antecedents of unspecified gender. They noted that the generic *he*—that is, *he* used to refer to an antecedent without a specified gender—had been established in other studies to be interpreted not as gender-neutral, but as masculine. They investigated whether this effect was seen in the production of pronouns with gender-neutral antecedents, and whether this was correlated with age (and linguistic development). The study tested third-grade, eighth-grade, and college students. Participants were given 20 two-clause sentences with a blank space where a pronoun would be appropriate, and were instructed to fill in whatever best fit the sentence. Hughes and Casey categorized responses as either feminine (*she*), masculine (*he*), or "other" (*it*, *they*, *he or she*). Overall among the three age groups, the most masculine of the agentive nouns used was *criminal* (84% masculine), followed by *doctor* (73% masculine). The least masculine words were *anyone* and *teacher*, which were both 21% masculine. The only word for which "other" was the clear majority was *anyone*, across all three age groups. *Teacher* was the only word for which the clear major-

ity was feminine. Overall, across all three age groups and in all words except for *teacher*, a clear preference was shown for *he*. Hughes and Casey propose some possible explanations: world experience could have informed the choices for words like *doctor* or *criminal* (though they note that this cannot be the case for *child* or *student*), or students are still adhering to the prescriptive standard that dictates the generic use of *he*.

Moulton et al. (1978), Hyde (1984), and Gastil (1990) all also look at the perception of gender provoked by pronoun use. Moulton et al. (1978) prompted groups of university students to create a fictional story about a central character; the prompt each group was given varied by pronoun, and the stories were evaluated to determine what gender the groups' characters were in relation to what pronoun was used in their prompt. Moulton et al. (1978) found that groups created female characters 35% of the time when they were prompted with *his*, 46% of the time when prompted with *their*, and 56% of the time when prompted with *his or her*.

Hyde (1984) partially replicated this methodology (in addition to the fill-in-the blank task, as in Hughes and Casey (1986)) with elementary school children. The participants were interviewed individually rather than in groups, and were asked to create a fictional story. Again, the prompt for the creation of the story varied in pronoun: when *he* was used in the story prompt, 12% of the characters were female; when the prompt used *they*, 18% of the characters were female; when the prompt used *he or she*, 42% of the characters were female. While Moulton et al. (1978) relied on the experimenters' interpretation to determine the gender of the protagonist of the story, Hyde (1984) explicitly asked the participants whether the character was a boy or a girl.

Gastil's 1990 task was radically different from either Moulton et al. (1978) or Hyde (1984). Gastil (1990) also focused on the gendered interpretation of pronouns, but relied on mental imagery. University students were asked to read sentences containing a pronoun aloud, then to report any image that came to mind; they were then asked if any gender of the sentence subjects came to mind. Gastil found that participants reported mostly 'male' images prompted by sentences using *he*, more male than female images

in sentences using *he/she* from male participants, and more generic images in sentences using *they* from male participants.

Moulton et al. (1978) and Hyde (1984) included not only so-called masculine and feminine pronouns, but *they*. In the case of abstract usage such as in these studies, the use of *they* for a singular referent is apparently allowable. These generic and indefinite uses represent what I will propose to be the most conservative form of singular *they*, with the longest history of attestation and discussion in English. When comparing indefinite uses with definite uses, many studies find that the definite use of singular *they* (even when generic) is comparably less acceptable for many speakers.

The studies I have discussed here have all focused on either indefinite singular *they* (Hughes and Casey 1986; Hyde 1984) or the gendered interpretations of generic pronoun use (Gastil 1990; Moulton et al. 1978). I will now turn to studies exploring more constrained, definite uses of singular *they*; it is not coincidental that the studies I discuss in this section are quite a bit newer. Authors such as Newman (1997) have, in their discussion of generic singular *they*, specifically noted that specific/definite uses are ungrammatical; Curzan points out and contests Newman's judgment in a footnote, stating that both she and several informants find non-gendered definite antecedents such as *my cousin* perfectly capable of anteceding singular *they* (2003: 81). This disagreement in particular is one of a few breadcrumb hints in the pre-2017 literature that the grammaticality of dsT is subject to sociolinguistic variation. However, recent studies from Sanford and Filik (2007) and Foertsch and Gernsbacher (1997) both examine singular *they* from a processing (rather than sociolinguistic) perspective, aiming to find whether English speakers find *they* more costly to resolve when used with singular antecedents. Processing cost can be a proxy for grammaticality, but neither of these studies were aimed at uncovering variation between potentially different underlying grammars.

Foertsch and Gernsbacher (1997) performed two experiments using a self-paced reading test with different sets of stimuli to probe the acceptability of singular *they*. In the first experiment, the stimuli included indefinite antecedents like (5) below:

(5) A truck driver should never drive when sleepy. even if he/she/they may be struggling to make a delivery on time, because many accidents are caused by drivers who fall asleep at the wheel.

This first experiment found that reading times were slowest (indicating difficulty processing) for stimuli where a (stereotypically) gendered antecedent was followed by the opposite pronoun (e.g. *a nurse... he*); however they found that *they* was comparable with both *he* and *she* for the indefinite antecedents—there was relatively little slow-down, suggesting that singular *they* was not difficult to process; they also found that singular *they* was preferred (had the least slow-down) when antecedents were indefinite pronouns (e.g. *anybody*). In their second experiment, Foertsch and Gernsbacher altered their stimuli to use definite antecedents, as in (6) below:

(6) That truck driver shouldn't drive when sleepy. even if he/she/they may be trying to make a delivery on time. because many accidents are caused by drivers who fall asleep at the wheel.

In this experiment, they found that reading times were (as in the first experiment) much faster with same-gender pronouns and much slower with opposite-gender pronouns; however, in this case singular *they* landed in the middle for both feminine antecedents and masculine ones – *they* was slower than congruent gender, but faster than full gender mismatch. This second experiment did also include gender-neutral (but specific, definite) antecedents such as "*a runner I knew*." When paired with these neutral antecedents, neither *he/she* nor *they* produced a slowdown. Foertsch and Gernsbacher also report a norming study in which participants were asked to provide a pronoun (much like the fill-in-the-blank tasks of Hughes and Casey 1986, Hyde 1984)—in this norming study, 70% of participants used singular *they* with an indefinite antecedent, and 20% used *they* for a singular definite antecedent.

Sanford and Filik (2007) ask whether *they* and *them* with singular referents are as easily interpretable as other singular pronouns. They provide two possible accounts for the

allowability of *they* with singular antecedents: either *they* simply permits for gender-unspecified singular antecedents in addition to plural antecedents, or *they* initiates a search for plural antecedents which first fails, then allows for singular antecedents. Using fixation times on pronouns and antecedents from an eye-tracking task, Sanford and Filik test the processing times of sentences with number mismatches ([someone/some people][they/her]) against sentences which match *they* against a singular antecedent—if *they* truly allows singular antecedents initially, there should be no cost in processing time. 36 native English speakers with normal vision were asked to read sentences including 24 test sentences, all introducing either a plural or singular antecedent, then matching it with either a plural (*they*) or singular (*him/her*) pronoun. 50% of the sentences were followed with comprehension questions. They found that there was an effect on processing time when a singular pronoun referred to a plural antecedent—this was a baseline they were testing against, so it was expected—but in a measure of first-pass reading time, first-pass regressions out, and regression path, there were no effects of a plural pronoun referring to a singular antecedent. However, looking at total reading times, they did find evidence that reading times were lower for plural pronouns referring to plural antecedents and singular pronouns referring to singular antecedents. They conclude that while *they* is an allowable gender-unspecified singular pronoun, the cost of processing time indicates that the use of the pronoun first initiates a search for plural antecedents, before accepting singular ones—hence the slow-down in reading.

These two processing studies (Foertsch and Gernsbacher 1997; Sanford and Filik 2007) both apparently demonstrated that singular *they* is variably acceptable, but a dispreferred option when a gendered pronoun would be available. The stimuli used here are analogous to the example in (3)—definite (specific) antecedent NPs, but not proper names. The account of language change that I present in this chapter will suggest that, if they were replicated in 2018, these two experiments might show significantly different results. I will now turn to the newest studies publicly available on singular *they*, both of which include specific (named) referents like the example in (4); Bjorkman (2017) gives a mor-

phosyntactic account, and Ackerman (2017) gives data from an acceptability study.

Very recent work begins to look at the use of singular *they* when referring to definite (and sometimes specific) antecedents. Bjorkman (2017) is a syntactic analysis of a possible diachronic change in the morphosyntactic features of pronouns that have shifted to allow for definite antecedents of singular *they*. The old system of morphosyntactic gender for Bjorkman was privative, binary features that differentiated *he* and *she*, so that any definite antecedent would be referred to with either of these choices. Bjorkman proposes a change in the nature of features: rather than a forced choice, she suggests that gender features in English pronouns have shifted to optional adjunct features. That is, a pronoun may either be marked as masculine or feminine, but it also may be marked for neither gender. This essentially forces a reorganization of the pronominal paradigm in English to allow for a gender-neutral singular pronoun, which has surfaced as *they*.

Another work that addresses the question of singular *they* specifically in the context of nonbinary definite specific antecedents is Ackerman (2017). In a study of the acceptability of singular *they* coreferenced with names that were either ambiguously gendered, or unambiguously either masculine or feminine, Ackerman found that the anaphor *themself* was relatively unacceptable, but was more accepted when matched with a name that could refer to a person of either gender.

In this chapter I build on Ackerman's work in investigating dsT used with names, as well as augmenting the data in acceptability judgments with natural production data that contains many instances of dsT in interview settings. In the next section I will propose two related hypotheses, then present data that addresses them.

Based on the anecdotal evidence from linguistic research and from linguists' own intuitive judgments, I here propose that the cline in grammaticality from (1)-(4) represents a move through an incremental language change. That is, with (1) being the oldest attested form (cf Curzan i.a.), and with (4) being the form that has elicited the most controversy, I hypothesize that more speakers are likely to accept (1), fewer will accept (2) and (3), and fewer still will accept (4); in other words, there will be an implicational hierarchy where

(1) > (2)/(3) > (4). Additionally, since this is an investigation into an apparent ongoing change, I hypothesize that age will be the primary factor that decides who accepts what forms. This means that in a sociolinguistic investigation, I expect to find differences in behavior around dsT between speakers of different ages. This chapter has two research questions:

> Q1: Is definite, specific singular *they* becoming more commonly accepted and produced over time? (*Is there a change going on*?)
>
> Q2: Is there a clear hierarchy between different forms of singular *they* that follow the cline in (1)-(4)? (*I.e., do specific uses of singular 'they' constitute the newest uses, followed by definite/generic uses, and do indefinite uses constitute the oldest or most conservative uses?*)

I hypothesize that age will correlate with both *production* (in continuous speech) of dsT and *perception* (in a sentence acceptability task) of dsT. If there is a language change underway, the pattern in speaker ages should show an apparent time change—that is, older speakers will be representative of more conservative forms and intuitions, and younger speakers.

3.1.1 Apparent time and Lifespan change

The Apparent Time Hypothesis (ATH) is the hypothesis that it is possible to observe real-time changes in language using data collected at a single time-point; apparent time refers to the inferred change over time that is reflected in differences in the speech of speakers of different ages (Labov 1994; Sankoff and Blondeau 2007). The strictest version of the ATH rests on two crucial assumptions which are necessarily interrelated: first, it assumes that individual speakers do not change their rates of use of alternating variants (of a sociolinguistic variable) over the course of their lifespans; second, it assumes that rates of sociolinguistic variants are established during the critical period of language acquisition.

The second assumption is reasoning for the first, but both are necessary in order to take the ATH in its strongest formulation. However, it is possible to maintain the ATH while still accounting for individual speaker changes if studies are designed specifically to control for this possible confound—thus, it is not necessarily the case that individual change invalidates a hypothesized change in apparent time. In this section I briefly review how change in apparent time and individual change in real time have been differentiated in previous studies; it is on this basis that I will discuss both support for a general change over time as well as potential indications of individual change in Section 3.5.

Sankoff and Blondeau (2007) examine the ATH through a longitudinal study of the place of articulation for /r/ (which varies between an apical variant, [r], and a dorsal variant, [R]) in Montreal French. They used two data collection points, 1971 and 1984, in a panel longitudinal study (i.e., the same individuals participated in both time points) to determine whether variation in apparent time was in fact related to change in real time; and additionally to control for whether individual speakers changed over the course of their lifespan. By following up with individual speakers in a panel study after a span of 13 years, Sankoff and Blondeau were able to pick apart individual differences among speakers from differences in the community as a whole. They found that, as the community in general moved towards widespread adoption of the dorsal [R] form, individual speakers varied in terms of how much their behavior changed over time. The three general types of speakers were 1) speakers whose use of [R] had already been near-categorical in 1971, and who showed little individual change by 1984; 2) speakers whose use of the conservative form, apical [r], was near-categorical in 1971, and who also showed little individual change by 1984; and 3) speakers who had intermediate rates of [R] in 1971, who overwhelmingly increased their use significantly by 1984. This grouping suggests that, for speakers whose use of a variant is near-categorical one way or the other, the internal linguistic system is relatively stable and does not show significant change over the course of the lifespan; however, speakers whose use of a form is variable within the context of a larger ongoing change can show change over the course of the lifespan, and exclusively in

the direction of the larger community change.[2] (I.e., no speakers reduced their production of [R] in Sankoff and Blondeau's panel study.) What the cross-examination by Sankoff and Blondeau (2007) suggests for the ATH is that the initial assumptions are correct for speakers whose linguistic systems are stable; for speakers who already show variation (even at very low rates—above 17% of the innovative [R]), that pre-existing variation enables individual speaker change over the lifespan.

This reading of the ATH predicts that speakers with pre-existing variation will only show movement towards adoption of dsT (not rejection) over time. Thus an overall effect of age in production and acceptability of dsT will constitute support for my hypothesis, even with the understanding that individuals may not (yet) be at their personal maximum of dsT use for their own lives. Some speakers also report awareness of change in their speech patterns in their metapragmatic comments about singular they. Several speakers in this section report that they notice themselves using dsT more in recent times, or that they are intentionally trying to increase their use of this form; however, I did not receive comments from any participants to the effect of trying to intentionally rid their vocabulary of dsT if they already use it productively. In the Discussion section I will show how the apparent time differences in production and perception of dsT are consistent with the analysis that there is an overall change occurring.

3.2 Experiment one: production

In this section I present the first experiment, which is a production study using sociolinguistic methods. This was part of a larger study on pronoun use more generally, but in this chapter I focus exclusively on the production of singular *they* (and dsT specifically) by trans and cisgender English speakers in Seattle. The data I present in this section are all taken from audio recordings of in-person interviews, so I take this data to be roughly

[2]This is not necessarily true in both directions: speakers with variation at any given point can maintain a rate of stable variation (as can a community as a whole) without change over time. Change over time needs variation, but variation does not always cause change over time.

representative of the participants' conversational speech. I first discuss the elicitation method, a novel type of sociolinguistic interview that I formulated to elicit third person pronouns in English without drawing conscious attention to them. I then present the resulting use of dsT in relation to speaker variables. This experiment was not sufficiently powered to conclusively show an age difference (although it is very suggestive) so the following section will cover a much larger experiment on dsT.

3.2.1 Methods

Experiment one was a sociolinguistic interview study where interviews were conducted with transgender people and interview partners both in pairs and in separate, solo interviews. Transgender participants were recruited first, and half of the transgender participants were asked to bring along a friend or acquaintance to the interview. The other half of the transgender participants were paired up with strangers, who were recruited separately from the wider population. Experiment One was approved by the University of Washington Human Subjects Division, Study Number 00000277, under the title *Transgender Linguistics / Linguistics and Gender*.

Sampling

Because of the narrow aim of the study, sampling could not be random or representative. Instead, a judgement sample was collected of 22 participants. Participants were recruited in pairs: one transgender *referent*, and one *speaker of interest*. In order to study the effects of contact with transgender people, half of these pairs consisted of participants who were already social acquaintances, and half of these pairs consisted of participants who were strangers.

The recruitment criteria were that participants were over the age of 18, native speakers of English, self-reported transgender identity on the part of the referent, and, for half of the sample, some established social relationship with the speaker of interest. Participants

were recruited through university campus and local LGBTQ+ community groups and electronic mailing lists, including social media like Facebook and Twitter.

Age group	**Income**	*Masculine-aligned*	*Feminine-aligned*	*Other*	**Group totals**
Young (18-29)					13
	high (over 90K)	1	3	0	4
	mid (30K to 90K)	2	1	3	6
	low (under 30K)	2	1	0	3
Middle age (30-55)					7
	high (over 90K)	0	3	0	3
	mid (30K to 90K)	0	2	1	3
	low (under 30K)	0	0	1	1
Older (56+)					2
	high (over 90K)	0	0	0	0
	mid (30K to 90K)	0	0	0	0
	low (under 30K)	1	1	0	2
Group totals		6	11	5	n=22

Table 3.1: Experiment 1 demographic overview

Table 3.1 shows the breakdown of participants by age group, annual income, and gender group. The sample included many more young speakers, more feminine-aligned speakers, and more middle-income speakers than other groups.

Interview method and tasks

Informed consent: All participants were given written consent forms that explained that they are being asked to participate in linguistic research, that their anonymity and audio recordings will be protected, and that at any time during the interviews or the questionnaires they may choose to ask questions or cease participation. Participants were not at this juncture informed of the nature of the research interest—pronouns and transgender referents—because doing so may have influenced their use of pronouns or gendered language during the experiment. The initial consent form described the study as focused on language and social interaction between friends and strangers.

Pre-test questionnaire: Demographic information about the participants was gathered at this time, including information on age, gender identity, preferred pronouns, ethnicity, regional affiliation, socioeconomic status, and sexual orientation. All sections of the pre-test questionnaire were free-response questions. The pre-test questionnaire was administered as a written survey.

Dyadic Interviews: After confirming consent verbally, participants were interviewed in pairs: one speaker of interest, and one "referent" participant. For participants who were acquainted with each other, interviews focused on the relationship between the two participants, asking the speaker of interest to recount narratives from their acquaintance and about the referent participant. For participants who were strangers, the interview focused on asking the speaker of interest to learn as much as possible about the referent. The aim of this interview was to elicit as much speech as possible from the speaker of interest about the referent, in order to maximize the potential for pronoun use with the referent as the antecedent. Part 1 was audio-recorded.

Solo Interviews: After confirming consent verbally, the speaker of interest was interviewed without the referent present. For participants who knew their referent partner, interviews again focused on the relationship between the speaker of interest and the referent. For participants who did not know their referent partner, interviews focused on

recounting stories and information about the referent learned in the Dyadic Interview in Part 1. The aim of this interview was, once again, to elicit as much speech as possible from the speaker of interest about the referent as possible, in order to maximize the potential for pronoun use with the referent as the antecedent. In this case, however, the referent was absent from the interview. Part 2 was audio-recorded.

Attitudes film reaction: Participants were shown short film clips (2-3 minutes) portraying transgender characters and narratives, then answered questions about the clips (based on Kempsell Jacinto 2015). The questions included Likert-scale items with answers ranging from 1-6 to measure positive or negative attitudes towards transgender identities; there were open-ended questions about the clips in order to elicit further discussion. These questions were aimed at eliciting unconscious reactions to transgender identities that may differ from participants' self-reporting of their own prejudices. The film reaction portion was administered and recorded orally; Part 3.1 was audio-recorded.

Attitudes and contact questionnaire: A questionnaire was administered to the speaker of interest to ascertain attitudes towards transgender identities and sexual minorities (adapted from Herek 2009; Norton and Herek 2013), as well as items to ascertain motivation to control prejudiced reactions (based on Dunton and Fazio 1997). The questionnaire included Likert-scale items with answers ranging from 1-6, and feelings thermometer items where participants were asked to rate the favorability of their attitudes from coldest or least favorable (0) to warmest or most favorable (100). These questions allowed participants to self-report attitudes and motivation to control prejudice. The attitudes questionnaire was administered as a written survey.

Debrief: Participants were given a written information sheet explaining the topic of research, and had an opportunity to ask the researchers any questions. Participants could at this time (or at any time during the procedures) rescind consent to participate.

The film clip stimuli:

In the film response task, participants had an opportunity to answer some freeform questions (included in Appendix 1) about clips from the 2014 film *Boy Meets Girl* before answering Likert-style questions about the clip and characters. There were four clips, all shown on an iPad during the solo interview task. Most participants opted to use headphones—usually their own.

Boy Meets Girl, directed by Eric Schaeffer, is a romantic film that depicts a young transgender woman, "Ricky," navigating life and love in a small town in Kentucky as she pursues her dreams of becoming a big-time fashion designer. Clips were chosen to depict key moments in the film that highlight Ricky's transgender status and show her interacting with other characters. Each clip has two characters, and was less than 5 minutes long.

Film clips

In the first clip shown, the character "Ricky" is working as a barista when she meets "Francesca," a customer. Francesca suggests that men are afraid of (but secretly want) commitment, and Ricky instead jokes that what men fear/desire is "dick," a joke Francesca responds awkwardly to. In the second clip, Ricky comes out to Francesca as a transgender woman, and the two affirm that they are now friends. In the third clip, Francesca has a video call with her overseas military fiancé, David. During the conversation, Francesca says she's become friends with Ricky, and David reacts angrily, calling Ricky "*that thing*" and saying Francesca should not be friends with her.[3] The fourth and final clip shows Ricky measuring her bust and complaining to her cis male friend, Robby, that she's not pleased with the physical effects of her transition.

[3] David does not refer to Ricky using "her" in this clip; Appendix 1 contains a full transcript of the dialog from all flim clip stimuli used.

Data coding and analysis

The 33 interviews conducted (dyadic and solo) were transcribed and annotated for pronoun usage. The pronouns were coded for:

(7) Speaker
   a. Speaker codes: last letter indicates paired dyad; e.g. AA and AB are partners

(8) Phi features
   a. Number: SG singular, PL plural
   b. Person: 1, 2, or 3
   c. Gender: M, F, or N (neither)

(9) Genericity
   a. Y (yes, generic) or N (no)

(10) Referent (who the pronoun refers to—either interview partners, self, or others)
   a. Coded by either speaker code (e.g. AA), index letter (e.g. J), or generic noun (e.g. ACTORS)

3.2.2 Data

In this section I present results of Experiment one. I will measure production of dsT as a dependent variable in two ways: one way will be raw token count, which is simply how many times each speaker used *they* to refer to a singular, non-generic referent; the second way will be a proportion of all third person singular pronouns that a speaker used. These two ways allow us to pick apart whether participants who used dsT relatively little also avoided other singular third person pronouns, and also account for participants who chose to talk less during interviews altogether. I will then highlight certain specific uses of dsT that were notable within the data.

Speakers produced dsT at variable rates. Figure 3.1 shows the token counts for each speaker of dsT—note that some speakers used it more than 40 times, while several never used it at all.

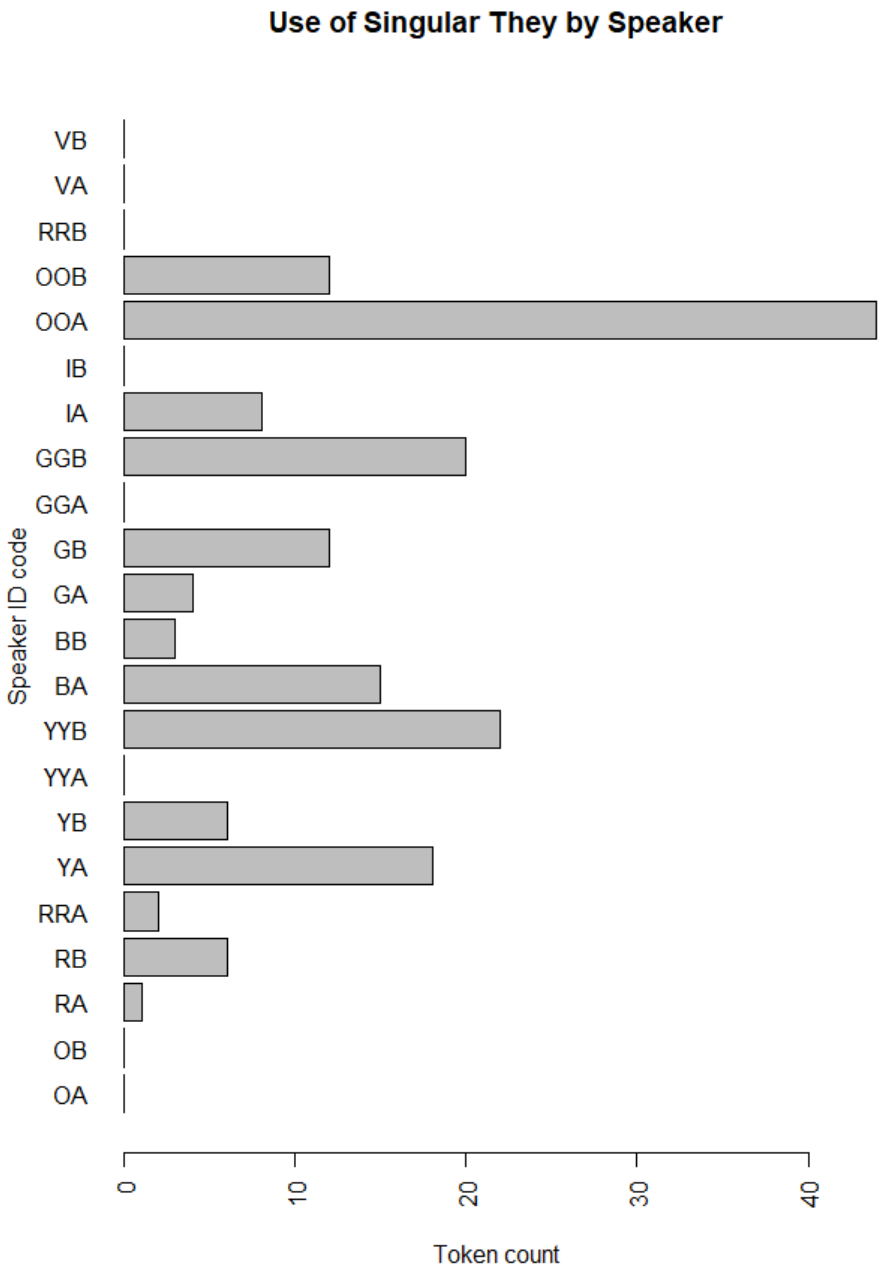


Figure 3.1: Token counts of dsT for each speaker

The following section explores which social demographic variables influence the rate of dsT for different speakers.

About the statistics

Because of the unbalanced sample population, I opted not to include reports of mixed model regressions, for fear that these models are overfitted and will lead to overinterpretation of these data. Instead, I will compare main effects of each social demographic variable individually using Kruskal-Wallis rank sum tests (due to non-parametric data).

Gender

Participants were free to write in any gender identity they wished on the demographics questionnaire. There were 11 different responses to this question ("androgynous," "female," "genderqueer," "male," "non-binary," "none," "queer," "transman," "transmasculine," "transmasculine nonbinary," and "transsexual woman"). The most common response was "female" (n=10). I binned the genders into three rough groups: masculine-aligned, feminine-aligned, and other (when no alignment was clear). Responses that included words like "female" or "woman" were coded as feminine-aligned, responses that included words like "male" or "masculine" were coded as masculine-aligned[4], and all other responses were coded as "other." The gender breakdown is shown in Table 3.2.

Feminine-aligned	Masculine-aligned	Other
11	6	5

Table 3.2: Participant count by gender

[4]I re-coded one participant who showed some confusion while filling out the demographics form, and may have confused "gender identity" (they listed "none") with "sexual orientation" (they listed "male"). Originally this participant was coded as "other," but I have recoded them as "masculine-aligned" to more closely reflect their apparent understanding of the question. Recoding this participant did not change the significance of the K-W test.

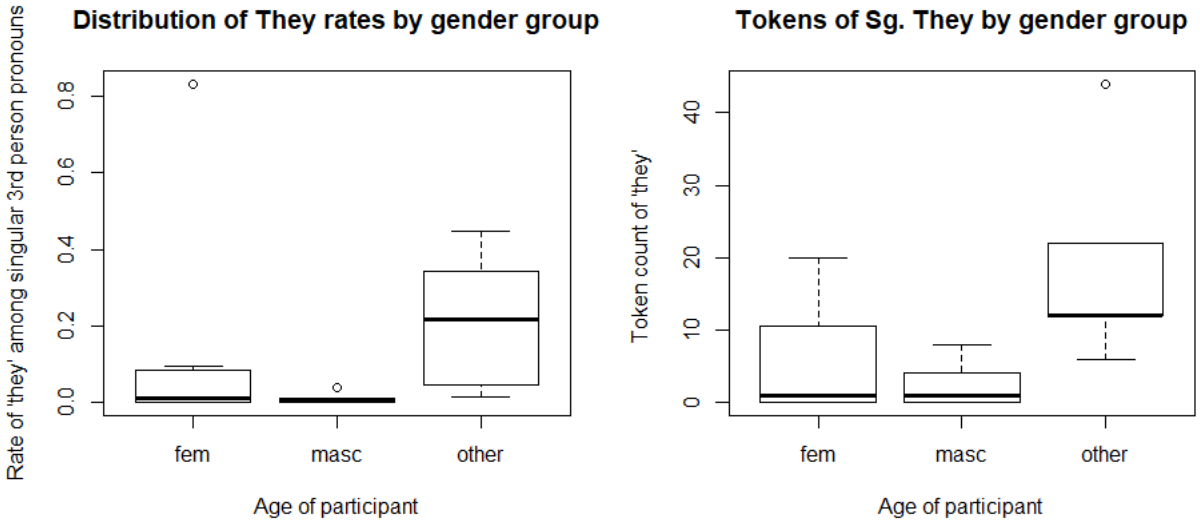


Figure 3.2: Use of dsT by gender of speaker

Figure 3.2 shows the rates of dsT production for each gender group along with the token counts of dsT for each gender group.

A Kruskal-Wallis rank sum test showed no significant relationship between gender group and rate of dsT (among other pronouns) ($\chi 2 = 1.544$, df = 2, p = 0.46); but there WAS a significant relationship between gender group and token count of dsT ($\chi 2 = 6.940$, df = 2, p = 0.03). A post hoc Dunn test showed that there was not a difference between men and women, but there were differences between nonbinary people and men (p = 0.04), and nonbinary people and women (p = 0.05).

3.2.3 Income

Because the prescriptive pressures on singular *they* may result in trends associated with socioeconomic status (if there is an ongoing change from above in either direction), I included socioeconomic class as a demographic measure of the analysis. As a measure of socioeconomic status I asked for annual income. The participants ranged from $9,000 to $300,000, with a mean of $75,636 and a median of $65,000. I binned incomes in three

categories: LOW < $30K, MID < $90K, and HIGH > $90K. A Kruskal-Wallis rank sum test showed no significant relationship between income group and rate of dsT (among other pronouns) ($\chi 2 = 0.267$, df = 2, p = 0.88); nor between income group and token count of dsT ($\chi 2 = 1.56$, df = 2, p = 0.46). The Appendix includes a graph representing this result.

3.2.4 Age

The ages of the participants ranged from 19 to 71 (mean = 32.5, st. dev. = 12.9). I binned the ages into categories: < 30 "young adult", 30-50 "middle age", > 50 "older". Table 3.3 shows the number of participants in each age group.

Young adult	**Middle age**	**Older**
13	7	2

Table 3.3: Number of participants in each age group

A Kruskal-Wallis rank sum test showed no significant relationship between age group and rate of dsT (among other pronouns) ($\chi 2 = 0.886$, df = 2, p = 0.64); nor was there a significant relationship between age group and token count of dsT ($\chi 2 = 0.384$, df = 2, p = 0.83). The Appendix includes a graph representing this result. These results may be influenced by low cell counts and an uneven age distribution. In order to look more closely at age despite low cell counts, I also analyzed age as a continuous variable.

Results of a linear regression indicated that there was not a significant correlation between (continuous, numeric) age and either rate of dsT (F(19), 0.2678, p = 0.61) or tokens of dsT (F(20), 0.1775, p = 0.68). However, below I show a smoothed Loess regression with age as a continuous variable. Figure 3.3 suggests that there may very well be a relationship between age and production of singular *they*, but the relationship is not linear.

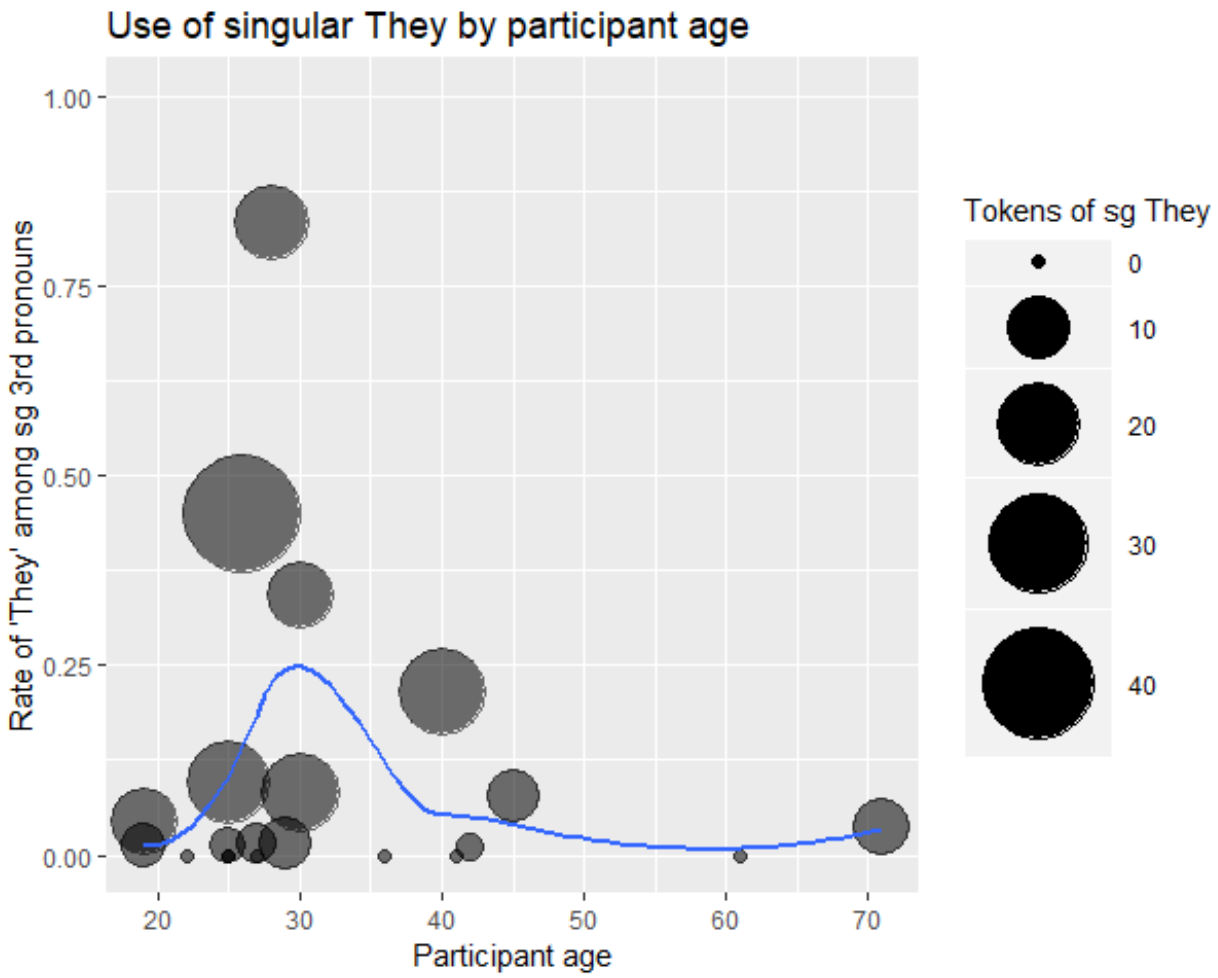


Figure 3.3: Proportion and tokens of dsT by speaker age

The greatest producers of dsT are all within a narrow age window between 20 and 35. Later in this chapter I will discuss the significance of this peak—as it stands, it is certainly suggestive of an apparent time trend that is very active among speakers with birth years between 1983 and 1998. These birth years quite closely match the birth years of the Millennial generation (1981 and 1996 according to the Pew Research Center[5]).

The results shown here for Experiment 1 are suggestive of social differences, but low

[5] http://www.pewresearch.org/fact-tank/2018/03/01/defining-generations-where-millennials-end-and-post-millennials-begin/

participant numbers make statistical confirmation difficult. Experiment Two presented below is a study designed at exploring dsT across a much greater number of participants.

3.3 Experiment two: perception

Experiment two is a sentence acceptability task where participants were asked to take an online survey about the naturalness of English sentences. Experiment two was designed using turktools and implemented using turkserver (Erlewine and Kotek 2016). Participants needed to read and agree to a consent statement before the survey, and were advised that they should be over 18 years old and native speakers of English in order to participate. Experiment Two was approved by the University of Washington Human Subjects Division, Study Number 00004635, under the title *Grammaticality of Singular They*.

Stimuli each consisted of two sentences—one including an antecedent, and one including a pronoun. Pronouns and names were randomly varied so that participants saw each combination of pronoun and antecedent, including varied masculine, feminine, and neutral names. Examples are shown in (11)-(13), and the full list of stimuli and filler sentences are included in the Appendix.

(11) **John** is very forgetful. **He(/she/they)** never remember(s) library due dates.
(Name + *he/she/they*)

(12) Students are very ambitious. **Every student** tries to write **her** essay perfectly.
(Quantifier + *he/she/they*)

(13) **The perfect spouse** is very thoughtful. **He** will always try to remember birthdays and anniversaries.
(Generic definite + *he/she/they*)

Participants were shown the stimuli and asked to rate the sentence using a Likert scale from 1 (very unnatural) to 7 (very natural). After being presented with 45 questions (including 15 target sentences and 30 filler sentences), participants were asked to answer

demographic questions (also included in the Appendix). Participants were recruited via online social networks. The survey was active for 2 weeks, and collected 884 responses.

3.3.1 Methods

Consent: Participants were recruited through social media and email lists. Recruitment included a link to the web-based survey. The web-site included a written consent statement that explained that they were being asked to participate in linguistic research, and that their responses were anonymous. Participants were advised that if they wished to cease participation, they could close their web browser at any time before or during the experiment. Participants were not at this juncture informed of the object of research – acceptability of singular *they* – because doing so may have influenced their answers to survey questions. The initial consent paragraph described the study as focused on "how different native speakers of English comprehend and rate the naturalness of a variety of English sentences."

Instructions and example: The web-site for the survey included instructions for answering Likert-scale acceptability ratings questions. Participants were advised to rate sentences for 'naturalness,' with a rating of 1 being the least natural and 7 being the most natural. Instructions indicated that "it might look like there's a mistake or typo in one of the sentences, or like something is wrong with the writing. If you think this is the case, you should rate the sentence lower."

After the instructions, participants were shown an example stimulus designed to be very natural, with advice that natural stimuli should be rated higher, followed by an example stimulus designed to be very unnatural ("My father is very secretive. My father is the only person who I don't know when was born."), with advice that unnatural stimuli should be rated lower.

Survey instrument: The survey presented all participants with 45 questions in a random order. 15 of these sentences were target sentences, including all possible pronouns combined with names of masculine, feminine, and neutral gender, as well as indefinite

and definite generic antecedents. All filler sentences were modeled to closely resemble the targets, so that all stimuli included two sentences.

The Likert scale instrument included HTML-styled radio buttons for numbers 1 through 7, with 1 being labeled "very unnatural" and 7 being labeled "very natural." After participants clicked on a radio button for a stimulus, the stimulus and radio buttons were hidden for that question. After completing all 45 questions, participants were asked to answer post-survey questions.

Post-survey: The post-stimuli questions first asked participants to comment on what basis they used for rating sentences - this question was a text entry box, allowing participants to comment as they wished. The post-survey then presented demographic questions, including questions about native language, age, gender identity and transgender identity, ethnicity, and dialect area. The full demographic survey index is included in Appendix 2.

Demographic make-up of the respondent sample

Overview: The survey received 884 responses over the course of 14 days in August 2018. I excluded responses from non-native speakers of English participants and participants who did not answer a majority of the questions, so I will here analyze the results of 754 responses. Table 3.4 shows the ethnicity, gender, and age of these 754 participants. The largest represented group were white women between the ages of 26 and 30. The majority of participants were white, and the majority of participants identified themselves as women.

MAN	under 20	20-25	26-30	31-35	36-40	41-45	46-50	51-55	56-60	61-65	66-70	Over 70
No answer	0	0	2	0	0	2	0	0	1	0	0	0
American Indian / Alaska Native	0	0	1	2	0	0	0	0	0	0	0	0
Asian	1	2	2	3	0	0	1	0	0	0	0	0
Black / African American	0	1	0	0	0	0	0	0	0	0	0	0
Native Hawaiian / Pacific Islander	0	0	0	0	0	0	0	0	0	0	0	0
White	1	17	29	27	18	28	11	4	11	7	3	1

WOMAN	under 20	20-25	26-30	31-35	36-40	41-45	46-50	51-55	56-60	61-65	66-70	Over 70
No answer	0	1	3	1	3	0	4	0	0	0	1	0
American Indian / Alaska Native	1	1	0	0	0	0	0	0	0	0	0	0
Asian	2	2	3	2	0	0	0	0	0	0	0	0
Black / African American	0	3	1	1	1	1	0	0	0	0	0	0
Native Hawaiian / Pacific Islander	0	0	1	0	0	0	1	0	0	0	0	0
White	6	57	89	50	45	37	34	22	24	11	7	5

NEITHER	under 20	20-25	26-30	31-35	36-40	41-45	46-50	51-55	56-60	61-65	66-70	Over 70
No answer	0	0	2	0	1	0	0	0	0	0	0	0
American Indian / Alaska Native	0	0	0	0	0	0	0	0	0	0	0	0
Asian	0	3	1	0	0	0	0	0	0	0	0	0
Black / African American	0	1	0	0	0	0	0	0	0	0	0	0
Native Hawaiian / Pacific Islander	0	0	0	0	0	0	0	0	0	0	0	0
White	3	29	20	12	2	1	2	0	1	1	0	0

Table 3.4: Participant demographic breakdown

3.3.2 Data

In this section I will first address the use of singular *they* across all antecedents (specific, generic, and quantified), then I will discuss the effects of different antecedents. The main social variable of interest will be age, since I hypothesized a change in progress and in the last section suggested an apparent time difference in the production of singular *they*. I will also discuss gender and transgender identity as social variables. Age, gender, and trans identity were all significantly correlated with ratings of singular *they*.

'They' compared with other pronouns

Comparing all uses of singular *they* with *he/she*, regardless of either antecedent or any social facts about the participant, shows that the acceptability of *they* is much more highly variable than for *he* or *she*, despite the inclusion of apparent gender mismatches with all pronouns. Figure 3.4 shows the variable acceptability of *they* compared with *he* and *she*—notably, the ratings of *they* have a wider spread than those of *he/she*.

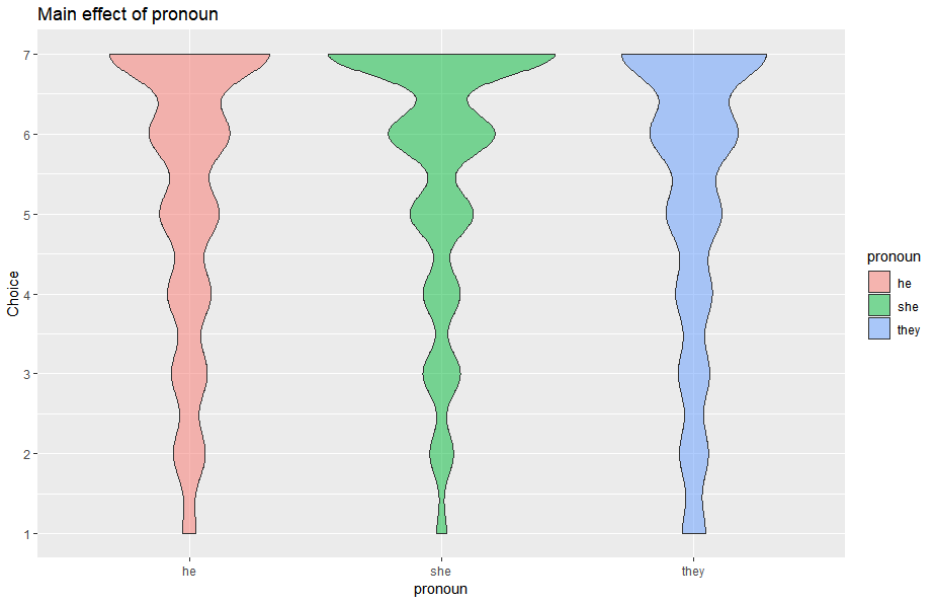


Figure 3.4: Rating of each pronoun by all participants

A one-way ANOVA test showed that there was a significant difference between the acceptability ratings of each pronoun above ($F(2, 10137)$, $p < 0.001$). Post hoc comparisons using the Tukey HSD test indicated that the mean rating of *she* was significantly different than *he*, and that the mean rating of *they* was significantly different than *she*. However, the ratings of *they* and *he* were not significantly different. In my discussion section, I will discuss the implication for the general acceptability of singular *they* across all speakers.

Looking at the effect of antecedent type on each pronoun, the ratings do appear to vary. Gender mismatches (e.g *Mary . . . he*) are rated relatively low for either stereotypically gendered name + *he/she*; this effect does not appear as prominent for gender-neutral names. Figure 3.5 shows the interaction between antecedent type and the ratings of each pronoun.

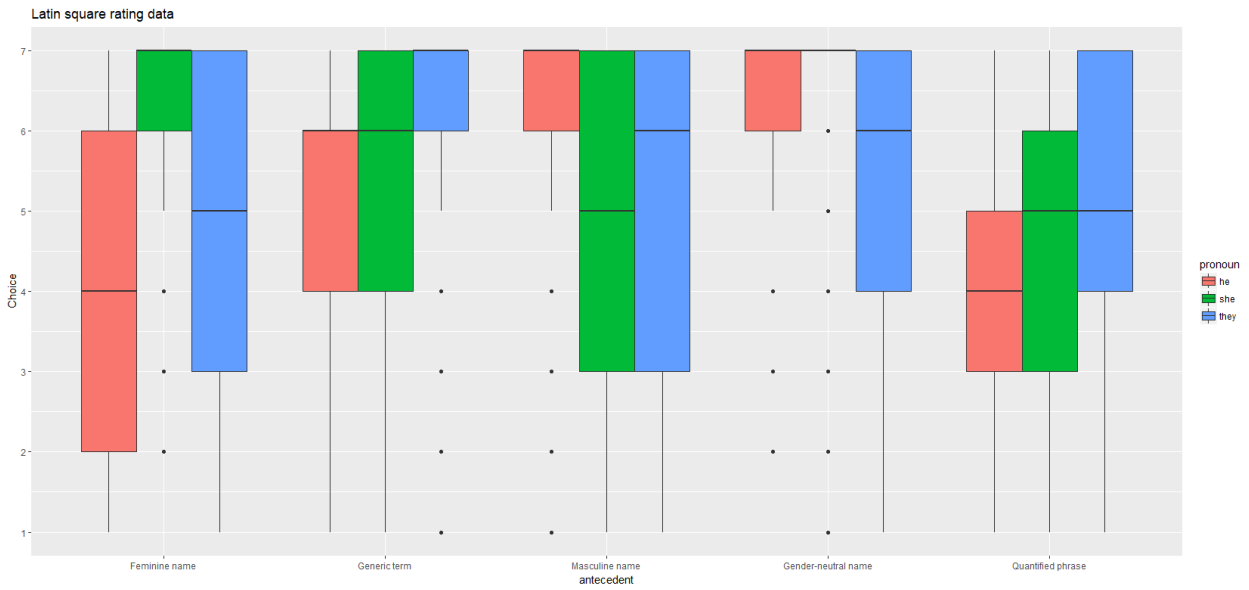


Figure 3.5: Rating of different pronouns by antecedent type

Singular *they* is highest rated when used with a generic definite term (e.g. *The ideal barista is very attentive. They will always make drinks carefully and quickly.*) It is lowest rated when used with a feminine name, but that is comparable to when it is used with a

quantified NP.

A one-way ANOVA test indicated that there was a significant difference in rating of singular They depending on the type of antecedent ($F(2,3379) = 78.72$, $p < 0.001$). Post hoc Tukey HSD testing showed differences between all types of antecedents except three pairs (neutral name VS masculine name, quantified NP VS masculine name, quantified NP VS neutral name). The full results of the Tukey test are included in Appendix 2.

In the discussion section I will return to the question of antecedent type as conditioning the (general) acceptability of singular *they*; for now it is worth noting that with a generic definite NP *they* is almost at the ceiling—meaning it was very widely considered perfectly natural and acceptable to most participants. It is also worth noting that masculine and feminine names are apparently less acceptable than other antecedents.

Social variables

In this section I will turn to examining the effect of social demographic variables of participants on the acceptability ratings of stimuli. For this section I will focus primarily on singular *they*, though I will give some attention to other types of gender mismatches at the end of this chapter.

Participant age

Because I have hypothesized a change in progress, the most important variable to examine is age of participant. When all antecedents were grouped together, age did have an effect on the acceptability rating of singular *they*. Figure 3.6 includes *she* and *he* for comparison:

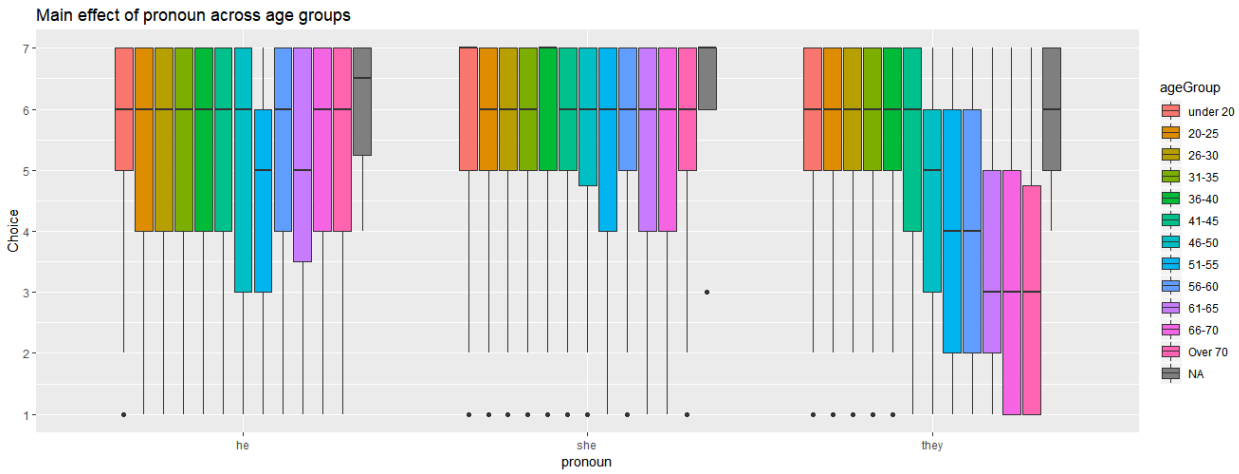


Figure 3.6: Rating of each pronoun by speaker age group

Looking only at singular *they*, the age difference becomes much more apparent. In Section 3.1, I hypothesized that generic and indefinite antecedents are older, more established uses, whereas definite and/or specific uses are more innovative. If this is the case, any age effect should differ depending on antecedent type—and this does appear to be the case. Older participants find singular *they* less acceptable across all antecedents, but to different extents—Figure 3.7 below shows a more notable difference in rating by age for proper names (regardless of gender) than for generic or indefinite antecedents.

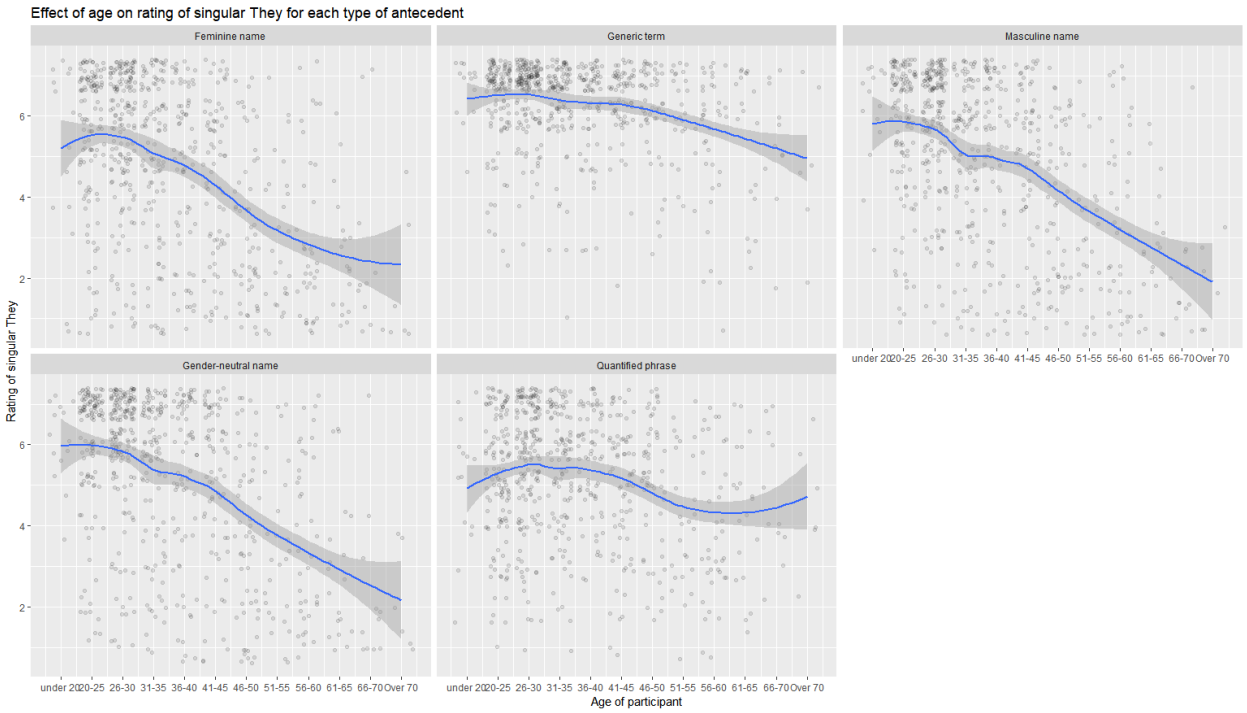


Figure 3.7: Ratings of singular *they* for each pronoun by age group

The age effect for names (regardless of their stereotypical gender) is more apparent than the age effect for generic and quantified NPs. In the discussion section I will relate this to my original hypothesis about how the change in singular *they* is progressing.

Participant gender

Participants were also asked to identify their gender (as either *man*, *woman*, or *other*); Table 3.5 summarizes the gender make-up of all participants.

Man	Woman	Neither	Declined to answer
175	420	79	5

Table 3.5: Participant counts in gender groups

A one-way ANOVA showed a significant difference in mean rating of singular *they* depending on participant gender ($F(2, 3357) = 66.43$, $p < 0.001$). Post hoc Tukey HSD test

showed that men and women did not significantly differ from each other, but participants who answered "Other" did differ from both men and women. Men and women rated singular *they* much more variably, while participants who were neither rated *they* almost at ceiling. Figure 6.5 in the Appendix includes a visual representation of this difference.

In Section 3.5 I will discuss the particular implications of the difference in gender, both in respect to the language change in progress and in respect to the specific social meaning of singular *they* as a sociolinguistic variable.

Participant transgender identity

Separately from gender identity, participants were asked whether they identified as transgender. Table 3.6 shows the makeup of transgender and non-transgender participants for each gender category. While transgender participants made up 11% of the study, they constitute 86% of the "Neither" gender category. A large majority of transgender participants answered as *neither*-gender, as well. Any insights in this section about the relationship between transgender identity and acceptability of singular *they* are therefore going to be inextricably related with the observation made in the previous section, that *neither*-gender participants accept singular *they* much more readily than others.

	Not trans	**Trans**
Man	167	8
Woman	409	9
Neither	28	49

Table 3.6: Participant count by transgender identity

In this section I look first at transgender identity alone, and then look at the interaction between transgender identity and gender identity as they affect ratings of singular *they*.

A one-way ANOVA test showed a significant effect of transgender identity of the participant on mean ratings of singular *they* ($F(2, 3348) = 136.3$, $p < 0.001$). Tukey post hoc testing showed that transgender participants rated singular *they* higher than non-trans participants; this is also readily apparent in Figure 3.8 below.

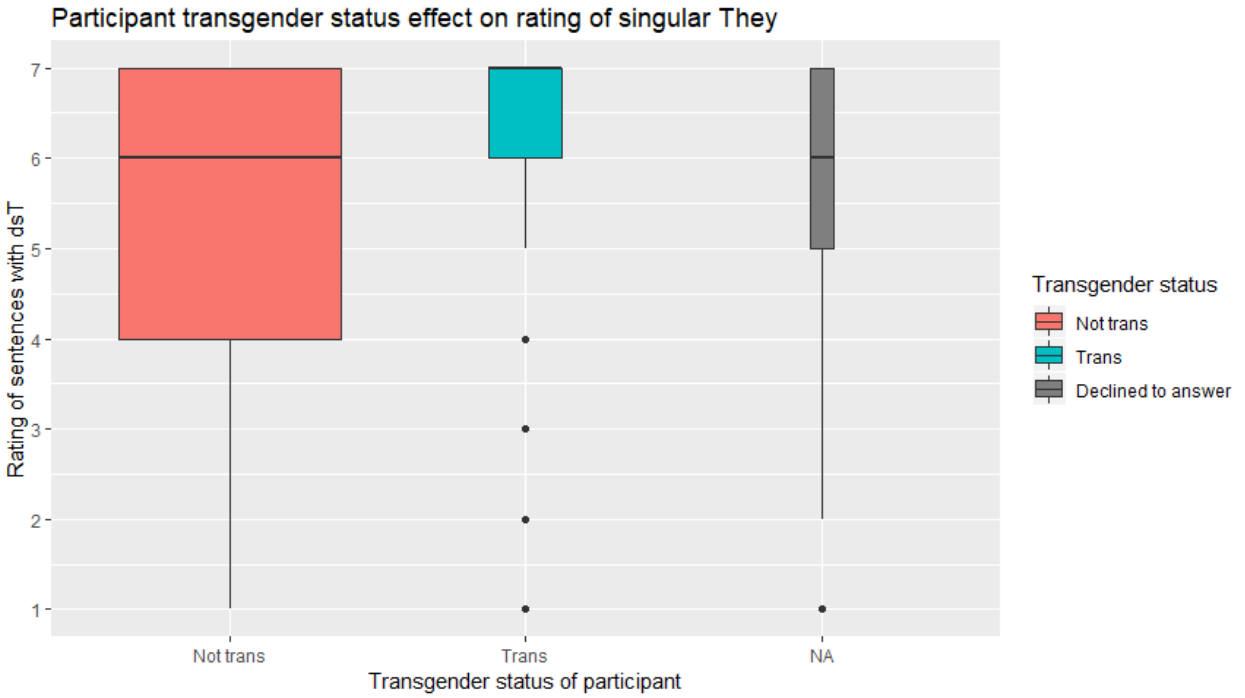


Figure 3.8: Rating of singular *they* by speaker transgender identity

In examining the combined effect of gender identity and transgender identity, it seems clear that these data have considerable overlap between *neither*-gender and transgender participants; Figure 3.9 shows however that transgender participants of all genders rate singular *they* almost at ceiling, while there is variability in the non-transgender group

Later in this chapter it will be necessary to determine whether it is the case that transgender and/or *neither*-gender participants are leading change generally, or whether the particular social meaning of singular *they* is especially prominent within transgender/nonbinary gender groups. If it is the case that trans/nonbinary participants are generally out at the forefront of a grammatical change irrespective of its social meaning, then it seems reasonable that they would show an extremely advanced version of the same grammatical steps that the larger population is taking—that is, trans/nonbinary participants should still show some effect of antecedent type in their acceptability ratings

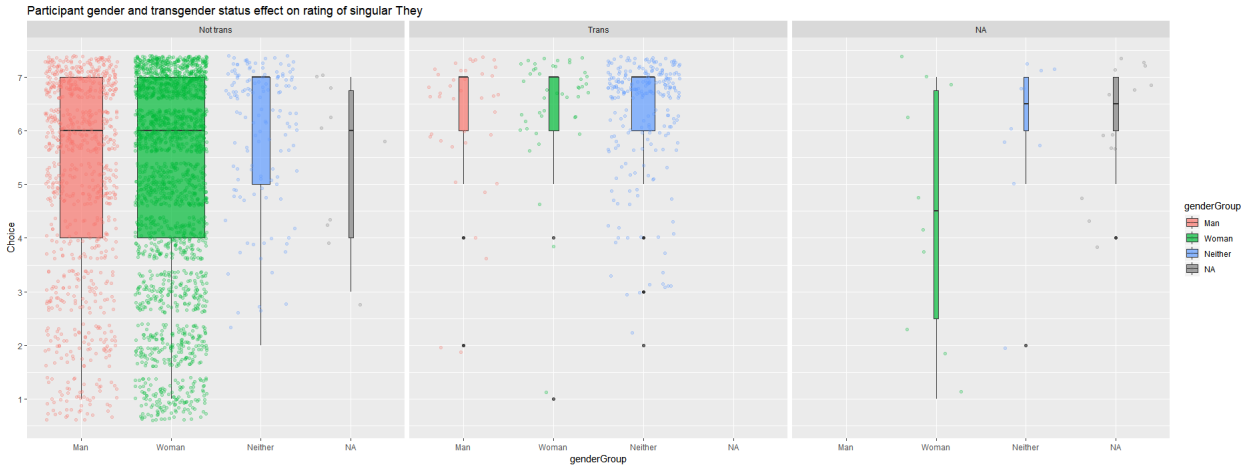


Figure 3.9: Intersection of gender group and transgender identity

of singular *they*. Figure 3.10 shows that transgender participants rate *they* very highly for all antecedent types; compared to non-transgender participants, transgender participants show even more consensus for specific names than for generic and quantificational terms – this does *not* pattern with the same order of increasing acceptability we see in the general population.

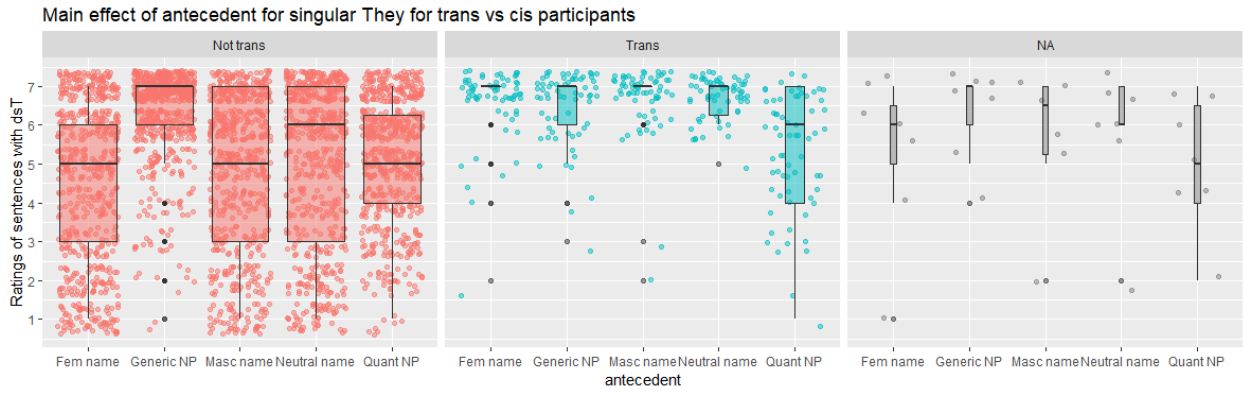


Figure 3.10: Rating of singular *they* by antecedent and speaker transgender status

A one-way ANOVA test indicated that there was a significant difference in rating

of singular *They* by transgender participants only depending on the type of antecedent ($F(4,324) = 22.38$, $p < 0.001$). Post hoc Tukey HSD testing showed differences between quantified NPs and all other antecedent types – transgender participants rated quantified NPs significantly lower than any other antecedent. This is not the pattern shown among all study participants as a group, nor among non-trans participants. In the discussion section I will discuss the ramifications of this finding for a probable social meaning of singular *they* with specific antecedents.

Comparing the effect of (other) demographic variables for transgender participants as compared to the experiment participants group as a whole may also inform us of whether transgender speakers are progressing along the same route of language change as the rest of the population. In a logistic regression model of singular *they* ratings across all participants, age and antecedent type were significantly correlated with ratings of singular *they* ($p < 0.001$); however, in a model of only transgender participants, these effects largely disappeared. This may, however, be partly because the study included very few older transgender participants. Appendix 2 includes the full regression tables.

3.4 Summary of Findings

In the experiments I have shown in Section 3.2 and Section 3.3, two patterns emerge: first, it is apparent that dsT is an established part of the grammar (both in production and in acceptability) for many speakers; second, some social factors (such as age and gender) appear to influence the perception or rating of dsT in a way that suggests this is a socially-sensitive variable. While Experiment 1 primarily shows that dsT is robust and attested in natural speech, the data were not sufficient to determine the relationship between production and social variables. Experiment 2 gives less detail about the use of dsT by individual speakers, but shows a more reliable statistical relationship between higher ratings of dsT and social variables including age, gender, and transgender identity.

In the next section I explore how these results can be interpreted in order to answer my initial research questions (about dsT undergoing a grammatical change); age is the so-

cial variable most under question when investigating a possible change in apparent time data, but I also discuss the relevance of gender identity in this context. Section 3.5 also investigates the possible social meaning of dsT that is supported by these data, especially as it is reflected in metapragmatic comments by experiment participants.

3.5 Discussion and Analysis

In this section I explore the change in progress that these data support. I will first analyze speaker demographic variables of interest, with particular attention to age, gender, and transgender identity. I will then explore the social meaning of the language change, pulling both from quantitative results and from analysis of metalinguistic comments about singular *they* in its various uses. Two general patterns emerge: the social meaning of generic uses of singular *they* relating to gender equality, and the meaning of specific uses of singular *they* relating to non-binarist gender etiology. Orders of indexicality map these disparate social meanings to the differences in production and acceptability of singular *they* with different antecedents. Finally, I will conclude this chapter with a section investigating particular productions of singular *they* in Experiment One that warrant more in-depth sociopragmatic exploration; this will lead into Chapter Four.

3.5.1 Social variables of interest for determining change in progress

This section analyzes speaker variables that significantly differentiate speakers' production and acceptance of singular *they*. The section on age will revisit the apparent time hypothesis; I will then discuss how age was grouped in the two experiments and how to analyze age in a socially meaningful way. I will then propose a "drop-off" point where there is an (apparently sudden) change in acceptability of singular *they* that aligns with generational divides. The section on gender will question whether it is appropriate to analyze certain genders as "leading" the change in progress, looking first at differences between men and women, then at differences between binary and other genders. Finally, I will discuss the meaningfulness of transgender identity as a social variable that signif-

icantly predicts acceptability of singular *they*, again asking the question of whether it is appropriate to try and identify groups as "leading" grammatical change more generally.

Age: Apparent time hypothesis

In the beginning of this chapter I introduced the apparent time hypothesis as an analytic tool for extrapolating real-time change from synchronic variation. The two experiments that I have presented here have showed varying results: in Experiment One, age impressionistically appeared to be a factor in production of dsT but did not have the statistical power necessary to reject the null hypothesis. In Experiment Two, age was a statistically significant predictor of rating of singular *they* for all antecedents together. The results from Experiment Two also showed variable age effects for each antecedent type. While the apparent time hypothesis has been explicitly tested in variables of phonological variation, fewer grammatical variables have been tested—Bailey et al (1991) found real-time evidence confirming apparent-time change in double modals and quasi-modals, but it is not clear whether these variables are comparable to singular *they*; this is partially because the variation in singular *they* is not completely syntactic, but also deeply tied into pragmatic and social considerations of pronoun use; Chapter 4 explores those complications thoroughly.

In determining whether a difference in age can be indicative of a change over time, we must also consider the possibility of lifespan change; Cukor-Avila (2000) found evidence of intraspeaker variation over time in grammatical variables, so we cannot rule out the logical possibility. However, the changes she found were in production, rather than perception—and so may be more comparable to Experiment One than Two. Furthermore, based on the findings from Sankoff and Blondeau (2007), it is also *possible* that individual speakers' grammaticality judgments of singular *they* may change over time—in Section 3.5.2 below, I include metalinguistic commentary from Experiment Two participants that indicates that speakers certainly *try* to change their use or acceptability of singular *they*. Speakers also, however, report that such change is *difficult*, which may be as much

due to social biases as due to grammar (as in (25) below). The fact that speakers make reference to *trying to learn* to accept singular *they* intentionally does not explain the actual direction of the age variation seen in Experiment Two—if speakers were learning to use singular *they* over the course of their lives, then we would expect that older speakers were in fact more adept at it, which is the opposite pattern to what Experiment Two shows.

It is not unreasonable, therefore, to conclude that most (but maybe not all) of the variation in age in acceptability of singular *they* is a cross-sectioned slice of a change happening in real time. The results of Experiment Two are particularly suggestive of an apparent time change—especially the difference in age variability with different antecedents. At the outset of this chapter I proposed that the change is progressing stepwise, with generic and indefinite antecedents being paired with singular *they* first, and more definite and specific antecedents following later. The pattern found in Figure 3.7 above supports this hypothesis.

Emic Age

This section investigates the social context around the variation in speaker age that correlates with ratings of dsT; in order to propose a socioculturally-supported explanation for language change over time, in this section I examine how speaker age can be separated into socially-meaningful categories in their native context (*emic* age, a term analogous to the phonetic/phonemic distinction; Pike 1967). My analysis of socioculturally meaningful (emic) age in these data will be based upon generational analysis.

If we accept the proposal that the age variation does indicate an apparent time change, which is reflective of change in real time, then the question becomes: when did the change start, and when will it reach completion? The answer to this is dependent on an emic grouping of speaker age—identifying a range of birth years that signals greater variability is much less meaningful without a social understanding of *why* a grammatical change might happen at that point in time. Additionally, Experiments One and Two elicited age data in different ways, and grouped them differently for statistical analysis; in order to

reliably compare the results of the two experiments, it is necessary to determine what age ranges are socially meaningful for the communities in which the experiments took place.

Generational analysis is the social-scientific episteme of separating groups in a community into age cohorts by birth year, where the boundaries of each cohort are defined by prominent historical events and context. For the investigation of linguistic change over time I will be working loosely within the social-structural model of generational analysis (Biggs (2007) reviews this) where age cohorts are organized with the goal of understanding social change. The birth years represented in Experiment One are 1947 – 1999; Experiment Two collected age data in five-year bins: the youngest participants marked themselves as "under 20," (but over 18 to participate), and the oldest participants as "over 70"—this age range means the birth year range of Experiment Two is ~ 1947 – 2001. I will follow The Pew Research Center's analysis of generational divides in the United States[6], shown in Table 3.7 below—that means that the participants from both experiments combined comprise four generations.

Born 1946-64	Baby Boomers
Born 1965-80	Generation X
Born 1981-1996	Millennials
Born 1997-2012	Generation Z

Table 3.7: Birth years for generations

In Figure 3.7 in the previous section, there appears to be a sharp divide between relative stability of ratings up to age 35, dramatic variability between ages 36 and 55, and relatively stable low ratings over age 55.

Experiment Two was conducted in 2019, meaning the birth year cut-off for broad acceptance of singular *they* falls roughly at 1983—around when the cut-off between Generation X and Millennials is commonly placed. The second transition point is at the birth

[6] http://www.pewresearch.org/fact-tank/2019/01/17/where-millennials-end-and-generation-z-begins/

year 1965. This is the cut-off between Baby Boomers and Generation X. The age binning in the demographic questionnaire for Experiment Two was not explicitly designed for generational analysis, but future investigations (especially around singular *they*) would do well to focus on the apparent transitional periods in apparent time—birth years around 1980 and 1960 appear to be noteworthy.

Accepting the loose relationship between the cut-off years above and the S-shaped curve falls within the assumptions of social-cultural generation analysis as related to social change; the fact that singular *they* is a sociolinguistic variable that is highly salient and tied to social referential meaning (through gender etiology) makes sense in this perspective. The grammatical change necessary for dsT to emerge was likely not spontaneous or random, but occurred in parallel with social changes in the organization of how English speakers view gender in general—the social change and the linguistic change are likely mutually reliant upon one another. I will return to the social meaning of the change below, using metalinguistic commentary as a way of probing possible indexical meaning of dsT. In the next subsection, I will turn to interspeaker variation by gender and transgender status, exploring whether these social variables are necessarily tied to the change in singular *they* because of the general nature of language change, or whether these relationships are linked directly with the variable in question.

Gender

In many reviews of language change, sociolinguists will look to gender in a binary way (are men or women leading the change?), sometimes folding in multiple dimensions of analysis to explain apparent paradoxes (are middle-class women in New York expected to behave like working class women in Martha's Vineyard?). No large mainstream studies of language change have included a third ("other" or "neither") gender category in the way that I have done in Experiments One and Two for this chapter—even in instances where linguists have managed to recruit any nonbinary people, it has been in such sparse numbers that these participants were excluded from the analysis. This section therefore

is asking two questions: first, is one particular gender group out of the three behaving noticeably different than the others; and second, is it reasonable to expect that differences are a matter of being further ahead or behind in the midst of a linear process? I will show that in fact the inclusion of a third, non-binary gender group illuminates exactly the problem with viewing language change (even with a clear apparent time pattern, as I showed in previous sections) as a necessarily linear process throughout which all speakers perform the same steps in the same order. The differences in gender found in these experiments are tied innately with both indexical and referential meanings of the variable, and this becomes even more obvious when working with more participants who don't fit neatly into traditional gender categories.

It is not necessarily clear that we should expect the change in singular *they*, which is syntactic and (socio)pragmatic, to pattern along with changes in phonological or phonetic variables, or even with other grammatical variables; dsT is not an unconscious reflex, and the rate of its production depends not only on sociolinguistic factors but on matters of reference. There is no way to confirm that higher uses of dsT are necessarily due to the identity of the speaker, since the token rates will always be confounded with the need to refer to people outside the gender binary. Thus, while this section will examine the differences between genders (from two angles), I do not suggest that the patterns of dsT use shown here will be generalizable to other sociolinguistic variables.

In Experiment One, there was not a significant difference between women and men in either token or proportional use of singular *they*—however, there was a difference between men/women and nonbinary participants in token production of dsT. The same results were found in Experiment Two: there was not a difference between men and women, but there was a difference between the nonbinary category (which is not assumed to be homogenous) and men and women. This is further evidence that it is not appropriate to compare the differences in gender to previous sociolinguistic studies, since vanishingly few include analysis of the three gender categories that I have included.

The differences between nonbinary and binary genders in both experiments, on the

other hand, is suggestive of the social meaning of singular *they*; as metalinguistic comments will show in Section 3.5.2 below, speakers are aware of the tie between singular *they* and nonbinary genders explicitly. It therefore follows that nonbinary groups produce dsT more (by token), and that they rate it higher in acceptability tasks. The popular perception that dsT is being used and promoted by nonbinary speakers appears to be confirmed in the results of these two experiments.

Transgender identity

The relationship between nonbinary and binary gender participants does gloss over another important distinction: not all nonbinary people are transgender, and not all transgender people are nonbinary. Experiment Two explicitly separates out self-identification of transgender status as a separate question from gender identity, allowing us to probe these variables which are often otherwise conflated.

First, I noted in the previous section that transgender men, women, and nonbinary people rated singular *they* more uniformly and higher, while non-trans men and women show much more variability. This suggests that, independently of nonbinary identity, transgender people of all genders are more accepting of singular *they*. This may be due to effects of social proximity (transgender people are more likely to have a lot of nonbinary friends), or it may be a reflection of underlying gender ideology (transgender people are less likely to ascribe to a binarist view of gender).

I earlier posed the question of whether it might be the case that transgender people are "further ahead" on the progression of grammatical change that allows dsT; however, the differences in ratings of singular *they* with different antecedents among trans participants did not pattern similarly to the ratings among non-trans participants; while non-trans participants rate proper names (of any gender) lower when paired with singular *they*, transgender participants rated them *higher*. The apparent inversion of the effect of antecedents suggests that transgender people are not necessarily progressing along the same "schedule" of grammatical change, but rather have developed their own indepen-

dent norms. This strongly suggests that transgender ratings of singular *they* are reflective of social variables (such as the possibilities posed above—social contact or gender ideology) rather than particular advancement along the same grammatical railroad as the general population. These facts, coupled with the metalinguistic comments I show in the next section, point to the social meaning and "zero-th indexical order" of speakers who are the highest users of dsT.

3.5.2 Social meaning of the change

This section turns to exploring the social meaning of the change directly. I will first present the metalinguistic comments elicited from Experiment Two as evidence of social salience; then I will turn to separating out particular social meanings for different uses (stages) of singular *they*; finally, I will use a model of orders of indexicality to propose an analysis for how the change originated and how it has begun to spread.

Metalinguistic comments and awareness of change

In Experiment Two, participants were given an opportunity to comment on the stimuli that they rated. After rating all the stimuli, but before they were presented with demographic questions, participants were asked to share what factors influenced how they decided on their ratings of each sentence. The answers were collected through a free text entry field in the online experiment; the full text of the prompt for this field is included in Appendix 2 along with demographic questions.

The responses to the comment field included many metalinguistic comments about grammar—because the filler items composed a majority of stimuli, many responses addressed ungrammaticality in those items. In this section I will focus primarily on comments about the target stimuli, which included various antecedent types paired with the singular pronouns *he*, *she*, and *they*. Figure 3.11 below shows a word cloud generated from the text corpus of all participant comments; the size of the font is proportional to

the frequency of the word in the corpus.

Figure 3.11: Wordcloud of words used in Experiment 2 comments

Gender, pronouns, and names were among the topics that many participants commented on. 37% of comments volunteered about the acceptability included the word "gender", 20% included the word "singular", 6% included the words "singular they" (only 3.6% included the word "trans"). This suggests that despite the filler items providing some distraction, respondents were basing their answers at least partially on considerations of the confluence between pronouns and the apparent gender of their referent. Below I include examples of comments in answer to the question asking respondents why

they picked certain ratings for different sentences.

(14) "Lack of pronouns, mis-ordering of words. I would say that sentences that use a singular they or unexpected gendered pronoun aren't as unnatural as some might think, and shouldn't be considered grammatically incorrect."

(15) "Some of them had unusual pronouns which **sound slightly unnatural but are gradually becoming more acceptable**."

(16) "I still find unexpected uses of he/she/they weird but gave them a middle 4 because **I know the rules are changing and why**" [emph added]

(17) "Whilst I know **the use of *they* as, a pronoun is growing**, I find it doesn't sound right."

Respondents who commented about dsT in particular made various comments about its grammaticality (both for and against), suggesting that as a variable undergoing change dsT is very salient and its users are aware of their use of it. Examples (16) and (17) show explicit comments on the respondents' awareness of ongoing linguistic change. These kinds of comments are expected, given the amount of metalinguistic commentary that the variable receives generally (discussed further in Chapter 4).

Other comments particularly noted the use of generic *he* as less acceptable, and *they* as the preferred generic pronoun

(18) "Sometimes the pronouns felt forced. For example, when a genderless subject was introduced in the first sentence it felt unnatural to assign a gendered pronoun (he or she) to the subject in the second sentence. Most people I know use 'they' if they don't know the gender of the person they are referring to or if they are talking hypothetically about a generic person."

(19) "I mostly rated sentences lower if there was no specific gender implied but a ""he"" or ""she"" was used as a generic pronoun. **"Their" as a generic pronoun is preferred.**"

That respondents are aware of an ongoing change, and that they notice generic versus specific uses of singular *they*, again suggests that the change is salient and conscious.

In the cases of comments on generic *he*, many commented that they found the construction unnecessarily gendered. This speaks to the older origins of singular *they* before the widespread use of dsT—as a generic pronoun, it is the only truly gender-neutral pronoun to use with non-gendered antecedents. Popular use of generic singular *they* with its surrounding discourse of gender neutrality makes it ripe for reanalysis as a gender neutral *specific* pronoun –what had to change, then, was the conception of individuals being able to be gender-neutral.

In cases where respondents commented specifically on dsT, many made reference to their own queer/transgender identity, or the presence of LGBT+ people in their close social network. They largely mentioned this in support of dsT:

(20) "I heartily support ""they"" pronouns for individuals, and not assuming gender based on names. **I'm queer.** I'm good at spelling and grammar."

(21) "some questions used names and pronouns that are not commonly used together, and referred to people with a singular 'they' which **may only have been noticable to me because I am transgender**"

(22) "**I've spent enough time in queer/trans/non-binary social contexts** at this point that that stuff is natural for me now and remarkable only to the extent that I'd expect some other folks to take exception."

(23) "My answers to they/them for a specific person have shifted much more positive in the last few years, **thanks to nonbinary friends**."

(24) "I noticed a bigger brainstop with heavily female names and gender neutral 'they' than with male names, no surprise there (if it helps, **I have a lot of genderqueer friends** who I've had to ask or even switch pronouns for over the course of knowing them, but I also don't have a lot of white queer friends, so a name like 'Reba' or 'Susan' I think still throws me for a loop cuz I would associate it with someone more conservative.....I think)" [7]

This explicit association between specific singular *they* and gender/queer identity suggests that a driving force behind dsT is indeed the association with individuals who have intentionally made an effort to carve out a space for identity outside binary gender.

However, many comments do *not* make reference to close personal friends, and instead suggest that the use of dsT is associated with a generally tolerant, liberal mindset. Respondents often did so by way of commenting on gender biases that they did not want to endorse. Examples of such comments include;

[7] This comment strongly suggests that race/ethnicity are part of the social context through which speakers construct nonbinary (or any) genders; future studies with a more diverse participant pool will be able to address this directly.

(25) "I'm working on ""they/them"" singular pronouns feeling like second nature to me, but I'm not there yet. I want to get there. I get irked with masculine pronouns as the default, and I now understand why alternating masculine and feminine singular pronouns does not solve the problem of inclusivity."

(26) "For me this survey served to point out my own gender biases as to what sounds unnatural to me when it comes to pronouns"

Comments from participants who are "working on" acquiring dsT seem to be related to two factors: first, the political desirability of using dsT to show a commitment to gender equality, and second, the apparent difficulty in doing so naturally. These comments are likely from respondents who are somewhat further behind in the change, but who are aware of the "endpoint" of the change (full use of dsT for proper names) and find that endpoint to be an explicit goal of theirs. This also suggests that lifespan change is necessarily at least partially at play here: respondents (as in (27) and (28) below) who say that they previously did not use dsT but have come to accept it or use it themselves are presumably reporting on their change since adulthood.

(27) "I noticed that I found sentences with the singular *they* less natural, although I use the singular *they* myself." [Italics added]

(28) "I usually refer to non specific people in academic writing as "they", for example when talking about 'prisoners', but do not personally know anyone who chooses to be referred to as 'they' (although tangentially I follow some people on twitter who do). To me 'they' in a standard sentence reads unnaturally when referring to a specific individual, although **I accept that is probably an anachronism which I will learn to change**."

The metalinguistic statements in these comments do not carry information about whether speakers are actually changing their behavior over time—only that speakers perceive themselves to be doing so. These comments, like other types of self-reported data, cannot be taken as accurate representations of either the type or frequency of actual dsT useage over any individual's lifespan (cf Milroy and Margrain 1980 on self-reported data). The comments alone do not necessarily rule out a change in time, either. The apparent time difference I showed in the previous section does indicate that, if older speakers have

adopted dsT (and some have) then this may have happened well after the end of their critical period. Future research into this variable should include longitudinal studies, of panels if possible, to compare rates of dsT use from speakers at different points in their lives.

In this section I have shown metalinguistic comments from respondents to the acceptability survey probing the grammaticality of singular *they* with definite and specific antecedents. These comments give insight into the probable social meaning of the dsT variable for various speakers participating in the change. In the next section I will turn to the indexical meaning of specific and generic uses of singular *they*, which I will then expand into discrete indexical orders.

3.5.3 Different meanings for different levels of the variable

It follows from the referential meaning of singular *they* that different underlying sociocultural gender epistemologies are implicated by generic versus specific uses. I will briefly summarize these different referential meanings (to be much more thoroughly excavated in Chapter 4) before moving on to how these can be projected towards indexical meanings along with the apparent time data and metalinguistic comments I have shown thusfar.

As I discussed in the introduction, generic uses of singular *they* are frequent, and attested very far back (cf Curzan 2003 i.a.) in the history of English. Explicit prescriptive grammars prescribe both against and in favor of singular *they* as the preferred generic/indefinite pronoun; both accounts typically pit generic singular *they* against generic *he* as the (apparently) obvious competing form. This was reflected in metalinguistic comments from participants of Experiment Two, as in (28) above. The underlying assumptions for supporting the generic use of singular *they* are more conservative than those needed in order to use dsT: a speaker need only assume (or assert) that there is not a "default" gender in English, and that both *he* and *she* bring in assumptions of gender that are not necessary or useful in indefinite/generic uses. That *he*/*she* introduce a gendered interpretation is experimentally confirmed by a great deal of the literature (Hughes and Casey

1986, Moulton et al 1978, etc). It therefore is perfectly possible to argue (prescriptively) in favor of the generic use of singular *they* on the basis of equality of the (two) sexes—that argument usually postulates that generic use of *he* either erases the possibility of women from supposedly-neutral sentences (in the interpretation of *he* as not actually generic), or generic *he* otherizes women as a "marked" gender (in the interpretation that *he* is indeed generic, but shouldn't be). Referential uses of generic singular *they* are therefore implicitly linked with the desire for gender equality between men and women. It is *not* necessary to challenge the existence of men and women as binary and exclusive categories in order to use singular *they* in a generic context.

Definite and especially specific uses of singular *they* (dsT), however, *do* need a radical restructuring of the underlying understanding of gender categories in order to refer. In using singular *they* with a proper name, a speaker is either failing to assert a gender of a referent, or asserting a non-gender of a referent. (There are pragmatic considerations as to how an interlocutor differentiates between these two potential meanings, which is the core focus of Chapter 4). In the latter case, asserting a non-gender, the speaker must explicitly be working with the understanding that an animate, human, singular referent can be neither a man nor a woman. Linguist Geoff Pullum has noted an apparent difference (for him) between using singular *they* to refer to a non-human referent, versus using dsT to refer to a human:

> "*I would now say that although* *Chris left their pen *still sounds dreadful for some reason (perhaps because* ***whoever Chris is, he or she really does have a gender****), nonetheless it is possible to have a singular they with a singular proper name antecedent*." [emphasis added] (Pullum 2003)

The implication for Pullum is that gender is indeed the sticking point—a proper name can be coreferential with singular *they* if the referent is appropriately excused from having a gender. For the innovative users of dsT, the possibility of non-gendered humans is clearly a central concept—and the very high ratings of proper names with dsT by non-

binary and transgender participants in Experiment Two points towards the communities who make up the Avant Garde in the realm of gender.

3.5.4 Orders of indexicality

This section analyzes the apparent change in progress around dsT through the lens of orders of indexicality (cf Silverstein 2003, Eckert 2008, Johnstone and Kiesling 2008). In this case, I want to carefully differentiate between the term 'indexicality' as it is used in semiotics, syntax, and semantics, particularly when talking about referential pronouns (which 'index' particular referents) and the use of indexicality by sociolinguists. This differentiation is important because there will still necessarily be a link between the symbolic, sociolinguistic meaning of dsT (however abstracted) and the semantic, referential meaning of dsT referring to a known entity. Indexicality for sociolinguistics, on the other hand, refers to the link between certain forms of language (sociolinguistic variables, usually) and the social meaning that they convey about the speaker—for example, in Johnstone and Kiesling's work, /aw/-monophthongization may index a Pittsburghese identity. This sense of indexicality is necessarily dependent upon its legibility and interpretation in the sociocultural context in which a speaker (and/or hearer) is embedded.

As I suggested in the previous section, users who can produce or accept dsT referentially must have a mental conception of social categories that can accommodate a human referent who is "*neither he nor she*." The fact that transgender and nonbinary speakers are the most frequent users and highest raters of dsT, I start from the understanding that this system of mental categorization is linked with the life circumstances of transgender and nonbinary speakers. It follows that the LGBTQ+ community is (a decentralized) starting point for tracing the innovation of dsT. In order to trace the diffusion of the variable from its least-abstracted use through progressively more abstract symbolic uses, I will consider indexical orders as an axis orthogonal to the axis of apparent time (via age). This section will first review the model of indexical orders before applying them to the dsT variable; at the end of this section I will return to the matter of apparent time and indexical orders

in relation to one another.

In its sociolinguistic use, indexical orders are a theoretical framework for understanding the potential layering of abstraction that results from the diffusion of a sociolinguistic variable. The *first order* of indexicality is the theoretical base of the variable: for Eckert (2008) or Johnstone and Kiesling (2008), the first-order users of a sociolinguistic variable are those users for whom the variable does not necessarily convey any social meaning. Rather, it is a variant form that exists within a community that has arisen through language-internal factors, such as vowel shifting tendencies. If the rate of use of the variant is stable within the community of first-order users, then the variable will not necessarily convey a particular social meaning.

The second order of indexicality is removed from the first order by abstracting social meaning from its use. For Johnstone and Kiesling (2008), the variable in its first order of indexicality was the pattern of monophthongization among rural Texan English speakers. Abstracting from that variation which, community-internally was simply a fact of language, speakers extrapolated the tendency from rural Texan speakers by ascribing social meaning to that population. In the case of Johnstone and Kiesling's work (2008), rural Texans were evaluated by outsiders as authentic and local, so monophthongization could be used as a second-order index of authenticity and ties to locality.

It is possible to trace higher levels of abstraction through the social meaning ascribed to users of a sociolinguistic variable in a process that Eckert (2008) calls a "result of an ideological move." Thus, a third-order level of indexicality for monophthongization would abstract away from authenticity and local ties by extrapolating through ideological links: if an authentic/local/Texan identity is related to an association with conservative social values, a politician from NYC may take up monophthongization to invoke an identity of a down-home trustworthy country man with conservative social values.

The process of abstraction taken up by speakers in a social context is not (probably) as clean-cut as this, but for a theoretical model I liken indexical orders to any type of categorization where at some point theoreticians are going to necessarily have to draw a

boundary between what uses qualify as what order of index. In expanding and building a theoretical framework around this model, Eckert (2008) suggests a theoretical construct of an *n+1* order of indexicality. If first, second, and third-order indexes of a sociolinguistic variable exist, then the process of 'stepping' from a lower order to a higher one can be pinpointed as a specific type of sociolinguistic process. This *n+1* process is one where the *nth* order is interpreted as socially meaningful, and its social meaning itself is deployed for further meaning (Silverstein 2003). Eckert (2008) provides a useful example:

> *The emergence of an n + 1st indexical value is the result of an ideological move, a sidestepping within an ideological field. In order to understand the meaning of variation in practice, we need to begin with this ideological field, as the continual reconstrual of the indexical value of a variable creates, in the end, an indexical field. An indexical field is a constellation of meanings that are ideologically linked. As such, it is inseparable from the ideological field and can be seen as an embodiment of ideology in linguistic form. I emphasize here that this field is not a static structure, but at every moment a representation of a continuous process of reinterpretation. [. . .] As noted above, the Martha's Vineyard fishermen, in appropriating the centralized variant of /ay/, were not simply claiming to be Vineyarders but were making a claim about what a Vineyarder is. I would argue that as disagreements about the future of the island became more prominent in daily life, the terms of those disagreements entered into the local ideological field, available to be pointed to with the use of a linguistic variable already associated with Vineyarders.*

In examining the variation in singular *they*, I anticipate that there will be interaction between epicene vs. referential uses and the order of indexicality that can be attributed to its use by any speaker or group of speakers in context. I will attempt to differentiate between epicene and referential when possible to avoid confounding the analysis—but because of the grammaticality cline I have observed in this chapter, epicene uses will necessarily feed into the possibility of referential uses. The purpose of this section is to inspect the spread of dsT as it progresses through the social evaluation of its use, followed by reanalysis of "*what it means to be a* [*nth* order user]." Particularly, this section will use a theoretical means of extrapolation modeled after Eckert's *n+1* device in order to reconstruct earlier indexical orders which cannot be detected in historical corpora or through

the data presented in this chapter. In this section I will make reference to 1^{st}, 2^{nd}, and 3^{rd} order indexes and users of the sociolinguistic variable, with the disclaimer that these numerical terms are meant to represent subsequent orders of abstracted social meaning without necessarily implying temporal linearity.

Because *they* is a third person pronoun, it necessarily must have emerged in contexts where speakers referred to a third party; this means that earliest uses of dsT will not necessarily have been produced by transgender or nonbinary people themselves. Instead, the conditions that require third person reference imply that the earliest users of dsT (in its referential form) would be people whose social relationships either encouraged or necessitated its use. LGBTQ+ communities have historically often been predicated around difficulty with the binary gender system—this includes both people whose identities don't map neatly into binary categories (in contemporary terms, the 'transgender umbrella') and people whose sexual and romantic connections aren't neatly legible in heterosexual paradigms (gay/bi/pan/queer people). These communities have a direct link with the actual semiotic and referential meaning of dsT—there is and has been a need for referential pronouns that are neither *he* nor *she*. This is both to accommodate people who aren't easily referred to by *he*/*she*, but also to accommodate the needs of people to de-emphasize gender when, for example, discussing a same-gender partner in a potentially homophobic context. The use of dsT by speakers in close networks with LGBTQ+ people is a good candidate for analysis as the 1^{st} order indexical use of the variable, in part because the data from Experiments One and Two show that transgender and nonbinary people are by far the most prolific and comfortable users of the variable, but also in part because of the analogy of orders of indexicality given by Silverstein, Johnston & Kiesling, and Eckert.[8] This use is consistent with other first-order indexes in that its use 1.) emerged from language-internal forces and 2.) does not carry a special social meaning within the pop-

[8]This assumption implicitly assumes that close friends of LGBTQ+ people are likely to be LGBTQ+ themselves; anecdotally I can confirm that this is especially true in the transgender/non-binary communities that I have encountered and been a part of, but it is not necessarily always true. Future social network analysis of transgender/non-binary speakers will shed light on the matter.

ulation, and externally simply indexes membership of the population. I will first explore the full line of logic that extends from point 2, and then necessarily return to point 1 to address the question of whether any pronoun can be said to "not carry a special social meaning."

If the *n+1* process (where a variable gains new social meaning through abstraction fueling its diffusion) applies to the 1^{st} order index of dsT as indexing simply membership in the LGBTQ+ population, the 2^{nd} order level of indexicality should be predicated upon ideological moves around LGBTQ+-community-membership. Thus, dsT will have taken on meaning through the social evaluation of what it means to be gender-transgressive in the way that community members are. Abstracting from gender/queerness to qualities associated with gender/queerness (among people who are not necessarily gender/queer themselves) is likely to have been transmitted along the ideological channel of gender hegemony. Thus the index goes from "nonbinary genders exist and I am embedded among a community where that is common" instead towards "nonbinary genders exist and I wish to align myself with them." This abstraction ends up indexing social proximity to genderqeerness or nonbinaryness that can include people who have ideological ties with the community—the broader LGBTQ+ community and its allies outside of the subgroup of nonbinary/transgender people constitutes this. This 2^{nd} order indexes a very specific political orientation around gender transgression and queerness, which means that cisgender gays and lesbians or bisexuals, or binary transgender people, may be 2^{nd}-order users rather than 1^{st}-order. The 2^{nd} order indexical meaning of dsT is therefore LGBTQ+ people and their allies who share the abstracted meaning of an ideological opposition to hegemonic binary gender. The comments in (22)-(28) above support this: in (27) a respondent claims use of dsT for their own, but in (24) a respondent comments about their friends (suggesting that this respondent doesn't prefer dsT personally, but uses it a great deal in their daily life). The social diffusion suggested by these comments indicates that the second-order indexicality of dsT does indeed index its users as close personal friends or allies of the LGBT+ community from which dsT originally emerged. Thus the particu-

lar metalinguistic comments from Experiment Two that I have presented provide support for my analysis of these orders of indexicality.

The 3^{rd} indexical is a reinterpretation from the uniting ideological orientation of being queer or a queer ally (that orientation being opposed to hegemonic binary gender) where queer-allyship is interpreted as being correlated with progressive and liberatory political alignment in general (not just around gender). In this step of abstracting, the connective tissue is the political identity built around tolerance of those different from oneself. Using dsT in this 3^{rd} level of abstraction does not require that speakers personally know any nonbinary/genderqueer (or other trans) people, nor even that speakers are themselves LGBTQ+ or closely tied to that community. Instead, this is an order of abstraction that allows speakers to mark themselves as progressive and not homophobic or transphobic. This is a definitionally oppositional identity, defined by a performative investment in not being perceived as harboring unpopular or undesirable prejudices. This indexical order casts a much wider net, and can include a broader portion of the general population than either of the previously-discussed orders. This also is the group where the ideological investment is more distant from the actual semiotic need filled by dsT, meaning speakers in this group may not even know any nonbinary or transgender people at all. This separation also rightly accounts for the speakers who report that they are "trying" to incorporate dsT into their grammar despite difficulty—the comment in (25) is of this type. (Inasmuch as there is any relationship between increasing indexical orders and apparent change over time, it may also be that these speakers include those who are newest to the variable, or who are older and further behind in the general change.)

Returning to the issue of the 1^{st} order indexical meaning of dsT: I earlier suggested that, like other 1^{st}-order indexical uses of sociolinguistic variables, dsT had no particular social meaning and was a result of a language-internal change. This is an oversimplification based on comparison with phonological variables—for Eckert, Johnstone and Kiesling, and others, there is of course no symbolic "meaning" in phonological variables because, signs being arbitrary, something like monophthongization or vowel centraliza-

tion as phonological processes cannot take on semantic load without some kind of serious morphosyntactic reanalysis. However, dsT is in fact a "load-bearing" word, in that the apparent change over time is at least partially responsible for, or caused by, a shift in meaning of the word itself. Moreover, dsT did not just emerge fully-formed in the genderqueer community in the 1980s (or at any other time) without some speakers actually doing the grammatical innovation required—but actually identifying the path of that innovation is difficult to observe directly. The fact that coreference for pronouns is not easily automatable means that it is very difficult to use automatic methods to search historical corpora for the first/earliest attestation of dsT *referentially*; all of the attestations for singular *they* in works tracing the pronoun's history have only shown epicene uses.

Because of the impossibility of using corpus data to investigate the initial innovation (presumably from language-internal factors) that introduced dsT into English, I instead propose that the method of extrapolation Eckert describes can in fact be reversed: rather than analyzing the *n+1* process of adding ideological/social meaning onto a sociolinguistic variable, I will use an inferential process of suggesting a (reconstructed) context and meaning for the use of a variable at one *less* level of abstraction. This form of inference is purely theory-internal and not a process that speakers of a language are likely to utilize; essentially, I am going to more deeply investigate the literal (pragmatic and semantic) meaning of dsT by *removing* an associated social meaning. Just as applying Eckert's *n+1* process to 1^{st} order indexes predicts 2^{nd} order indexes, I will apply *n-1* to the 1^{st} order of dsT to infer a sort of "zero-th" order, a reconstructed meaning for dsT under the layers of sociolinguistic meaning.

The concept of a 0^{th} order indexical meaning is only relevant for morphosyntactic variables that carry some underlying semantic or pragmatic meaning—this process could not be applied to monopthongization, for example, because there is no semantic content in the sub-morphemic units involved. This is also an extrapolation based on the assumption that semantic and pragmatic meanings are one of the fundamental elements of language that can drive linguistic change—put another way, I am assuming that if speakers

need to express a meaning and conventionalize it, and if a linguistic resource does not already exist in a way that serves that need, they will invent or repurpose some part of the language to fill the need. The way that these innovations likely manifest are through language-specific analogy. In the case of dsT, the pre-existing epicene uses of singular *they* were available for analogical extension to serve the needs of the language users.

In my discussion of the 1^{st}-order users of dsT above, I suggested that nonbinary/genderqueer people and their close friends had a special need for a non-gendered, animate pronoun. This need has been commented on by various language experts before, and pronouns have been coined with the intention of filling that need. What this suggests is that such a need is not restricted to nonbinary/genderqueer communities. The point of innovation—using singular *they* referentially—may have in fact occurred at several points throughout history. The step of moving from a purely analogical use of dsT towards a conventionalized use of dsT within a community is the necessary diffusion required to move from a 0^{th} order use (which is an isolated use brought on by an idiosyncratic need) to a 1^{st} order use (which is a use by a population) is still necessary. That step is one that is facilitated by shared needs and social norms—in order for a nonbinary/genderqueer-centric community to take up dsT in the first place, there must have been a point at which the variable successfully conferred a shared semantic and pragmatic meaning beyond the intentions of a single speaker. The shared social need, for dsT in particular, is the need to refer to an entity without ascribing binary gender to the referent. Individuals have various causes to need this expressive ability, but also the nonbinary/genderqueer-centric community has historically been where multiple people can share this need consistently enough to support conventionalization, moving the variable from its 0^{th} order use to its 1^{st} order use.

In order to confirm the hypothesis that I have proposed here, where dsT is most likely to have sprung from LGBTQ+ communities that have particular need for gender-neutral specific pronouns, several field, corpus, and experimental methodologies may be informative. Because my data have been based largely on United States (and specifically Seat-

tle) English speakers, it may be possible to observe English-speaking communities that are not yet at the point of saturation that I have observed; it is, for example, hypothetically quite possible that dsT will go from obscurity to prominent use in a more isolated English-speaking community. Observing this innovation and spread in a smaller and more closely-attended community in real time (or shortly thereafter in apparent time, depending on study design) would be highly informative for how dsT serves the communicative needs in a particular context. Due to the high level of discussion and metalinguistic awareness through internet communication, however, it is likely very difficult to find a community where dsT can innovate from *purely* language-internal mechanisms; instead, it is likely that the innovation will be initially transmitted online and later diffused through in-person social behavior. Some very recent research has also suggested that acceptance and use of dsT is correlated with direct social acquaintance with non-binary people (Ackerman et al. 2018 a.o.); this itself is strong evidence that LGBTQ+ identity/social proximity is at the very least a crucial part of the puzzle.

Online discussion and diffusion may, however, provide the potential for a part-corpus, part-online ethnographic approach like that used by Zimman and Hayworth (2018). In that study, Zimman and Hayworth built and analyzed a corpus that consisted of a Livejournal community of transgender English-speaking members, and tracked the change in use of certain trans-related terms over the life of the community. This approach has the advantage of giving realtime data and significant interactional/social context to support explanations of why changes are happening.

Any corpus-based research into dsT will be highly reliant on this type of relational and interactional context, since the gender of referents can only be estimated with this contextual knowledge (if at all). It is not necessarily possible to determine, for example, whether dsT with no local antecedent is anteceded by a specific person or not. Additionally, as I discuss in Chapter 4, dsT can be used not only to refer to non-binary referents but also to obscure or de-emphasize gender even when a referent has a (known or unknown) binary gender identity. Even aside from the labor-intensive task of hand-analyzing antecedence

and reference for pronouns in corpora, not all tokens will be obviously connected to the actual gender identity of the referent at all.

Other rich grounds for corpus explorations would be serial conversation-based broadcasts that are broadcast over a long period of time; podcast transcripts could serve this purpose very well. As an example, *My Brother, My Brother, and Me*, contains a years-long archive of conversational speech by (mostly) the same participants; this type of archive would be very informative for examining changes over time in individual speakers. I can (somewhat anecdotally) report, for example, that the speakers *My Brother, My Brother, and Me* seem to use dsT more over time between 2015 and 2019. This is in part because the "advice" questions and jokes are largely based around anonymous askers; in earlier episodes, the McElroys largely opted to assign either *he* or *she* to hypothetical question-askers either arbitrarily or based on stereotypical or contextual information in the questions submitted. In more recent episodes, however, the McElroys appear to have opted for dsT in more cases—and in recorded live shows, audience members apparently spontaneously began stating their pronouns before asking questions.

Finally, experimental investigations into the link between LGBTQ+ identity and dsT use may be modeled off of existing studies on dsT such as the ERP study by Prasad, Morris, and Feinstein (2018); in order to specifically probe the effect of proximity or social familiarity with LGBTQ+ communities, Prasad et al.'s experiment could be fully replicated with the simple addition of post-test interview questions about participants' friendships and closeness with trans/non-binary people.

The methodological suggestions I make here can all potentially build off of existing orthogonal research on language (especially pronouns) and transgender identity; if it is the case that future studies consistently show links between use/acceptance of dsT and social connections with transgender and nonbinary individuals, I would consider those findings to constitute significant support for the hypothesis I have set out in this section.

While this chapter has been aimed towards investigating the variability and social meaning of dsT (thus indexical orders 1 and above), Chapter 4 following this is a more

rigorous investigation of the semantic and pragmatic content of dsT itself (the 0^{th} order). In particular, I will pay attention to *why* speakers may have individual needs for an ungendered third person singular pronoun, and how they utilize that resource to accomplish conversational goals.

The 0^{th} order use of dsT that I have inferred here is the use of singular *they* in a way that did not involve a "great leap" of grammatical machinery at any point. Rather, the use of singular *they* to refer to a particular referent is one that can be easily transmitted through subtle ambiguity. In Chapter 5 I will discuss the mechanism of reanalysis through ambiguity, particularly as it pertains to the internal syntactic and semantic structure of a third person pronoun.

3.6 Conclusion

In this chapter I have shown experimental evidence that there is an ongoing language change around singular *they*. I reviewed previous literature exploring both the grammaticality and attestation of older forms of singular *they*, most of which were with indefinite or generic antecedents; I also showed the gap in existing literature around the use of singular *they* for a definite, specific antecedent (dsT). Experiment One showed differences in production of dsT by different speakers, patterning by various demographic speaker variables including age and gender. Experiment Two was a large-scale acceptability survey of singular *they* compared with *he* and *she* for various types of antecedents; the data from Experiment Two showed that there was an age difference in ratings of singular *they* but not for *he* or *she* (despite comparable gender mismatches), and that other speaker variables such as gender identity and transgender status were also related to the ratings of singular *they*. My analysis of the data from both experiments posits an apparent time difference reflective of an ongoing change in real time. I also discussed the social meaning of dsT, particularly around its links with transgender/nonbinary communities, and proposed an analysis using orders of indexicality to track how the change would be transmitted as a social mechanism.

The next chapter also deals with singular *they*, especially dsT—this will be a socio-pragmatic analysis of actual attested uses of dsT, in alternation with other pronouns (*he* and *she*), as a linguistic resource that speakers can exploit to accomplish various conversational ends. The final chapter of this dissertation ties together the syntactic phenomena shown in Chapter 1 with the sociolinguistic and pragmatic phenomena in this chapter and the next; Chapter 5 will show how a single syntactic analysis can correctly account for and robustly explain all of the variation shown.

Chapter 4

PRONOMINAL SHIFTS IN DISCOURSE

4.1 Introduction

The conversational maxim *Maximize Presupposition!*, from Heim (1991), proposes that when a speaker has an option between two propositionally-equivalent utterances, they must choose the one that has the strongest presuppositional content. Phi-features on pronouns have been proposed to be presuppositional (Sauerland 1998), which implies that speakers should always choose the most 'specific' phi-specified pronoun available to them. Under an analysis like Bjorkman's (2017), *they* is radically underspecified for morphological features: it has no number, person, or gender features, and all other pronouns are differentiated from *they* by the presence of those features. If phi features are presuppositional, and *they* has none of those, then *they* should never be used except in cases where every other pronoun is infelicitous.

This is clearly not what happens. As shown in Chapter 3, particularly with the production study in Experiment One, speakers opt to use *they* when other pronouns (namely *he* or *she*) would theoretically be available. The current chapter is aimed at exploring the sociopragmatic consequences of the availability of singular *they*. For the purposes of this chapter I will focus primarily on instances where speakers are of the innovative type described in Chapter 3—that is, they have dsT (definite, specific singular *they*) available to them grammatically. It is these speakers around which Bjorkman modeled her morphosyntactic analysis, which I will further be building upon in this and the following chapter. The data from this chapter will include both attested and constructed (based on attested) examples of uses of pronouns where *Maximize Presupposition!* should supposedly prescribe one particular pronoun, as well as examples where pronoun use ex-

presses some meaningful content besides "biological sex" (e.g. gender performance, political stance, and other expressive uses).

The aim of this chapter is to pinpoint where the static, presuppositional analysis of pronouns fails to account for the diverse and expressive uses that are typically not analyzed in formal texts like Heim and Kratzer (1998) —the reader will notice that a *lot* of my examples are from or about transgender speakers, gender-nonconforming speakers, drag queens, and so forth. These speakers are frequently excluded or disregarded in traditional analyses, but a complete and explanatory theory of language use in context should be able accommodate behaviors that are non-normative but apparently grammatical and productive.

This first section will cover the relevant background needed in order to account for the data I present in Section 4.2.3: I will review Gricean maxims, some fundamentals of politeness theory that I will employ, and Du Bois' conception of stance (the Stance Triangle, which will allow us to triangulate evaluations of third person referents who may or may not be participants in a conversation). I will then explore the interaction between the availability of dsT with maxims of quantity and relevance, politeness (around norms of 'guessing' gender when pronoun use is not obvious for the speaker), and stance (including examples of misgendering and, at the other end of the spectrum, gender play). I will end this chapter with an exploration of how sociopragmatics can and should interact directly with the (morpho)syntax, which will provide the needed background for my syntactic proposal in Chapter 5.

4.2 Part 1: Pragmatic Frameworks and Background

4.2.1 Background on Maxims:

The propositional semantic content of an utterance is the main contribution of the utterance, and one that has truth conditions within the context of the utterance.

(1) a. Nick ate a cupcake.

b. *Proposition*: eat(Nick, cupcake)

Entailments of utterances like (1) are the meaningful consequences of the proposition. A proposition α entails a proposition β iff it is the case that when α is true, β must also be true. The proposition in (1) entails, for example, *The cupcake has been eaten*. Presuppositions are a particular type of entailment where a proposition will not be felicitous if another proposition is not true; presuppositions are the preconditions for propositions to be interpretable. A proposition α presupposes a proposition β iff it is the case that if β is not true, α is not felicitous. The proposition in (1) presupposes *Cupcakes can be eaten* or, even more generally, *Cupcakes are physical objects*. The propositional content and entailments of an utterance are generally taken to be semantic issues, but utterances convey meaning that is not necessarily related to propositional content or entailment. This section reviews how pragmatic or contextual inferences can contribute meaning without necessarily adding entailments (or presuppositions) to the semantic content of an utterance; I will primarily review the framework given by Grice (1975).

H.P. Grice, building on JL Austin's speech act theory (Austin 1975), proposed that conversations were guided by certain principles of cooperation between interlocutors. In general, cooperative interlocutors can be trusted not to lie, deliberately mislead, or interject with random unrelated statements. The principles Grice set forward have been the basis for extrapolating conversational implicatures—these implicatures are neither part of the overt, propositional content of an utterance, nor assumed to be part of the common ground understanding between interlocutors (presupposition) (Grice 1968, 1975). Implicatures are extrapolations of meaning that infer from the context of an utterance, and from what a speaker chose *not* to utter. Gricean implicatures can very strongly communicate meanings that become conventionalized in sociocultural contexts. For example, in (2), the propositional content of the utterance does not necessarily *entail* or *presuppose* the implied meaning—instead, that implied meaning is taken from social context. ("Why else would you be asking?")

(2) Context: I come home from work and ask my spouse, "**Have you eaten yet?**"
Implicature: *I'm hungry and I'd like to have dinner with you*.

In Section 2.1 I will use Gricean maxims to explore possible conversational implicatures that can arise from particular uses of pronouns—particularly the choice between singular *they* and *he/she*.

Conversational maxims are a part of how listeners will infer information based not only on the propositional content of the speaker's actual utterance, but also inference based on alternative ways to express the same content—what a speaker *doesn't* say is often just as informative as what they *do* say. When multiple possible utterances have the same propositional content, then the choice between them is informative to listeners; in this chapter, I will challenge purely presuppositional accounts of pronoun gender in English, paying particular attention to where optionality is introduced (or, more accurately, increased) with the emergence of singular *they* as possibly referring to a specific person. Specifically, I will argue that the addition of singular *they* to the pronominal paradigm of English enables the gender features of pronouns to be interpreted through conversational inferences rather than presuppositional conditions.

In the case of pronouns, I will continue down the line of reasoning I set out in Chapter 2: rather than an absolute (and binary) "*x is female*," I there take *she* to mean "*it is appropriate to call x she*." If this seems circular, it's because I am deliberately embedding the arbitrariness of social convention into my denotation for pronouns—this is an attempt to abstract away from "sex" as some static, inherent, or obvious property of referents, highlighting instead the nature of gender as performatively constructed through social behavior (including language; Butler 1993 i.a.). This also pushes third person pronouns in English closer to honorific systems—the difference between *she* and *they* is not exactly the same as the difference between *Usted* and *tu*, but comparable means of analysis will prove useful here (see also the discussion of Brown and Gilman 1960, Raymond 2016 in Chapter 1). In what follows I will discuss briefly some attested examples of singular *they*

used in online discussions, which will vaguely illustrate the extent to which pronouns are a matter of appropriateness and inference; I will then move on to a more detailed analysis of alternations between pronouns for a particular (specific) referent, using particular Gricean maxims to analyze possible contextual meanings for each pronoun.

The conversational maxims that will be the most important for this discussion will be *Quantity*, *Quality*, and *Relevance*.

The maxim of *Quantity* is given in (3) below. The maxim can be synthesized to mean, essentially, that speakers prefer to provide just the right amount of information—not too much, nor too little. Minimizing information may be a principle of economy, which is something that crops up in various other levels of linguistic structure; economy can be related to the tendency for assimilation, pronominalization, or ellipsis, for example. Its counterpart—maximizing information—may be a principle of clarity, also observed at various linguistic levels; clarity may be a driving force for redundancy in a language, including redundant elements like vowel harmony or morphological agreement.

(3) QUANTITY (Grice 1975:45)

a. *Make your contribution as informative as is required (for the current purposes of the exchange)* [=Maximize Quantity!]

b. *Do not make your contribution more informative than is required* [=Minimize Quantity!]

The principle's two sub-parts given in (3) are relevant to gendered pronouns because, for speakers who have the grammatical option of using dsT wherever *he* or *she* would also be available, there is always a choice between a *more vague* and a *more specific* pronoun. Following the logical path of dsT's lack of gender features morphosyntactically, dsT can and does get used by speakers to refer to referents who are of a binary gender identity. The choice, then, depends on whether it is conversationally more advantageous to include more information (if *he* or *she* are options, they are more informative ones—obeying Maximize Presupposition! apparently) or less information (since dsT gives no information

about gender at all. This optionality is a natural consequence of the grammatical change that occurred in order to enable dsT in the first place (see Chapter 5 for details). This enables speakers to adjust the amount of information in triangulation between Quantity and other conversational maxims.

Grice's other maxims that are in contest with Quantity are Quality, Relation, and Manner, given in ((4)-(6)) below. *Quality* assumes that cooperative interlocutors avoid giving misinformation unless otherwise compelled to; *Relation* assumes that conversations have continuous connections between topics, and that introducing new topics would ideally be a natural consequence of previous topics. *Manner* is the maxim of appropriateness.

(4) QUALITY [is] a supermaxim – 'Try to make your contribution one that is true' – and two more specific maxims: (Grice 1975:46)

   a. *Do not say what you believe to be false.* [=Don't lie!]

   b. *Do not say that for which you lack adequate evidence.* [=Don't guess!]

(5) RELATION "*a single maxim, namely, 'Be relevant."'*

(6) MANNER "*Be perspicuous"*

   a. *Avoid obscurity of expression*

   b. *Avoid ambiguity*

   c. *Be brief (avoid unnecessary prolixity)*

   d. *Be orderly*

In ranking the relative importance of these conversational pulls, speakers have the constant option of prioritizing one maxim over another. A speaker might tell "a little white lie" (*Sorry I couldn't come, I was ill*) to avoid giving too much information (*I didn't come to your party because I was in bed all day crying because I have depression and parties make me miserable*). Likewise, a speaker might give a deliberately underinformative answer (*I have two cats*) to avoid saying irrelevant facts (*I have one cat who is my own and I live with another one but it technically belongs to my roommate*).

In terms of pronominal gender, various rankings of these three maxims will allow various strategies for including more or less gender specification in a conversation by way of pronouns. A speaker may use *they* in order to prioritize relevance by intentionally being more vague when the gender isn't important to the conversation—this may be advantageous if you know your interlocutor to be a notorious sexist, for example, and you want to get through a whole conversation without trying to defend the fact that your boss is a woman. Relatedly, one may use *they* to flout Quantity by intentionally obfuscating the gender of the referent—it's not that it's not relevant to mention your spouse's gender, but if for example you are gay in a hostile environment then you know your options are either to lie about the gender of your spouse or face homophobic retribution; this is a subtle difference in reasoning, and one that rests on the internal motivations of the speaker who is aiming to optimize the language in order to achieve their conversational goals. One final (and very common) example of using *they* (when either *he* or *she* might theoretically be available) is in the case of referring to someone whose gender you're not exactly sure of—either you haven't met them, or you can't read their gender expression, or something similar. In this case a speaker prioritizes *Quality* above *Quantity*—in using *they* one can avoid guessing, and in doing so avoid a possible censure for guessing wrong.

This last example, in which Relation and Quantity are both supersumed by Quality (and specifically avoidance of making a conversational gamble) demonstrates the submaxim of Quality given in ((4)b) above.

Grice's wording for ((4)b) does not specify whether or how guess-aversion may be related to politeness; I will elaborate on that relationship in Section 4.2.2 below. Instead, this aspect of avoiding accidental misgenderings is rooted in the potential embarrassment of saying something untrue. Using *they* rather than *he* or *she* does not necessarily relieve the problems of misgendering related to misattributing or misidentifying a significant part of a person's identity; rather, it is a conservative guess that attributes as little information as possible. (7) below gives an attested example of an internet poster on a question forum who uncomfortably alternates between *they* and *he/she*, demonstrating

this tension:

(7) Q: *Is it rude to ask someone if they're male or female?*
A: I think it depends on the situation and one's relationship with **them**$_a$. I answered a similar question where I had a client that I didn't know **their**$_b$ gender. I wanted to take **him/her**$_c$ on a business lunch but was afraid to not knowing **their**$_d$ gender. And didn't feel comfortable asking.[1]

I use subscripts to allow reference to each individual pronoun in the post; however presumably (7b,c,d) are all coreferential. In (7), the answer to the question uses an instance of a generic *they* in (a), then a definite specific *they* in the next sentence in (b), then *him/her* in (c), then went back to *they* in (d). This alternation suggests that the speaker would prefer to be able to attribute some binary gender, but simply cannot collect enough evidence to make a confident guess. Using the coordinated *him/her* in ((7)c) does suggest that the speaker is reluctant to consider potential non-binary options; it may also be a reflex of a speaker who is not consciously aware that dsT is part of their grammar, and the *him/her* is a type of hypercorrection to try and avoid it.

This type of hypercorrection seems to be an effect of mismatch between conscious and unconscious beliefs about grammar, a phenomenon which is not unprecedented: as a widely-circulated internet joke in the service of pro-singular *they* language commentary, speakers will occasionally use singular *they* in the process of saying that *they* is strictly plural. There are several examples of this available online:

(8) "I mean, I'll call someone whatever pronouns they want as long as it's she/her or he/him"[2]

(9) **Carol Off**: Professor Peterson, why have you said you don't recognize another person's right to determine what pronouns you use to address them?

---

[1] Source: https://www.quora.com/Is-it-rude-to-ask-someone-if-theyre-male-or-female-If-so-what-are-some-alternative-ways-of-finding-out-someones-gender

[2] Source: https://twitter.com/mczephzeph/status/1021204286682718209

> **Jordan Peterson**: That's right. I don't recognize that. I don't recognize another person's right to decide what words I'm going to use, especially when the words they want me to use, first of all, are non-standard elements of the English language and they are constructs of a small coterie of ideologically motivated people. They might have a point but I'm not going to say their words for them.[3]

The example in (9) is particularly popular fodder for metalinguistic humor because Peterson's use is, in fact, dsT (used for a referent of unknown gender, since they're hypothetical) while the example in (8) is a (grammatically more conservative) indefinite use of singular *they*. The comedy in discussions of utterances like (8)-(9) derives from the apparent contradiction in statement of refusing to use a form *while in the act of using that form*; this type of linguistic unawareness may be an indication that, despite the very high salience of singular *they* and its surrounding social commentary, speakers are not at all reliable in their own reports of whether or not they use it themselves. Returning to the use of *him/her* in (7), then, it seems possible that although dsT is apparently part of the speaker's grammar enough to produce it several times, the speaker's own discomfort and awareness of more prescriptivist-sanctioned forms forces the alternation. In all three of these instances shown in (7)-(9), the use of dsT is NOT a (respectful and conscious) way to refer to a non-binary referent, but rather an instance of referring to a referent for whom gender features are not available—due to a lack of clear social cues, or because the referent is hypothetical.

Of course, using dsT to avoid potentially guessing wrong can backfire; if your interlocutor infers that you are intentionally avoiding guessing, they may assume that your discomfort with guessing is a reflection of some gender transgression. Part of the important power of conversational maxims comes from the assumption that interlocutors are playing by basically the same rules; thus, meanings can be inferred from the choice between various options. In the case of third person pronouns in English, there are very few

[3]Source: https://www.cbc.ca/radio/asithappens/as-it-happens-friday-edition-1.3786140/i-m-not-a-bigot-meet-the-u-of-t-prof-who-refuses-to-use-genderless-pronouns-1.3786144

options and many more potential relative rankings of maxims that may have motivated them. If only the three maxims discussed here (Quantity, Relevance, and Quality, leaving aside Manner) are potential factors, this still yields six distinct possible rankings that may have motivated the use of a single pronoun. In the next section, I will discuss in more detail data (constructed and attested) that demonstrate the flexibility with which speakers deploy different rankings of conversational maxims in both generating and inferring meaning from pronouns.

4.2.2 Background on Politeness Theory

In Section 4.4 I will move to explore politeness theory, primarily looking at the conception of speech acts as part of a negotiation between interlocutors to preserve the 'face' of both themselves and their partner. Goffman (1967) conceives of politeness in terms of 'face,' or the need to maintain social and emotional esteem through interactions. Brown and Levinson (1987) expand on this, separating the social and emotional needs of interactors into two categories; *positive face* is an individual's need to be well-thought-of, to maintain a positive and congruent self-image with their identity; *negative face* is an individual's need to be free from obligations or impositions upon their will. In this theory, 'politeness' refers to the cooperative system where all participants are attempting to serve the needs of both their own and their interlocutors' face, positive and negative. In this cooperative system, it is beneficial for participants to attempt to meet the needs of their partners—however, sometimes trade-offs are necessary, and interactors cannot meet *all* of each other's needs.

When looking at the way that face needs intertwine with gender identity, expression, and referential language, it seems at first that the primary face needs to be attended to will be those of positive face—these are the desires for people to be spoken of kindly, and to maintain their own positive self-image and identity. When a person is misgendered, this is a direct threat to their positive self-image. It is not the case that certain gendered pronouns are inherently insulting and others are inherently deferential or respectful; if

such interpretations did exist, this would be dependent not only on the context of a single conversation but also the context of the general sociocultural setting shared by interlocutors. Rather, any respect or disrespect in effect through the use of gender features is due to the implication or statement that a referent is not being interpreted as a gender congruent with their identity, or that their interlocutor does not see them the way they see themselves. By this measure, using *she* to refer to someone who sees themself as obviously male is just as threatening to the referent's positive face as the reverse. The same context-sensitivity can be applied to number or person features that are conventionally attached to respect or politeness in pronouns—so, for example, *thou* (a singular pronoun) could be considered inappropriate enough that *"I thou thee, thou traitor"*[4] was a way to convey a serious insult. The formal difference in number features are not *inherently* linked to politeness, but are capable of conveying politeness or impoliteness within the conversational and sociocultural contexts in which they are embedded.

In addition to the ways that positive face needs that can be served through the use of pronouns, negative face needs will also be addressed: in this case, metacommentary from some speakers suggests that corrections of misgendering constitute a violation against their negative face needs—that is, it is an imposition upon the free will of a speaker to try and dictate to them what pronoun they ought to be using.

I will also need to gently amend the domain of interactional politeness in order to adequately cover the phenomenon of third person pronoun use. The amendment is only that referents, either absent or present, are also considered to be interactants in a conversation whose face needs must be considered when making conversational moves. It may be the case that the needs of third person referents are generally ranked lower than the needs of the speaker or the addressee, but they are still present and active in the social calculation.

In Section 4.4 I will present primarily data from interactors who are strangers, because the assumed politeness in interacting with a stranger comes more from social convention

[4]Sir Edward Coke to Raleigh, reported by Mencken (1945) a.o.

than from personal relationship history. In these cases, speakers will use maxims of politeness to attempt to choose a gendered pronoun—these maxims can be variably ranked for different speakers and in different situations, thus causing variable behavior. The politeness constraints that contribute to choosing a gendered pronoun include (at least) (10)-(12) below:

(10) *Don't ungender!* Failing to attribute a person's gender to them or nullifying a person's gender is an imposition on their positive face.

(11) *Don't misgender!* Asserting an incorrect gender for a person is an imposition on their positive face.

(12) *Don't correct me!* Dictating what pronoun a speaker should use is an imposition on their negative face.

It is worth noting that social meaningfulness of (10) itself is derived from conversational implicatures; if a speaker uses *they* (or avoids pronouns completely) then a hearer may infer that this under-informativeness is a sign that the speaker is willfully withholding gender information; thus, any calculation involving (10) will also be predicated on how conversational participants interpret the meaningfulness of ranking Quantity, Quality, and Relevance in various ways. For this reason, discussions of ungendered pronouns and degendering throughout this chapter will necessarily be dependent upon the interaction between conversational maxims and politeness constraints. Likewise, the social meaningfulness of (11) and (12) is necessarily related to a speaker's stance, both towards the addressee and referent in particular and towards gender ideology in general. The next section explains the model of stance that I will use in this chapter to account for these patterns.

The dataset used in this chapter includes examples selected to allow testing of the proposed possible rankings of (10)-(12), since each possible ranking may correlate with different speaker behavior.

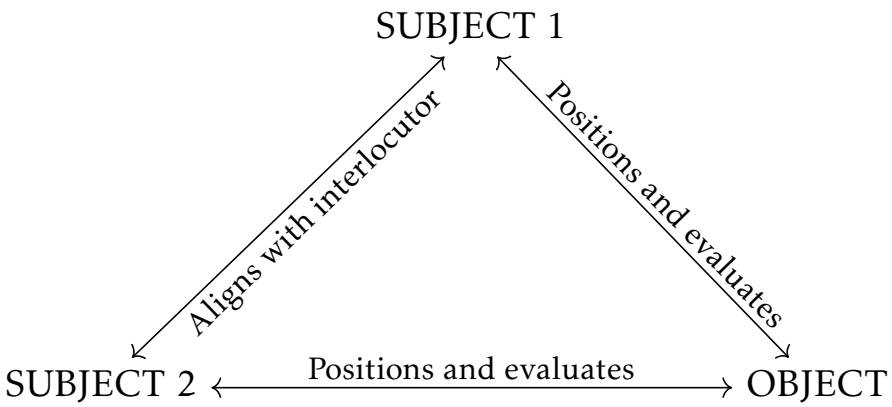


Figure 4.1: DuBois 2007 Stance Triangle

4.2.3 Background on Stance

This section introduces a theoretical model of stance that accommodates the relationship between speaker, addressee, and referent, as well as showing how gendered language can be evaluative or can position subjects in relation to referents. Chapter 1 discusses more general background on the theory of stance I am following.

DuBois' conception of the stance triangle incorporates three necessary components for stance-taking: the stance-taker, the object of stance, and the alignment of the stance with other stances taken within the conversation (DuBois 2007). These three components are the arguments necessary for a stance to be comprehensible in context; the relationship between these components compose the type of stance itself (which can be evaluative, positioning, aligning, epistemic, affective, etc.). In a stance-taking act, a subject evaluates or positions themselves in relation to an object and aligns themselves with relation to another subject (an interlocutor) who is themselves positioned in relation to the stance object. This three-way relationship relies on two or more subjects who share an object. In the case of pronouns, these roles align closely with the different persons. Section 4.2.3 below is DuBois visualization of the stance triangle that results from these relationships (2007:163).

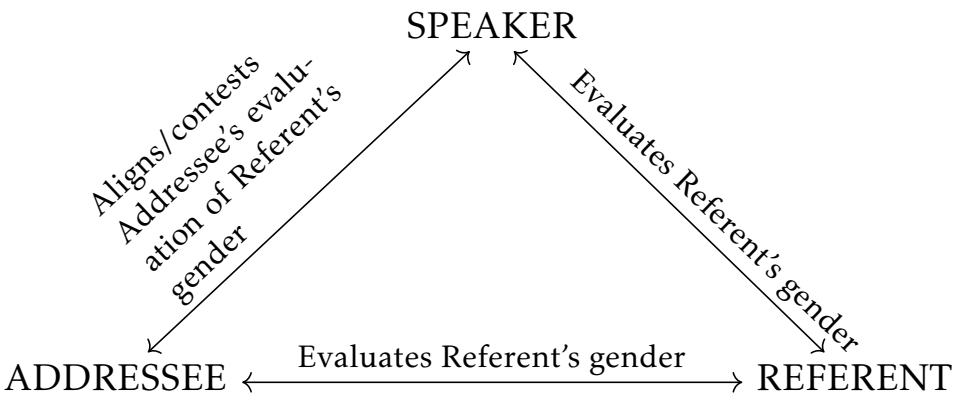


Figure 4.2: Evaluation of gender and alignment between speakers

The stance-taker, or stance subject, is often either explicitly linked to the speaker of an utterance by a first person pronoun, or is implicitly linked to the speaker in sentences without a first-person pronoun. The use of (animate) third person pronouns will necessarily position a person as a stance object.[5] Because third person pronouns in English can be used to convey information about gender, and because gender is a social categorization that is subjectively evaluated, the use of a third person pronoun can be a way for a stance subject to evaluate or position themselves in relation to the third person referent (stance object). Section 4.2.3 below shows a reformulation of DuBois' triangle, adapted to target gender specifically.

In Section 4.5.5, where I discuss 'closet pronouns,' I will show how the relationship between the evaluation of a referent's gender by a speaker and by an addressee can influence each other in such a way as to change pronoun use.

In this section of the chapter, I have briefly summarized the three pragmatic models that will be needed to account for gendered pronoun alternations: conversational maxims, politeness constraints, and stance triangulation will all play a part in any speaker's choice of a third person pronoun in English. In the next section I turn to data which

[5] Use of second-person pronouns can also potentially position an interlocutor as a stance object; this would be a useful framework for analyzing social uses of second person pronouns like T/V pronouns in German and Romance.

demonstrate the usefulness of these frameworks.

Part 2: Data and analysis

4.3 Data about Maxims

The introduction of dsT into the grammar of speakers also introduces optionality for all pronouns: because dsT has no gender features, speakers can (and do) use *they* to refer to referents even when the referent has a known, binary gender. For speakers without dsT available in their internal grammar, animate singular referential pronouns are limited to *he* or *she*; presumably the forced binary choice limits the actual effects of optionality to relatively extreme circumstances. With dsT as a potential competitor, however, more subtle forces can sway pronominal choice, meaning that any choice gains more weight from conversational implicatures. This section will show that, given a choice between a binary pronoun and dsT, there is no 'default' or completely unmarked choice: rather, using either *she* or *they* for a particular referent will necessarily force the listener to attempt to reason why the speaker made this choice. The inferential process by which the listener construes implicit meaning is guided by variable rankings of the conversational maxims of Quantity, Quality, and Relevance.

Ranking *Maximize Quantity* (as stated in (3a)) over Relevance or Quality will result in more specific gendered pronouns, even when gender may not be relevant to the conversation, or even if the speaker is not absolutely sure that their choice is the correct one. Likewise, ranking Relation over Quantity will result in a more vague pronoun, even if the speaker has access to more specific information that they *could* opt to share. The way that a listener infers a speaker's meaning comes from a combination of knowledge of these rankings and awareness of social context. Below I will demonstrate the various construable readings from a single pronoun choice.

I propose a thought-experiment of relative simplicity to demonstrate the effects of various rankings of these constraints. Suppose a speaker ("Kirby") is talking to Avery about

Laurel[6]. Avery is two degrees of social separation from Laurel but they know each other online. Kirby has the option, theoretically, of calling Laurel either *she* or *they*. Kirby's utterance is specific:

(13) Laurel$_1$ told me that ____$_1$ got a job in San Francisco.

What pronoun Kirby chooses for (13) does not change the propositional content of the utterance. Laurel told Kirby that Laurel got a job and that the job that Laurel got is in San Francisco. The propositional content alone is simple enough. However, what pronoun the speaker chooses (between *she* and *they*, specifically, leaving aside *he* entirely right now) conveys much more (implicit) meaning. The (primary, not exhaustive) possible rankings[7] of Quantity (minimize/maximize), Quality (Don't lie!/Don't guess!), and Relation follow; I will discuss the relative rankings only inasmuch as they convey implicatures through the use of the pronoun.

(14) Laurel$_1$ told me that *they*$_1$ got a job in San Francisco.

They in (14) is completely unspecified for gender; this means that it does not necessarily presuppose a nonbinary gender, but also does not necessarily presuppose any other gender features either. Presupposition can be tested by Infelicity if Avery challenges it: "*#Hey wait a minute! I didn't know Laurel was a they / is nonbinary* / whatever else." [8] Instead, Avery must decide what possible constraint rankings could have resulted in *they*.

[6] This section draws heavily on conversations I had with Edwin Howard, Leah Velleman, and Brooke Larson. My understanding of presuppositionality benefitted greatly from these conversations, and it was in the course of those conversations that I developed my understanding of the examples and inferences I discuss here.

[7] The format of constraint rankings and the tables below are intentionally reminiscent of Optimality Theory (Prince and Smolensky 1993 i.a.); however I am not assuming that any of these constraints are universal or actually necessarily part of the narrow language faculty; these considerations are socioculturally bound and certainly look very different in various social contexts. The tableaux are mainly an effective visual communicator for the audience of this dissertation.

[8] This very obviously varies between different speakers; however, in my experience when a speaker uses they in conversation with an interlocutor who is unaccustomed to its use, the interlocutor's first assumption is NOT that the referent is non-binary gender. That interpretation does exist, but it relies on inference rather than presupposition.

The use of *they* may be due to the various underlying rankings in (15), and implicatures that arise will be extrapolated from which use is most contextually likely.

(15) a. Relation (5) > Max Quant (3a) = *if it's not relevant, don't include gender.*

b. Don't lie (4) > Max Quant (3a) = *using a more specific pronoun would be a lie – prioritize truth over specificity* (meaning, Laurel is nonbinary or uses *they/them* pronouns for any other reason)

c. Don't guess (4) > Max Quant (3a) = *the speaker is not absolutely certain of Laurel's gender and wants to avoid guessing at the risk of guessing incorrectly.*

Constraint ranking (15a)	Maximize Quantity!	Relation
☞ *They*	* (not being as specific as you could be)	
She		*! (only include information if it is relevant)

Table 4.1: Constraint rankings for Gricean Maxims in conflict: don't include irrelevant gender

Constraint ranking (15b)	Don't lie!	Maximize Quantity!
☞ *They*		* (not being as specific as you could be)
She	*! (*she* is not correct/appropriate)	

Table 4.2: Constraint rankings for Gricean Maxims in conflict: referent is nonbinary

The inferences shown above all arise specifically when singular *they* is part of the speaker's grammatical repertoire. Importantly, however, the availability of singular *they* also introduces optionality into *all* utterances involving singular pronouns. If conversational maxims are a type of inference made based on a speaker's choice among several,

Constraint ranking ((15)c)	Don't guess!	Maximize Quantity!
☞ *They*		* (not being as specific as you could be)
She	*! (gender is not certain)	

Table 4.3: Constraint rankings for Gricean Maxims in conflict: avoid guessing Laurel's gender

propositionally-equivalent alternatives, and if singular *they* does not introduce any gender features at all, then *they* must ALWAYS potentially be an option available to speakers. Instances when speakers opt *not* to use *they* therefore necessarily rely on Gricean inference to construct the social meaning conveyed by *he* or *she*. So, opting for a more specific pronoun does still rely on relative rankings of the conversational maxims.

(16) Laurel$_1$ told me that *she*$_1$ got a job in San Francisco.

As with the test in (14) previously, it is possible to test for presuppositionality by challenging the assumed common ground implied in (16): "*Hey, wait a minute, I didn't know that* [*Laurel is a she / Laurel is female / Laurel is out at work / you knew that Laurel had transitioned* / whatever else]". Crucially, these challenges to the introduction of *she* into the assumed common ground are predicated on whether it is appropriate for a speaker to call Laurel *she* (NOT that Laurel "is" a woman in any objective or definitional sense—that is a secondary leap of inference dependent on gender ideology.) In relation to *they* as a potential alternative, however (why use *she* when *they* is an option?) the relative rankings of constraints can still imply information:

(17) a. Max Quantity > Don't guess/Relation/etc = *put as much information in as you can—she* is more informative than *they*

b. Relation > Min Quantity/Don't guess/etc: *include specific information as it is relevant to the context. Laurel's gender is related to her new job in some way*

Constraint (17a)	Maximize Quantity!	Relation
They	*! (less specific)	
☞ *She*		* (gender is not relevant)

Table 4.4: Constraint rankings for Gricean Maxims in conflict: be specific about who you're referring to

Constraint (17b)	Relation	Minimize Quantity!
They	*! (withholding relevant info)	
☞ *She*		* (not as concise as you could be)

Table 4.5: Constraint rankings for Gricean Maxims in conflict: include relevant gender information

The inferred interpretations I have shown in this context all rely on alternations between an ungendered (and thus less specific) singular pronoun and a specifically gendered one. The implicatures shown can all be exploited by speakers for various conversational goals.

In Chapter 3, I introduced data from an experiment that was aimed at prompting speakers to use pronouns to describe real referents. This experiment yielded many instances of dsT by various speakers for various purposes. Three participants in fact used dsT essentially as a universal pronoun—they never used any other third person singular pronoun to refer to anyone. Two of these participants, OOA and OOB, used dsT to refer to various absent third parties (mutual friends of theirs). While discussing mutual friends, OOA and OOB were attempting to describe their respective relationships to a particular friend; both speakers only ever used dsT, and OOB in particular noted that they didn't want to say a friend's name (in front of the interviewer or while being recorded):

> *OOA: Well, one [mutual friend], we, the one we mentioned earlier we know through our - well I know **them** through soccer, and that was how **they** got they got me*

introduced into the activist group, I don't know how you know ***them***

[. . .]

OOA: Who were your friends in the class?

OOB: I don't wanna say their names!

The use of dsT by these speakers could be motivated by various factors: it may be that these speakers (both of whom were nonbinary and preferred to be called *they*) had many friends who were nonbinary and preferred to be called *they*—if this was the case, then the use of dsT shows that the speakers both rank Quality (and *Don't misgender!*) higher than other potential factors. If, however, the use of dsT was instead an extension of the speakers' preference for anonymity, then they may have been prioritizing Relevance (*'Don't include irrelevant details'*) over accuracy for their friends' pronoun preferences.

What this attestation shows more generally is that the use of any given pronoun is only interpretable by a hearer insofar as the hearer has sufficient context (discursive and social) to infer how the speaker is likely to behave. The fact that, as an interviewer and a stranger, I can't discern whether OOA and OOB were intentionally obfuscating the gender of their friends or simply using their friends' preferred pronouns is an artifact of my relative lack of context; sufficiently rich contexts will not cause such ambiguity.

In the following sections I will go forward with the assumption that speakers always have a choice in what pronoun they will use for a particular referent, and that their choices can reflect particular conversational aims or constraints. I will first discuss issues of how speakers negotiate the social situation described around ((13))—particularly the case of a speaker not knowing a particular referent's gender or pronoun preferences. This case will rest on three different politeness constraints that I will propose: *Don't misgender!*, *Don't ungender!*, and *Don't correct me!*. I will then go on to explore more affective uses of pronouns to convey stance, both in an evaluative and gender-transgressive context; in these cases, the implicatures shown here will become more complicated. I will

finally go on to discuss situations in which different speakers (even in the same conversation) will have different relationships to a particular referent that are reflected in different pronoun uses.

4.4 Showing Politeness

One of the reasons that the discourse sensitivity of the gender features of pronouns has been under-explored has been the simple fact that, for most people most of the time, it is never really in question which pronoun to use about any given referent; many speakers are accustomed to knowing from past conversations the conventional pronoun in use around their acquaintances, and many speakers are equally accustomed to relying upon impressions of a new person's body, voice, clothes, number of piercings, romantic entanglements, etc., to give them a consistent set of clues that 'add up' to a conventional pronoun use for strangers. However, readers may also recall times when they are not sufficiently secure in their 'guess' (due to insufficient information, or conflicting information) as to which pronoun would be appropriate for a referent. This happens much more around people whose gender expression is non-conventional or ambiguous in some way, which can happen for various reasons—different body shapes, racialized bodies, or absent bodies (e.g. in internet communication) along with non-conventional clothing, romantic behavior, or other social cues can make speakers feel less secure in their 'guess,' which may only go so far as to make them aware that they are 'guessing' at all.

Of course, everyone is always guessing everyone else's' pronouns unless they ask—and norms around asking pronoun preferences are in flux at the time of this writing. In this section I will largely abstract away from the possibility of asking ones' referent their pronoun preferences. Instead, I will examine strategies that speakers use to establish their pronominal convention (and adjust it as an acquaintance progresses) when they're not sure. I will show two possible first-guess strategies—under-shooting and over-shooting—before moving on to strategies that speakers use to adjust their first guesses based on information from other interlocutors or other social cues.

4.4.1 How do you guess when other interlocuters aren't involved?

When in a situation where other interlocutors are either not present or not helpful, speakers must rely completely on their own ideas of gender and politeness in order to choose the first pronoun. In Part 1, I proposed three constraints of politeness that inform speakers on the correct social move in this instance[9]. I will repeat the first two here (the third being mostly left to the section on stance).

(18) *Don't ungender!* Failing to attribute a person's gender to them is an imposition on their positive face.

(19) *Don't misgender!* Asserting an incorrect gender for a person is an imposition on their positive face.

Given these two constraints, speakers generally have to decide between attempting to guess a binary pronoun (and if so, which one?) and avoiding binary pronouns (thus, dsT or total pronoun avoidance). In the pronoun production experiment, speakers demonstrated various possible behaviors that derive from different rankings of textitDon't ungender! and *Don't misgender!*.

In the data from Experiment One, some speakers opted to avoid binary guesses when possible. This was more often the case when speakers were referring to transgender referents than cisgender ones (though not universally; I discuss an exception below). I will here discuss two examples from the conversational data of this experiment—one referring to a fictional character, and one referring to a real (but absent) referent.

The pronoun production experiment included a task where participants answered open-ended questions about film clips that included both trans and cis characters. In the first clip, a transgender woman (Ricky) and a cisgender woman (Francesca) meet for

[9] While the wording in (18)-(19) makes particular reference to gender, these principles should be generalizable to other instances where social identity or relationships are encoded in language; a generic version might be *Don't remove identity!* or *Don't misidentify!* – this could be applied to considerations around honorific marking and pronouns, for example.

the first time. Ricky is not explicitly revealed to be trans, but she implies it in her conversation. Some participants reacted to either this implication or some aspect of Ricky's gender expression by avoiding using *she*, opting instead for *they*. One participant, RA, used *they* when discussing Ricky in the first clip, noting immediately that Ricky might be transgender:

> *[RA about Ricky in clip 1] I'm pretty sure that Ricky is trans, I'm also pretty sure that* ***they*** *are close friends with the man who was sitting in the cafe listening to this conversation but not participating. Um, also pretty sure that there was attraction and flirtation between both Ricky and the customer.*

This represents a conservative guess: when presented with possible evidence that a casual guess might not be accurate, RA opts instead to avoid binary pronouns. It is also worth noting that this was the only time RA used a pronoun to refer to Ricky at all; pronoun avoidance is also a strategy that aligns with the goal of not attributing the wrong gender (at the cost of attributing no gender at all). RA can be said to prioritize *Don't misgender!* over *Don't ungender!* very regularly.

Another participant, VA, apparently ranked these constraints the other way—she used *she* to refer to Ricky right away:

> *[VA About Ricky in clip 1]*
>
> *Um. I felt like she was like too direct.*
>
> *Not caring about the other person's feeling. Like not minding, let's see, maybe trying to make the other person feel uncomfortable.*

What makes the ranking more apparent was the fact that VA also opted to prioritize *Don't ungender!* over *Don't misgender!* when talking about her interview partner, VB. VB was a transgender nonbinary person who prefers *they* (and will accept *he*), but VA exclusively referred to VB as *she*. Thus while in the example above (about Ricky) VA

guessed correctly, she guessed incorrectly later in the interview—the fact that she opted to guess at all suggests that *Don't ungender!* is an important constraint for her. It is also interesting to note that VA only used two instances of *she* to refer to VB over the course of the interview—this may have been a partial avoidance strategy, even though it was overridden by the necessity of guessing one of the binary options.

Not all stranger pairs in the experiment opted to guess a binary pronoun at the cost of potentially misgendering their referent—in fact, one participant opted to avoid guessing in a way that ended up misgendering her referent nonetheless. The guess avoidance in this case again appears to be related to the speaker (GGB)'s impression of the referent (GGA)'s transgender status. In this pair, the speaker GGB exclusively used *they* to refer to GGA when being interviewed separately (and used no pronouns when being interviewed together); this was not a reflection of GGB's actual identity (he uses *he* exclusively), but was apparently related to GGB's assessment of GGA as *some* type of transgender, based on the pair's discussion in the interview together. Thus, *they* may have been GGB's 'safest guess' from her perspective.

In the pair interview, GGA brought up their transgender identity in a discussion of political events that were ongoing at the time, noting a discomfort with the association between anatomy and gender:

> *GGA: [. . .] I also, all the like, the, the hats made me kind of uncomfortable to be honest, and just the focus on like, genitalia, as a theme in it, is-*
>
> *GGB: What- what hats?*
>
> *GGA: Oh, they were wearing like, like pussy hats, 'cause they're, um*
>
> *GGB: Oh really?*
>
> *[...]*
>
> *GGA: I guess as a trans person it made me uncomfortable*

GGB: Yeah. That seems Kind of- I don't know. Maybe not thoughtlessly insensitive, maybe it was thoughtfully insensitive, but it's kind of, it's quite narrow-sighted, I think it's- oh, okay, I didn't know about that That's not so good

GGB seemed to accept GGA's transgender identity in a generalized way, but never found out what *specific* gender GGA identified with. Thus, *they* served as an apparent marker of transgender identity that is agnostic about specific gender. GGB made no metapragmatic comments about pronouns specifically, but used *they* consistently in the solo interview. She also expressed "wariness," apparently related to their understanding of the nature of the study as a whole being related to gender and transgender identities.

K: And then just last question. How do you think the first interview went? How were you feeling? How do you think [GGA] was feeling?

GGB: Um, I imagine, we were both thoroughly apprehensive about what it would involve. Um, ***I'd imagine particularly because of, like some gender issues and some trans issues can be sensitive****. I imagine we were both kind of, a bit.. I don't know. not wary. But yeah, apprehensive I think? [. . .]*

In this instance, the connection to transgender issues appeared to have convinced GGB that she should not try to 'guess' a binary pronoun for referring to GGA (and she didn't try to ask him when they were together); so *they* is apparently a conservative guess by being both underinformative (as *they* can be used to refer to people of binary genders—future sections will have more data on this) and, through its underinformativeness, related to transgender identity. GGB's behavior in this interview constitutes one possible reaction to being forced to 'guess' a pronoun when encountering a new person whose gender expression is unclear and whose gender identity is unknown.

These attestations suggest not only that *Don't ungender!* and *Don't misgender!* are both active (but differently ranked) constraints that inform speakers as they navigate unfamiliar pronoun situations, but also that the behavior around these rankings is most notable

when a referent is transgender or gender-nonconforming in some way—that is, it becomes more obvious that speakers are always necessarily 'guessing' at pronouns when put in situations where they are more likely to guess wrong, or to feel uncomfortable about their guess. I will return to the significance of this trend in the section on stance.

4.4.2 "Narrowing down" pronoun choice

When speakers opt to prioritize *Don't misgender!* over *Don't ungender!* they may use an underspecified gender pronoun or no pronouns at all when referring to a referent; however, as speakers learn more information about referents, they can incorporate the new information into their calculations of what pronoun is most appropriate. In these instances, speakers will start with an underspecified pronoun (dsT) and move towards a more richly specified pronoun (*he* or *she*) as they collect information sufficient to override the possibility of violating *Don't misgender!*.

In the conversational data from Experiment One there were several instances of speakers starting with *they* and moving to *she*, particularly when discussing the fictional transgender woman Ricky. This was a consequence of accumulation of context: in the first film clip, Ricky implies that she is trans but doesn't outright state it, while in the second film clip, Ricky explicitly comes out as a transgender woman. Participants in the experiment whose first 'guess' regarding Ricky was *they* often switched to using *she* after she came out as a trans woman—presumably because the evidence in favor of Ricky identifying as a woman was at that point sufficiently rich to merit a more specific guess.

RB is one of the participants who began the film clips interview calling Ricky *they*, but updated to calling her *she* immediately after watching the second film clip:

> *[RB about Ricky in clip 1]*
>
> *Uh, so I think **they** were probably a little more abrasive than I would ac, than, um. **They** were a little abrasive.*
>
> *[RB about Ricky in clip 2]*

the fact that Ricky switched to texting the details, when, um, ***she*** *then, previ- or, then subsequently was saying that* ***she's*** *very comfortable*

The pattern of starting with a more vague gender and updating towards a more particular gender is evidence that speakers were prioritizing *Don't misgender!* until such time as they felt comfortable enough to 'risk' a guess.

4.4.3 "Updating information" in pronoun choice

Prioritizing *Don't ungender!* over *Don't misgender!*, however, means starting off with a binary gendered guess, but potentially switching pronouns later in the conversation when the context provides sufficient evidence that the first guess might not be correct (or appropriate, or accurate within a speaker's understanding of binary gender). In this case, we would predict that a speaker would show the pattern from above in reverse: start with *she* and switch to *they*. This was attested in the study once. IA, a participant who did not guess at first that Ricky was transgender, updated his pronouns after watching the second clip—he had previously been referring to Ricky only as *she*.

K: How do you feel about Ricky?

IA: I really like ***her****.* ***She's*** *very authentic, and uh, sincere, and self aware and unassuming about who* ***she*** *is, or who* ***he*** *is, or haha,* ***they*** *are. So yeah. I like* ***her*** *authenticity and integrity.*

After this brief foray through the possible gendered pronouns of English, IA returned to calling Ricky *she* for the rest of the discussion about the film clips, and never explicitly suggested that she was not truly a woman. This initial 'shock' alternation is interesting because it is very overtly metapragmatic—the repetition to include different pronouns highlights the pronouns themselves as markers of possible gender. The use of *he* and *they* as possible (but ultimately rejected) options for reference to Ricky suggests that IA was, at least momentarily, debating the definition of Ricky's womanhood. The rotation

suggests that IA needed to at least entertain the possibility that *she* was not the most appropriate choice, meaning that the constraint D*on't misgender!* threatened to overrule *Don't ungender!*.

4.4.4 Asking for help

In some cases, speakers will not feel secure enough in their commitment to either *Don't ungender!* or *Don't misgender!*, and will reach an impasse. This can result from an internal politeness system where these constraints are not ranked higher or lower with respect to one another, but are equally important to the speaker. If a speaker reaches an impasse, they are faced with either total avoidance of the issue (which by default prioritizes *Don't misgender!*) or they are forced to ask explicitly. Only one participant asked about pronouns explicitly in the experiment: BA, who was paired with a stranger, did not feel comfortable about using a pronoun to refer to her partner (who was out of the room)—so she asked the interviewer (me):

*BA: And **she** goes by "**she**", right? I'm not incorrect on that?*

K: I have no comment. I did not ask

This exchange is interesting because during the joint interview, BA and BB had been sharing experiences of what it was like to be women in a male-dominated industry—which apparently was not sufficient evidence to convince BA to commit to using *she* without reassurance. (Despite the fact that I did not give BA an answer, she did opt to use *she* for the rest of the interview.)

This section has focused on the interaction between two constraints on politeness around gender when they are apparently in conflict. I have shown that speakers will use various strategies for navigating a social situation in which they need to 'guess' a pronoun for a particular referent when there is no linguistic antecedent to depend upon. Speakers who prioritize the constraint *Don't misgender!* will opt to avoid gendered pronouns altogether, while speakers who prioritize *Don't ungender!* tend to hazard a guess. Both types

of speakers show adaptive strategies for altering pronoun use when new information is introduced into the context, which results in instances where a speaker will switch what pronoun they use for a particular referent over the course of a single conversation.

4.5 Stance and Gender

The previous sections have shown how manipulating certain constraints, such as politeness constraints and conversational maxims, can reliably predict variable speaker behavior with respect to gendered pronouns. This section will turn to stance, paying attention to particular ways that speakers use gendered pronouns to position themselves relative to their referent and to the relationship between their referent and their addressee. The first subsections will be more focused on positioning and evaluation between a stance subject and a stance object, while the later subsections will move towards showing how that evaluation can be brought into (or against) alignment with an addressee.

4.5.1 Personal relationships: affective and evaluative stance

When referring to a third person referent with whom some or all of the audience are familiar, pronoun use can be sufficient to constitute a stance-taking act, where the referent of the pronoun is the object of stance. These alternations are rarer in mainstream English contexts, but much more common in social situations where some or all of the participants are transgender, nonbinary, or otherwise exist at the margins of the gender landscape. One example, anecdotally reported by many linguists,[10] is a 'distal *they*.' This is the use of dsT to signal social distance—which is a phenomenon dependent on second-order abstractions of inferred meaning that are constructed through manipulations of (for example) conversational maxims. If all participants in a conversation are in accord with both the availability of dsT and with the understanding that dsT can be used to signal irrelevance, then using dsT about a referent whose binary gender is known (and whose

[10] Thanks again to Leah Velleman for informative discussions about this.

identity is known) can also be a signal of social/relational irrelevance. That is, using dsT may be a signal that a person is unimportant or uninteresting based on the implicature that if their gender were relevant, the speaker would have used a gendered pronoun.

One notable example of gender-switching in one of these cases took place in a pair interview with RRA and RRB, acquaintances who both identified themselves as transgender. In discussing a mutual acquaintance, they used both *he* and *they* within the space of a few utterances to refer to the same person:

> *RRA: His partner at the time was also dating this other person that was in our group. Um, and **they** have a very, um, **he's** a very strong and kind of controlling personality, and so **he** had kind of taken over like the whole thing, and [...]*
>
> *RRB: **He** started blaming different people to different people, and just like the whole thing fell apart. So. But that's okay.*
>
> *Moved on to better things.*
>
> *RRA: Ha. Yes. it kind of, that was kind of one of those things where it just- and that same person, I would see **them** more often than I would see [RRB] and **they** were trying to like convince me of these like negative things [...]*

In this exchange, the referent is a stance-object for several stance acts (underlined above); RRA introduces the referent and evaluates him ("*strong and controlling*"), RRB also establishes an evaluation but explicitly positions themself with regards to the situation ("*But that's okay"*), and RRA concludes the topic with an additional evaluation ("*trying to convince me of these like negative things"*).

The use of *they* and *he* alternate back and forth in this exchange: a simple 'updating of information' (in either direction) is not explanatory here. Instead, I note that both speakers use *he* when reporting particulars of the acquaintance's bad behaviors, and *they* when discussing more vague aspects or times when the acquaintance was less socially involved with them.

This analysis assumes, of course, that neither RRA nor RRB are using pronouns as a way of intentionally misgendering the referent. It's also a very subtle and potentially

complicated reading of the exchange, but more overt instances of misgendering make the relationship between stance and pronouns more obvious.

4.5.2 Malicious misgendering

Aside from conveying information about a personal relationship, pronouns can also be deployed to express a stance on the referent's gender identity—a referent's validity, authenticity, attractiveness, politics, and various other personal traits may be wrapped up together with their gender identity or expression. In Section 4.2.3, I gave an example of using pronouns to intentionally derail sexism or avoid homophobia; but pronouns can also be used to accomplish the opposite. Pronouns can be used to imply (a stigmatized, socially-undesirable) gender-nonconformity (in the case of using the opposite binary pronoun as the referent's gender identity), or to emphasize the implication that a referent's gender is related to their actions (e.g. a speaker repeatedly using or emphasizing *she* when discussing a person's bad behavior, potentially to imply that women in general are badly behaved. Speakers are aware that using 'crossed' binary pronouns constitutes a serious insult, although in cisgender circles it is assumed to be somewhat a puerile one. (20) gives commentary from an internet discussion board[11] amidst a heated discussion of politics, in which one participant used gender-crossing as an insult to a (non-involved) referent:

(20) Never heard anyone in middle or grade school call someone a "useful idiot." I did, however, hear kids who would invert other kid's gender, calling a guy "she" if he had long hair, or a girl "he" for being a tom boy. Never heard it since, even in high school. Until social media. And especially in the comment section under every youtube video featuring Maddow. A blast from the past.

The commentary in (20) reflects the position that gender-crossed pronouns (henceforth 'misgendering,' a deliberately vague term throughout this chapter) constitute a 'cheap shot' when attempting actual verbal sparring or debate. This may however be

[11] https://forums.anandtech.com/threads/rachel-maddow-just-became-the-most-watched-show-on-all-of-television-poor-hannity.2554160/page-2

related to the fact that the referent who was misgendered, television pundit Rachel Maddow, is lesbian and gender-nonconforming but (crucially) cisgender.

Rachel Maddow's cisgenderism protects her from what is seen as much more 'acceptable' misgendering-as-insult. In cases where referents are transgender, there is a much greater divide in the discussion on this forum and similar ones (or, e.g., Twitter) about whether or not misgendering is an appropriate or fair insult; it is also extremely common in the lives of transgender people to be misgendered intentionally as a form of verbal censure. This is a more extreme way for speakers to use pronouns to convey stance about a particular referent. It is also relevant that misgendering (via pronouns or otherwise) has a significant negative effect on the mental health and well-being of transgender people (McLemore 2013, 2015); pronominal misgendering aimed at trans people therefore constitutes a form of serious (and frequent) verbal harm.

Intentional misgendering can convey affective information about the referent that would be unusual to find in a completely neutral, grammatical (and presuppositional) element of language; using the wrong pronouns can show a speaker's prejudice against transgender people in general, disapproval of their particular gender presentation, or a kind of disindentification with gender altogether through 'degendering.' I'll first show examples of malicious misgendering, then degendering.

In Conrod (2017) I collected tweets from Twitter to create a small corpus of text where speakers referred to Chelsea Manning by either her preferred or dispreferred first name, and by either *he* or *she*. In these tweets I found not only that those misgendering Manning (using *he* with either name) had significantly more negative attitudes about her, but also that those tweets were proportionally much more related to her transgender status. Figure 4.3 below shows the mean proportion of positive/neutral/negative tweets for each combination of name and pronoun (*she* + *Bradley* was unattested), and Figure 4.4 shows the proportion for each category that was related to her transgender status.

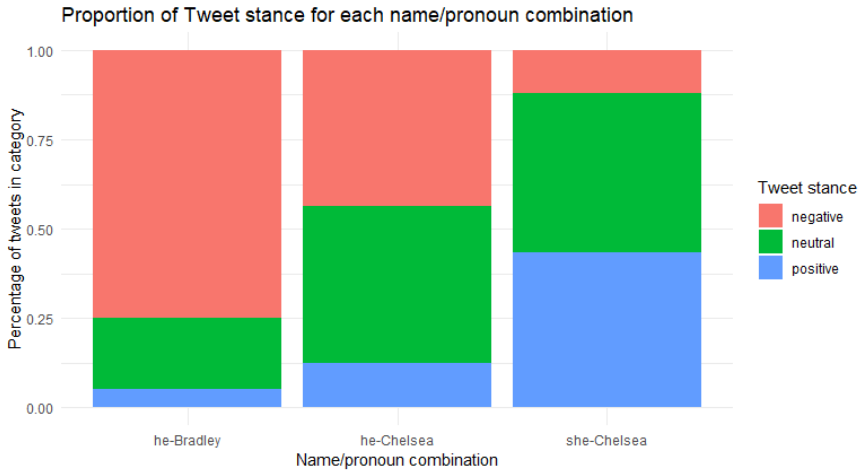


Figure 4.3: Tweets that misgendered Manning were more negative

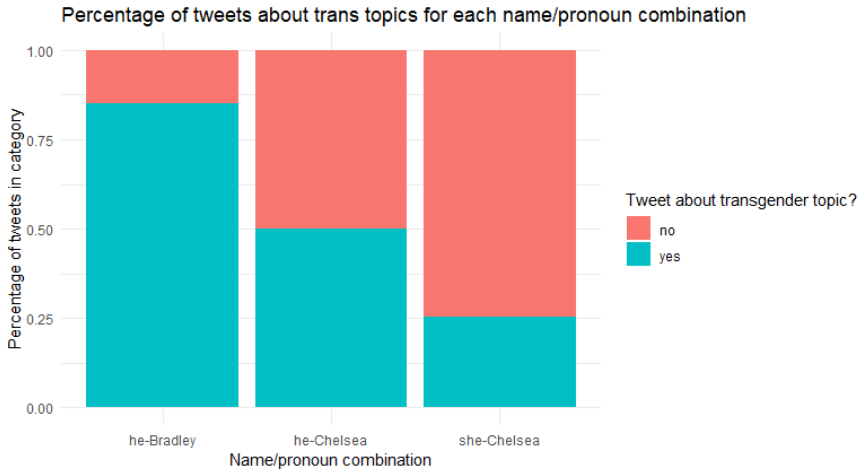


Figure 4.4: Tweets that misgendered Manning were related to transgender status

The relationship between negative evaluations of Manning and use of pronouns misgendering her is indicative that the pronoun use is either caused by negative stance, or caused by the same factors that make speakers evaluate Manning negatively. Likewise, the high relative proportion of tweets that were related to Manning's transgender status among tweets that misgendered her suggests that those who misgender a trans person are

also more likely to do so while/in order to draw attention to that person's trans-ness.

Tweets where speakers overtly commented on Manning's pronouns also make overt statements of evaluation. The tweet in (21) uses a depronominalization to avoid using *her* referentially, and to contest the use of *her* in reference to Manning in general. (22) uses two referential pronouns but the speaker still positions themself as an expert on Manning's 'true' gender (which is an epistemic stance overtly, but covertly evaluative).

(21) *Bradley Manning is NOT **a "her"**. Don't believe me? Check **his** DNA. It's right there in "X & Y"*

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

The tweet in (21) also shows alignment with regard to a previously-established (inferred) stance-taking act—that tweet was in response to another tweet in which a speaker used *her* to refer to Manning. The response in (21) is an explicit contestation of that use, suggesting that the speaker does in fact interpret referential pronouns to be stance-taking. This reply is an explicit example of alignment (between stance subjects, for DuBois 2007) where the speaker aligns themself epistemically in opposition to their interlocutor's positioning with regard to the (mutual) stance object.

It is worth noting that malicious misgendering is more frequent in utterances where the speaker is attempting to position themselves in opposition with the stance object—and in fact misgendering through pronouns is one resource among many that speakers may exploit in order to do so.

4.5.3 Inflexibility or inertia

Another, less overtly antagonistic type of misgendering can be attributed to a combination of stance-taking, politeness constraints, and what I'll refer to as 'inertia.' Many transgender English speakers will report that, despite polite and repeated requests for friends and family members to use certain pronouns, friends and family members will continue to 'accidentally' misgender their loved one. Friends and family members will

report causes for this such as 'it's hard,' a fascinatingly vague explanation. It could be that it is hard to reconceptualize a particular referent as being of a gender category you're not familiar with at all (in the case of nonbinary people) or of a gender category you haven't previously attributed to them (in the case of binary transgender people). Some have suggested that being asked to change their pronoun behaviors with regard to a particular referent is the equivalent of being asked to fundamentally restructure one's own unconscious grammar,[12] though this does not explain why switching from *he* to *she* should ever be a problem if both are otherwise grammatical.

Instead, I suggest that repeated misgendering from 'friendly' acquaintances (family, partners, etc.) is an artefact of two things: first, a resistance to recategorization of social categories in general (inertia or force of habit); and second, a higher ranking of the politeness constraint *Don't correct me!*, which constitutes an imposition on an interlocutor's free will (negative face). Since the second factor (preservation of negative face) is included in this model as a construct of politeness rather than stance, I will in fact argue that stance is expressed through the first factor, instead. Resistance towards recategorization is expressed through a conflict in the epistemology of a referent's gender: when a speaker repeatedly 'mistakenly' misgenders a particular referent, it is because the speaker (probably unconsciously) believes that their evaluation of the referent's gender is more accurate than the referent's stated preferences; if all speakers were at their core truly and completely dependent on the self-identification of their referents then misgendering would otherwise never occur[13]. The way this manifests is in speakers repeatedly ignoring cor-

[12] Pullum, G. "Suppose someone said they wanted any object pronoun referring to them to be positioned before the verb, as in French, rather than after the verb, as in English. Could you manage that? Could you them accommodate by making the requested change to the positioning of pronouns that them denote?" http://languagelog.ldc.upenn.edu/nll/?p=35641

[13] Inertia or force of habit are an under-explored and difficult-to-measure factor, but may well continue to contribute to repeated misgenderings as well. However, anecdotally I can report that speakers whose relationship with a transgender person only began after the person had transitioned do still exhibit the same kind of consistent misgendering behavior, and it is in fact almost indistinguishable from those who have known the referent since before their transition. In order to control for this possible confound future research will look more carefully into general cognitive functions like recategorization, in gender and elsewhere.

rections or other socially-oriented complaints, and effectively functions as a way for a speaker to show an opposed alignment with the third party they're referring to. This is not exactly how DuBois sets up the triangle, but I return to this issue after discussing an example.

One participant in the production experiment, GA, consistently misgendered his partner despite corrections, and spoke about the difficulty of meeting their partner's preferences. GA was interviewed with his romantic partner, with whom he cohabitated. During the pair interview, GA repeatedly used *she* for GB, despite GB's preference for being called *they*. GB corrected GA on this matter during the pair interview:

> *GA: [...]* ***she's*** *not able to, so I'm just suggesting paths that are more efficient. It's-*
>
> *GB: And also* ***they them****. He always forgets they them, too.*
>
> *GA: Right,* ***they****. Um. The pronouns are something I'm trying to get used to. [...]*
>
> *GB: But then the pronouns one was like I dunno, maybe a little before that, but I didn't really mind it until*
>
> *GA: Um.* ***She*** *didn't really correct me until about summer, I wanna say in the summer.*
>
> *GB: Probably.*
>
> *GA: Yeah. Really. Before,* ***she*** *didn't mind the* ***she*** *does, sometimes* ***she*** *doesn't, I*
>
> *GB: You're still saying* ***she****!*

During the solo interview, GA did not attempt to use *they* for GB at all, and specifically asked the interviewer (me) whether it was okay if he didn't try to match GB's preferences:

> *K: Um, great! So. I would like to ask you: How do you describe [GB] when you're talking to your other friends? Like, personality or looks or stories. How would you explain [GB] to a person you were just meeting?*

GA: ***Could I not use her official name*** *or something in this interview?*

K: Um. yeah, 'm going to. I am not gonna tell you what to do.

GA: Does it have to be kept... Um. I... I could go on and on for a while.

K: Yeah, no. Like as much as you want to say

[. . .]

GA: [. . .] it's more just about getting accustomed. What could I describe about ***her****?* ***Again, with the pronouns.*** *It's a habit to say that. She's really hard- I wanna avoid that, but*

K: Whatever you need to do

GA: Could I...? Okay. A hard-working person. [...] Once you get to know ***her****. I want to be able speak fluently without stuttering, so*

K: Again, I'm not gonna- I'm not gonna stop you from anything. So, if whatever you need to do to speak freely, you can do that

GA: Okay. Once you get to know ***her****, then* ***she's*** *[...]*

In this excerpt, GA repeatedly seeks validation and permission from me to "speak fluently without stuttering," which for him means using *she* and GB's dead name[14]. The use of GB's dead name suggests that grammaticality of singular, specific *they* is not the only contributing factor; GB's ratings of transgender people in both implicit and explicit measures of attitudes were notably less positive than other participants'.

Negative attitudes towards transgender identities appear to contribute to misgendering both in conscious and unconscious ways. In the case of Twitter users misgendering Chelsea Manning, misgendering is one among many linguistic 'weapons' they can wield

[14] A "dead name" is a term used mostly by transgender people for a name that is no longer consistent with one's gender identity, and is therefore dispreferred. I use this term to 1. Align myself with the transgender community, and 2. Emphasize the discomfort transgender people feel upon being called their dead name.

against her in order to sanction her political actions. In the case of misgendering an intimate partner, however, GA's reluctance appears to be at least as connected with his idea of the 'old version' of GB, the person he knew when the pair began their romantic relationship. His reversion to their former name reinforces the pattern—his resistance is as much to change or transition in a personal realm as it is to any possible political stance. In repeatedly using GB's old name and pronouns, GA is taking an implicit stance on how he thinks of GB as a person, and how he wants to describe them to others when he's unsupervised.

In the case of misgendering a referent with whom a speaker has a personal relationship, the referent also necessarily acts as a subject whose subjectivity the speaker must be aware of. DuBois's (2007) description of stance objects is largely restricted to non-persons, but to account for the use of third person pronouns there must necessarily be some blurring of the distinctions between interlocutor and stance object. This manifests in the example above, where GA must actively account to GB's stance on the matter of their own gender—GB's corrections constitute a stance-taking act that GA must then later align in opposition towards. In the solo interview, GA attempts to secure approval (and positive alignment) with his interlocutor (me) by explicitly asking about it. This suggests both that GA is taking into account his own epistemic and evaluative stance towards GB's gender, but also how his stance aligns with GB's and my own.

The examples discussed in this section on stance thus far have focused primarily on negative attitudes and stance towards referents; however, pronominal gender can also be a site of positive stance expression that enacts gender affirmation and celebration. The next section will examine how stance influences pronoun use within LGBTQ+ communities.

4.5.4 Gender play

In some cases, pronouns can be used for affective or relational purposes that build on gender as a social structure. This is one of the more overt instances of the performativity

of gender in language, where language and speech acts are part of the day to day behavior that constitute gender expression and relations. One of the more well-known conventions of this type is "the gay *she*," a convention where cisgender identified gay men will use *she* within gay mens' spaces as a way of expressing gender-nonconformity and gay identification. As gender is a very broad field with a lot of room for variation and sub-groups, small communities or social groups built around sub-genders can and do develop their own in-group vocabulary and indexical field through which gender can be played out, negotiated, and expressed.

In these sub-fields (very frequently in various types of LGBT+ social contexts), gender is less tied to hegemonic macro-identities, and gender presentation and expression can be much more dynamic or fluid. When such a sub-field is predicated on a particular kind of relationship with gender, then gender presentation or expression can itself be a prominent way of evaluating other group members. Contemporary popular depictions of drag culture call on this fluidity overtly; on reality television show *Ru Paul's Drag Race*, contestants are explicitly evaluated on their performance of (drag) gender—a performance which is predicated on stage skills, craftsmanship, style sensibility, and use of familiar tropes specific to the art. This explicit evaluation takes place both through judged contests, but also through video testimonials (from judges and other contestants) regarding the various competitors. These evaluations give valuable attestations of gendered language use, including pronoun use, being used as a linguistic resource for discussing group-internal gendered performance.

In one late-season episode of *Ru Paul's Drag Race*, one contestant, Jaymes Mansfield, was falling behind. The way that judges and fellow contestants referred to Jaymes in video testimonials was tightly linked to the adequacy of Jaymes' performance. In ((23)), a judge uses *him* and *he* while reporting that Jaymes is not really 'selling' the performance—a performance explicitly predicated on successfully presenting or invoking tropes of femininity for the stage. But in ((24)), another judge expresses a more sympathetic stance—and,

correspondingly, uses *she* to refer to Jaymes.[15]

(23) We are actually rooting for Jaymes and want **him** to shine, but he's gonna have to believe in **himself** to really sell this challenge

(24) Jaymes' audition tape was so funny, I got it. I understood the shtick. But I think that since **she's** been in this competition with the other girls, **she's** thrown off.

Pronouns are also used in the metacommentary around the drag scene as a way of differentiating between a drag performer (when the performer is a man) and the drag persona; This is reflected in the fan-curated *Ru Paul's Drag Race Wiki*, a mini-encyclopedia of the show's characters, performers, and events. The fan wiki routinely uses *she* to refer to the drag queen in character and on the television show, and *he* to refer to the performer in non-drag life.[16] (25) below shows the contrast, where both pronouns are ostensibly referring to the same person, Charlie Hides.[17]

(25) a. **She** was the third queen to be eliminated and ended up finishing in 12^{th} place. At age 53, **she** is the oldest queen to compete on the show.

b. Hides is also well known for **his** YouTube channel, which features hundreds of parody videos, and sketches where **he** impersonates celebrities.

The alternation between *she* (to refer to a drag queen's stage persona) and *he* (to refer to a performer off-stage) is one of the ways in which drag subculture splits what would in mainstream society be considered a single gender category into multiple sub-categories; using pronouns to express affect throughout a conversation is another way. LGBT+ social

---

[15] Transcript source: https://www.springfieldspringfield.co.uk/view_episode_scripts.php?tv-show=rupauls-drag-race-2009&episode=s09e02

[16] Wikipedia, the mainstream (crowdsourced) online encyclopedia, exclusively uses he to refer to Charlie Hides (https://en.wikipedia.org/wiki/Charlie_Hides) , as well as RuPaul (the show's host). RuPaul has expressed explicitly that he does not have a preference for particular pronouns; other former contestants on the show, such as Sasha Velour, have expressed explicitly that she is the only appropriate way to refer to their drag persona, and Wikipedia uses she uniformly to refer to Velour (Velour's tweet: https://twitter.com/sasha_velour/status/830305619886366720).

[17] Fan-curated wiki: https://rupaulsdragrace.fandom.com/wiki/Charlie_Hides

spaces, especially those where dominant gender ideology is challenged, are spaces where pronouns are effectively available for play, for expression, and for emotion. These uses are common, and the social rules governing them are implicitly understood by members of these spaces, even if they are not apparent to outsiders.

4.5.5 Closet pronouns and the stance triangle

One further complication of how gender and pronouns intersect in English is due to the fact that, because only third person pronouns are gendered, third person reference adds another dimension to the calculations speakers make about expressive gender and stance. While the previous examples (misgendering, gender play) rely upon the relationship between the speaker and the referent, there are many every-day situations where pronoun choice depends instead on the relationship between the addressee and the referent (absent or present). The most overt example of this is what I will call *closet pronouns*.

For transgender people, transgender identity can long predate any kind of social or biomedical transition; likewise, transition is a long process that happens gradually, at an irregular rate, and sometimes outside of our control. People whose gender identities are not legible in mainstream settings (including but not limited to nonbinary people) will sometimes choose the path of least resistance when it comes to pronouns—rather than attempting to correct acquaintances or colleagues, a person may instead submit to being referred to by whatever pronoun causes the least trouble. To give a personal example, I was traveling with a fellow nonbinary friend of mine across the border between the United States and Canada. My friend was not feeling well, so I did most of the talking. The border guard asked me what my friend does for a living, and I replied, "They're in art school." The border guard was confused, and asked, "Who's in art school? Is it just you two traveling?" At this point, I had the option of either invoking further scrutiny and possibly harassment upon myself and my sick friend by trying to explain singular *they*, or by simply picking whatever gender would make the border guard let us through the fastest. I opted for the latter, and simply repeated my answer using *he* instead. This calculation

was not made based on what I thought would make my friend the least uncomfortable (politeness) or based on my personal opinion of my friend's gender expression (stance), but instead based upon my attempt to guess what the border guard would think of my friend's gender opinion.

This third person calculation is why DuBois' stance triangle is needed in order to analyze third person pronoun use effectively: I needed to take into account the addressee's (border guard's) stance or relationship with regard to the referent (my sick friend). My friend has confirmed that they don't feel like this constitutes misgendering on my part—if anything, it's more that I was incidentally forced to play along with someone else's misgendering, even though the border guard never said a pronoun himself. Thus, the category of *closet pronoun* is a distinct term for a use where the speaker needs to incorporate their understanding of the addressee's understanding of the referent's gender, rather than directly referring to the gender that the speaker may know better. Figure 4.5 below shows a schematic of how third person gendered pronouns depend upon the triangulation between all three persons.

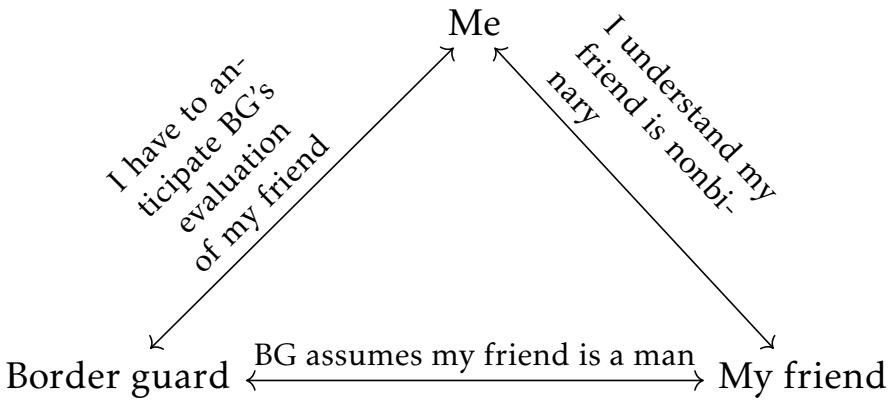


Figure 4.5: Different evaluation of gender = conflict

The relationships in Figure 4.5 can also be represented by variably-ranked constraints. In this case, the constraints are not general principles but are specific ways that each of the

participants in the discourse context relate to one another. The constraints as I give them in ((26)) apply specifically to me as the speaker (although, of course, (26a) is a specific example of the more general *Don't misgender!* constraint).

(26) a. *Don't misgender my friend*

b. *Don't surprise/confuse the border guard*

Relational constraints (26)	Don't misgender friend	Don't confuse guard
They		*! (guard doesn't understand dsT)
☞ *He*	* (friend doesn't go by *he*)	

Table 4.6: Constraint rankings for gender

A much more overt instance of closet pronoun use is when a transgender person is in the awkward period of social gender transition where they have come out to some people in their life, but not everyone. In this case, using the right pronouns with the right person may become a matter of immediate personal safety; when a transgender person is not out to their parents, for example, this is frequently because there is a danger that they will be disowned if their parents discover that they're trans. As another personal example, I am friends with a woman who is out to her close friends, but not to her colleagues or biological family. As a way of maintaining those boundaries (and her safety) it is therefore imperative that any of her friends who speak to her family do so using the pronouns that the family are accustomed to using. This means that there are social events with mixed groups (of friends plus family) where friends must use *he* to refer to this woman when they otherwise would only ever use *she*. These friends, when in safe company, not only use *she* exclusively, but are careful to correct each other if someone ever slips up. Using *she* is of great emotional importance to this woman, so her friends are exceedingly careful about it. However, the woman's personal safety is understood to be more important than her emotional well-being, and so closet pronouns are used when necessary. This is, again,

a (more extreme) example of the three-way calculation necessary for third person reference: the friends (speakers) base their pronoun use on their knowledge of the addressee's (family members) understanding of the referent's (woman's) gender.

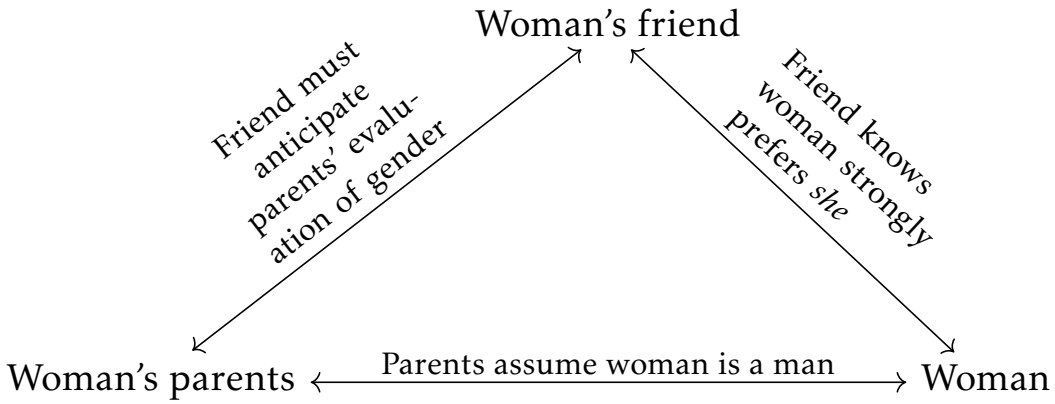


Figure 4.6: Negotiating pronouns when a friend is not out to her parents

As above, Figure 4.6 can also be described with conflicting constraints which are, again, specific instantiations of the social relationships and context.

(27) a. *Don't misgender my friend*

b. *Don't expose my friend to transphobic retaliation (= don't 'out' friend)*

Relational constraints (27)	Don't misgender friend	Don't "out" friend
She		*! (parents may disown friend!)
☞ *He*	* (friend prefers *she*)	

Table 4.7: Constraint rankings for gender with friends vs family

This triangulation is further evidence that complex social relationships are at play when speakers choose what gendered third person pronoun to use in English. For many, the social information conveyed by pronoun choice is part of the background, and warrants less attention than (overt) propositional content of an utterance; however, I have

shown that in many instances pronouns are the site of information conveyance with significant social weight and pragmatic import, and moreover that pronouns are flexible and contestable in a way that is not well-captured by a strict semantic presuppositional analysis. In the final section of this chapter I will set up the relationship between the pragmatic data that I have shown above and the syntactic analysis that will follow.

4.6 Conclusion

At the outset of this chapter, I cited Heim's conception of *Maximize Presupposition!* as a constraint that should, in its strictest reading, rule out singular *they* except as a last resort pronoun (when any other pronoun would be infelicitous). However, as I have shown in the previous sections, pronouns carry implicatures that can be extrapolated from aspects of stance (and affect), as well as Quality, Quantity, and Relevance, all of which enable speakers to use *they* and other pronouns with relative flexibility. Is it the case that the language change underway towards dsT (shown in Chapter 3 and further theorized in Chapter 5) is a movement away from *Maximize Presupposition!* as a ruling tenet for semantic felicity?

I would suggest that the data I have shown in these chapters instead suggests that *Maximize Presupposition!* was weaker than previously supposed all along—and that apparent changes around the norms of gendered pronouns in English are a move away from strict semantics towards an already-existing underlying sociopragmatic sensitivity. How, for example, could *Maximize Presupposition!* have ever applied to alternations such as T/V pronouns? R. Brown et al. (1960) give examples of mid-conversation alternations between *you* and *thou* that closely mirror the alternations in gender that I have demonstrated in this chapter. Neither *you* nor *thou* are more specific independent of context; nor can any implicatures be drawn from their use without context. What I suggest here is that *Maximize Presupposition!* is inherently an overly-strict (and in some ways over-simplified) formulation when it is divorced from context, even when apparent "true" presuppositions such as referent gender are at play. Rather than throwing out *Maximize Presupposition!*,

the best way to treat this social sensitivity is to treat it as one among many constraints that are in direct competition, which can be ranked according to a speaker's needs and desires in a particular conversational landscape.

In this and the two previous chapters, I have shown evidence that gendered pronouns are socially sensitive and cannot be analyzed as simply determiners with formal gender features. In order to set the stage for the final chapter, I must provide some scenery and lighting to hold the whole production together. Chapter 2 primarily (but not exclusively) focused on syntactic phenomena, while Chapters 3 and 4 have used sociolinguistic and pragmatic data that cannot be completely explained through formal features. It is an open question, then, whether the socio/pragmatic issues are relevant to a syntactic analysis. I argue that it is in fact necessary for a syntactic account of language to accommodate grammatical reflexes of social relationships; because of agreement phenomena, and in line with phenomena like honorific marking (in Japanese and other languages, as I discussed in Chapter 1), a syntactic analysis that can account for sociopragmatic behaviors is more richly explanatory and better suited for cross-linguistic generalization.

If this is a goal for the theory of grammar, then pragmatics (appropriateness, implicatures both intended and interpreted, and how stance will be conveyed) must be evaluated compositionally, in the same way that semantics must be evaluated compositionally (post-Spellout, in the LF component). This means there must be interface and interchange between the narrow syntax and the compositional sociopragmatics at the point of Spelling out a structure. Knowing that the evaluation of sociopragmatic well-formedness takes place at Spellout makes a specific prediction about syntactic behavior: that only whole phases are eligible for this process, since only whole phases are shipped to Spellout upon their convergence. In the case of DPs, this means that it isn't until we merge the phase head (D) and send the whole DP to the interfaces that we will attempt to calculate its appropriateness.

If, as I proposed in Chapter 2, *he* and *she* are essentially distilled versions of the *IS-CALLED(x)* predicate, it remains a serious question why anyone should ever report

UNGRAMMATICALITY specifically—not social inappropriateness or discomfort, not semantic infelicity, but syntactic ungrammaticality of the sort Pullum reports[18]—based on gendered mismatches between pronouns and antecedents. Certainly no one is reporting similar intuitions for the phenomenon of calling someone an improper or incorrect proper name, as in (28) below:

(28) The 45^{th} president, Donald Duck, is on TV.

Reactions to (28) would (very reasonably) range from a comic interpretation to a presupposition failure to simply an interpretation of (28) as a violation of the maxim of Quality—a lie. These are quite different than the reactions to (29), below, even though they equally rely on world knowledge.

(29) Donald Duck$_1$ adores her$_1$ nephews.

As I discussed in Chapter 3, participants in an acceptability study showed considerable variation in response to 'gender mismatches,' and some of them definitely reported ungrammaticality. Comments from participants include:

(30) "I notice singular 'they' which is not correct grammar (to most English speakers)."

(31) "The grammar was incorrect - the pronouns were not in agreement. They sounded like they had been translated. There were missing words. The names did not always match the assumed gender."

However, many participants also specifically commented on a gradient of ungrammaticality, noting that pronoun mismatches were 'less bad' than other distractors:

(32) "I also noticed that pronouns often didn't ""match"" with the gender typically associated with given name, but this didn't impact my naturalness rating as much as the other mistakes."

[18] In various Language Log posts, 1, 2, 3

(33) "The sentences with mismatched pronouns were more natural than those with missing objects. The singular/plural pronouns mismatch was less natural than the expected gender pronoun mismatch."

Chapter 3 discussed other factors influencing participants' conscious awareness of dsT as a sociolinguistic variable; the digital/downloadable appendix in Chapter 3 includes all comments made in the survey. Furthermore, it seems clear that some gender mismatches are much more broadly considered to be ungrammatical—in particular, anaphora with 'mismatching' gender features caused lower ratings and higher surprisal in Prasad et al's (2018) ERP study, which used bound reflexive forms (example reproduced in ((34)) below). It is also worth noting that Prasad et al found a P600 effect in their ERP study comparing different types of number/gender mismatches between antecedents and (bound) reflexives—this effect is traditionally associated with syntactic ungrammaticality.

(34) John decided to treat themselves to sushi. (Prasad et al. 2018)

If it is the case that gender 'matching' is part of what determines grammaticality for bound pronouns (as in (34)), then it remains a mystery why variation and mismatches should be slightly more tolerable—but not *perfectly* acceptable—in free pronouns (as discussed in Chapter 3).

To sum up some relevant facts: first, it seems to be the case that many speakers do find the binary gendered pronouns mutually incompatible most of the time –meaning that gender features are still optional adjunct features, abstracted into a calling-predicate, but otherwise still exist (and that, as this chapter shows, the meaning of those features is socially mediated). Second, Chapter 2 showed that in the use of predicative pronouns the predicate feature is propositional—that is, it's asserting gender rather than entering it into the common ground (through implicatures or by whatever other means). It is also worth noting that what makes something 'background' information (inferable through implicature or pragmatic presupposition) in DP is the existence of a referential D head that merges in some way with the nP (or highest nominal pre-D projection) (Gutzmann

and McCready (2014) discuss this, and I will expand in Chapter 5 on the relevance). The fact that the merging of a referential D head is crucially linked to interpretation of pronouns aligns with the statement above that sociopragmatic evaluation occurs at the point when a phase converges and is Spelled out.

It is also worth observing that what people are calling unacceptable in gender mismatches with free (referential) pronouns is particularly targeted at singular *they*—almost no one is seriously arguing that a gender mismatched name with *he* or *she* is actually ungrammatical. Moreover, Chapter 3 showed empirically that many more speakers find singular *they* grammatical when it is anteceded by a generic or indefinite DP. Thus, the grammaticality seems to derive from the specificity (or the referentiality) of reference—which is an artefact of combining with a referential D. When there are (optional adjunct) gender features (e.g. "masc"/"fem") attached to an antecedent then it is actually possible for speakers to detect a (sociopragmatic) anomaly in pronoun mismatches. These features aren't "grammatical" features per se—they operate much more like honorific "features" of pronouns like *Usted*—but these features must still be accessible by the grammar such that their meaning is evaluated compositionally and compositionally (phase by phase as spell-out proceeds).

When, however, there are no gender features available for an antecedent (for whatever reason—a speaker doesn't know, or the referent is nonbinary, etc.), there are two possibilities that differentiate speakers who report ungrammaticality and those who don't. For speakers that allow dsT as an option, the computation on matching and appropriateness is reliant upon conversational implicatures and other pragmatic factors that I have discussed in this chapter. For speakers that don't have dsT as a part of their grammar, they end up in a bind: their grammar doesn't have a way to resolve specific reference without the gender features—hence why these speakers only report ungrammaticality of singular *they* when it is referential.

The job of Chapter 5 is to provide a clear syntactic account that adequately explains this split; my proposal hinges upon the availability of *n* to D head-raising as the mecha-

nism that composes pronouns. I show how in cases of non-referential singular *they*, the pronoun does not incorporate a referential D, because the pronoun is still being interpreted as a variable. In these cases, the pronoun is still (covertly) behaving like a predicative pronoun, and the D restricts its reference quantificationally rather than referring to an entity directly.

Chapter 5
RAISING PRONOUNS

5.1 Introduction

This chapter will propose and give details for a head raising analysis of pronouns that can account for both N-like and D-like pronominal behaviors in English and cross-linguistically, as well as give insight into how the syntax can encode sociopragmatic features and relationships in a way that is constrained by principles of syntactic theory. The goal of this chapter is to clearly illustrate how head raising can cause a new pronoun to emerge, in addition to satisfactorily explaining previously-existing phenomena.

5.1.1 Review of previous chapters

In Chapter 1, I reviewed accounts of pronouns that argued for differing positions of pronouns in the nominal functional domain. Many foundational accounts (e.g. Postal 1966) propose that pronouns are determiners (D) and not nouns, supported by evidence from constructions like *we linguists* as well as analogy between other functional/lexical splits; D-pronoun accounts tend to either claim that pronouns are fully intransitive D heads with no NP complement (Postal 1966), or that pronouns are D determiners that select a phonologically null NP complement (Elbourne 2013). Other accounts attempt to explain both N-like and D-like pronoun behavior through mixed analyses: this includes analyses of category switching (Melchin 2015), N to D raising depending whether it involves strong or weak pronouns (Cardinaletti 1994), and my adaptation of head raising in the current chapter. The third type of account of pronouns is similar to Cardinaletti's, in that different sub-types of pronouns are claimed to be of different categories; Déchaine and Wiltschko (2002) provide the most well-known example of such an analysis. This chap-

ter attempts to retain the empirical strength of Déchaine and Wiltschko's generalizations and cross-linguistic insights found in each of these three types of analysis through head raising from *n* to D.

The first chapter also examined accounts of where in the syntactic hierarchy ϕ-features are located, in pronouns or in other types of DPs; this is important to the current investigation of English pronouns because I am making some strong claims about the nature of gender features in particular, and those claims need to be compatible with previous insights about the nature of ϕ-features. I reviewed proposals in which ϕ (or gender) features are on or above the D head (e.g. Sauerland 2013), gender on *n* (Kramer 2016), and gender as a split phenomenon where both *n* and D contribute different *types* of gender (also Sigurðsson 2018). In my proposal I will attempt to reconcile the properties of D-gender and *n*-gender by drawing parallels to honorification on pronouns; I will suggest that social gender enters the derivation at *n* but is evaluated for social appropriateness at D.

In Chapter 2 I provided evidence that English pronouns can act as predicates rather than as full referential DPs. I first examined pronominal relative clauses, where pronominal heads are necessarily construed as generic and combine with restrictive relative clauses at the nP level, denoting generic kinds; then I showed how English pronouns can be 'depronominalized,' appearing with an external article and with some other nominal modifiers (including but not limited to relative clauses). Based on evidence from PRCs and depronominalizations I proposed that English pronouns enter the derivation lower in the nominal domain, at *n*, and that the denotation of pronouns that *remain* in *n* (when they are blocked from raising) is a calling-predicate. The calling-predicate is a semantic predicate which connects naming conventions to named entities. Naming conventions, which I discuss more thoroughly in Chapter 2, are the social conventions by which names are associated with things; named entities are not necessarily all entities in the world, but rather entities that are associated with a particular name by this social convention.

In this analysis, pronouns are analogous to proper names in their power of reference

(or predication, when they fail to raise from *n* to D)—the difference being that names provide a root node containing phonological material which informs the calling-predicate, while pronouns are the purely functional skeleton of that, and are not merged with a lexical root. In this chapter I will show how a head-raising analysis can account for the properties of both predicative pronouns (which enter the derivation in *n*, and remain there) and referential pronouns (which raise to D).

Chapter 3 showed data that suggest there is a (semi-)grammatical change happening in American English, wherein the allowable antecedents for singular *they* are expanding to include more definite/specific uses. Apparent time data and acceptability ratings supported the hierarchy from (1) – (3) below:

(1) Everyone . . . they
(2) The professor . . . they
(3) Parker . . . they

Chapter 4 explored intraspeaker variation in different deployments of singular *they*, showing that the underlying feature specification allows speakers to exploit implicatures and pragmatic constraints around stance and politeness to achieve certain goals in discourse.

In this chapter I put forward a proposal intended to group the wide array of empirical phenomena under the shelter of a unified syntactic structure. This proposal is intended as an explanation for why pronouns appear to switch category, and how pronominal syntactic behavior can be linked to the pragmatic use of pronouns in context.

5.1.2 Detailed proposal

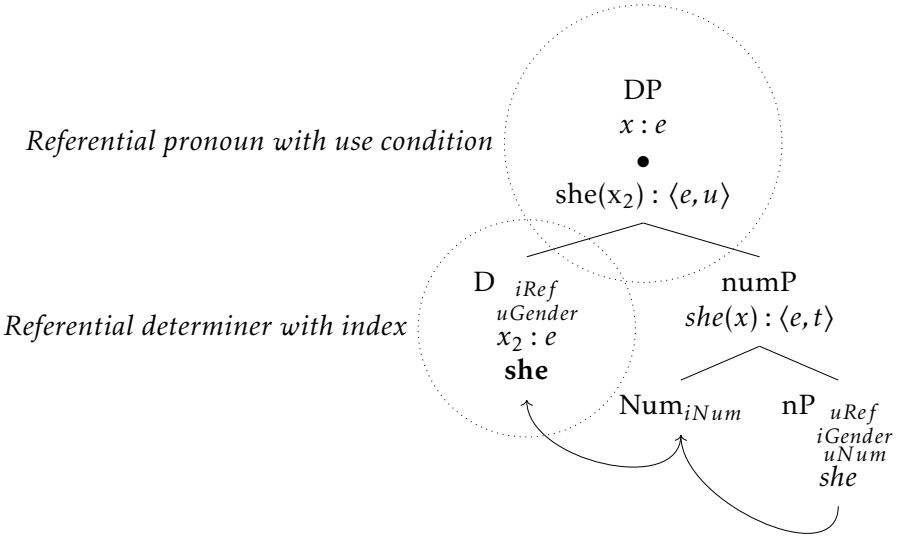
In this section I detail my syntactic proposal for *n* to D movement in pronouns.

Below is a schematic representation of the 'end state' of three sub-types of pronouns. In Section 5.2 I will elaborate further what these sub-types consist of and how they build off of cross-linguistic work. The additional notation in the structures in (4) indicate use-conditional semantic composition, which I also review carefully in Section 5.2; *e* indicates

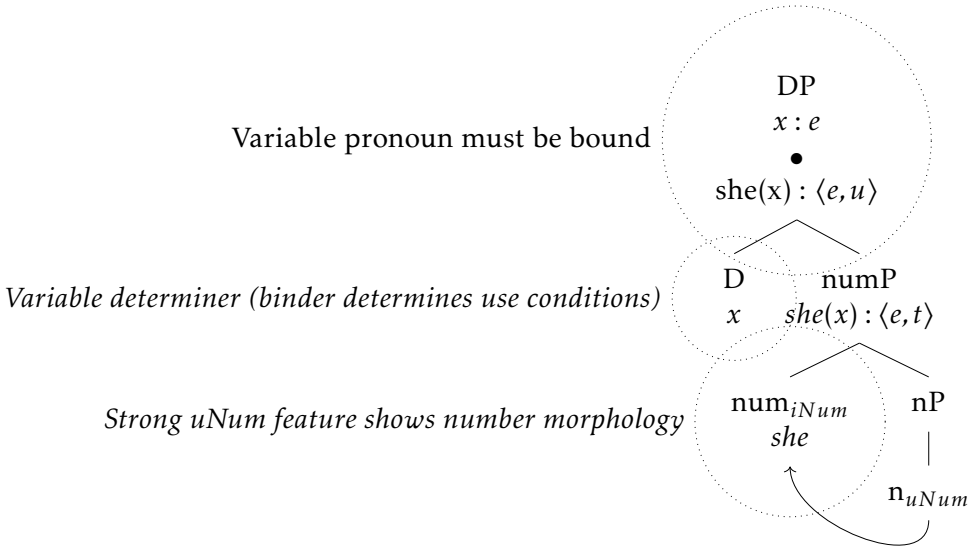
an entity, and u is a conditional variable introduced to account for *use* conditions (rather than truth conditions, traditionally notated as t). The structures below represent pronouns which operate as referential descriptions (4a), pronouns which operate as bound variables (4b), and the predicative pronouns that I showed in Chapter 2 (4c). These three structures are the result of three different landing sites, where pronouns end up after raising from their initial merge at n.

(4) **Sub-types of pronouns based on n-raising**

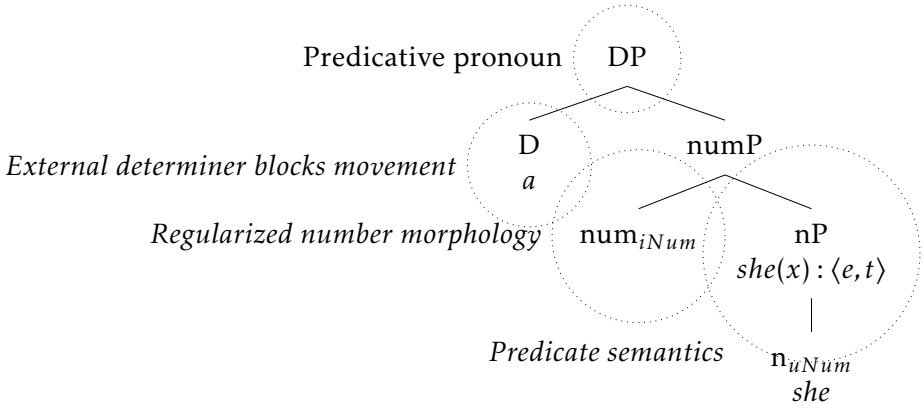
a.



b.



c.



The featural specification I assume in this analysis includes interpretable and uninterpretable Gender, Num(ber), and Ref(erence) features. Gender is interpretable on n to maintain the insights from Chapter 2 around the 'sense' of gender found in predicative pronouns; later in this chapter I discuss how the uninterpretability of Gender on D relates to the social evaluation of the feature. Number is interpretable on Num and uninterpretable on n, again to account for predicative pronouns where overt numerals in Num are not morphologically merged with pronouns. I further propose that the presence or absence of uGender on referential D constitutes the syntactic parameter which determines whether speakers have dsT as a grammatical part of their dialect. For speakers who lack the uGender feature on D, the vestigial gender features still present are not iGender (a formal feature, which is implied to be obligatory) but rather are "optional adjunct features" as Bjorkman (2017) suggests. In those cases, <Masc> or <Fem> may be present as independent features, rather than values of Gender.

The Reference feature I propose is the locus of indexation (with direct referents), person/participant features (for first/second person pronouns), and variable binding (for variable pronouns); it is interpretable on D and uninterpretable on n.[1] For speakers who lack uGender on D, the Reference feature shown in (4a) is the only remaining feature

[1] This chapter does not deeply explore matters of person features, but I will note that person features do seem to be subject to similar social evaluation as gender features; additionally, as I show later, $1^s t$ and $2^n d$ person pronouns can appear lower in DP as predicative pronouns. This warrants further exploration which is outside the scope of this chapter.

responsible for agreement and head movement to D in referential pronouns, while for speakers who retain uGender both Gender and Reference are implicated in this movement.

In the structures I show in (4) uninterpretable features can be checked or valued by establishing an Agree relationship, which can result in either movement (which I discuss below) or concord (which I explore in Section 5.3.2).

The features I propose above, combined with the ones Bjorkman (2017) proposes, can account for differences between English pronouns, here formalized as insertion rules (as in Distributed Morphology). The feature Reference can, for the purposes of this analysis, be valued with participant features (specifying person); 3^{rd} person pronouns are therefore simply those pronouns whose Reference feature is not valued as including a participant.

(5) 1^{st} and 2^{nd} person pronouns

a. [Reference:Speaker] [Number:singular] ↔ *I*

b. [Reference:Speaker] ↔ *we*

c. [Reference:Hearer] ↔ *you*

(6) 3^{rd} person pronouns (Bjorkman 2017:7)

a. [feminine] [singular] ↔ *she*

b. [masculine] [singular] ↔ *he*

c. [inanimate] [singular] ↔ *it*

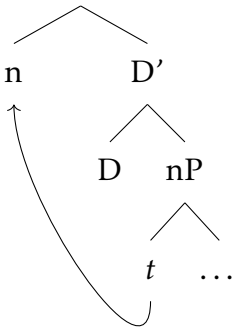
d. elsewhere ↔ *they*

For the remainder of this chapter I primarily focus on how gender features are valued (or remain unvalued) in my discussion of referential singular *they*, as well as the effect of head raising on how the pronouns in (5)-(6) can be used.

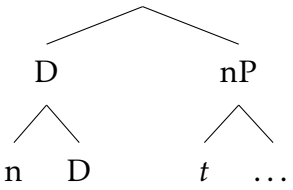
The head raising process that I will adopt is a two-step process based on Matushansky's (2006): movement itself takes place in the narrow syntax, and the moving head moves to merge with the root node; then m(orphological)-merger occurs in the morpho-

logical component, resulting in a composite head that thereafter moves as a unit. (7) shows a reduced schematic of this process (without including semantic notation or feature structures for simplicity).

(7) a. (*n* moves to merge with root)



b. (*n*+D morphological merger)



In the next section, I give the relevant background for the components of this proposal, including background on the three sub-types of pronouns, background on the syntactic processes involved in head movement, and background on use conditional semantics (especially as applied to referential pronouns, as I do above). Section 5.2 also discusses the advantage of differentiating between head movement that involves a phase head from movements that do not involve phase heads, with particular attention towards how phase edges mediate context sensitivity for gender features on pronouns.

5.2 Background

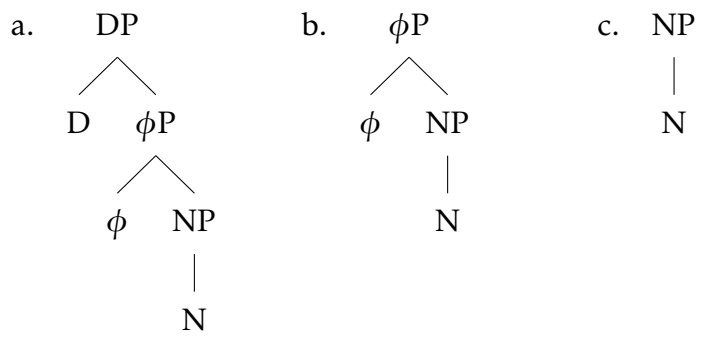
This section provides the necessary theoretical background for the proposal I give above, including precedent for analyzing sub-categories of pronouns, motivation for the two-

step head raising process, and explanation of how use-conditional semantics can be incorporated into the derivation. I also address the significance of D as a phase head, and show how use-conditional semantics can interface with pragmatic context when mediated by a phase edge linker (which connects a syntactic structure with pragmatic context).

5.2.1 Sub-types of pronouns

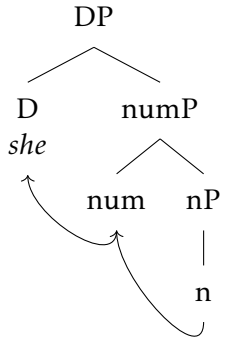
In Chapter 2 I proposed that depronominalizations and PRCs, grouped together as predicative pronouns, could be analyzed as *n* heads that 1) lacked a lexical root, and 2) were blocked from raising to D by other elements that base-generated in D. As I argue below, a pronoun that starts lower in the nominal domain *must* raise to D in order to become a referential, free pronoun. As a consequence, I am in a sense back-engineering Dechaine and Wiltschko's (2002) analysis of different pronoun types as constituted from different levels of functional projections; the difference is that while Déchaine and Wiltschko (2002) analyze non-free, non-referential pronouns as less than full DPs (PHI-Ps and nPs, for them), I instead will argue that the full functional spine of the nominal domain is *present* in all pronouns, but can be *null* or composed of other elements besides a definite D. Below is a schematic representation of the full typology taken from Déchaine and Wiltschko (2002), 'translated' to show head-raising instead of absent projections:

(8) (Déchaine and Wiltschko 2002)

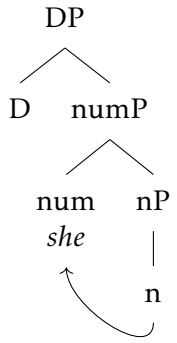


(9) **Adapted for head movement**

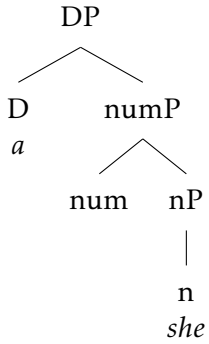
a. Referential pronouns:



b. Bound variables:



c. Predicative prns:



I follow Déchaine and Wiltschko (2002) and Kramer (2016), among others, in placing gender features (at least) on the *n* head. This is, first of all, essentially required for my derivations to converge for both referential and predicative pronouns. However the added benefit is that this provides an account of gender features that in some way allows for parallels to be drawn between languages with gendered noun-classes (grammatical gender, as it were) and languages with 'natural' gender (the social dimension that I discussed extensively in Chapter 4).

5.2.2 Head movement

In this section I will very briefly summarize Roberts' (2011) review of the history of theoretical conceptions of head movement from GB into the Minimalist Program, before outlining how my approach fits into the theoretical landscape sketched here. Head-raising as a theoretical apparatus straddles the line between syntax and morphology in part because of the close correlation between (apparent) head movement and rich inflectional morphology. I will discuss general accounts of head-raising not comprehensively, but rather as they are relevant to the necessary mechanism for n-to-D.

The 'traditional' account of head movement in the Government and Binding (GB) framework places several strict restrictions on how the movement may proceed: heads

can only move to heads (not phrases), heads may only move upwards under c-command (to bind the trace, in GB terms), head-movement must obey relative minimality (meaning it cannot skip intervening heads); and the cyclic effect of rolling up heads is achieved through head-to-head adjunction (Roberts 2011 reviews Koopman 1984, Travis 1984; and Baker 1985a, 1988 in much more detail).

In the transitional period between GB and Minimalism, some theory-internal problems with this account of head movement were raised (as in Chomsky 1995, 2001). First, since Chomsky claims that head movement never affects interpretation, there would need to be some mechanism for features to differentiate between heads and phrases in order to ensure that (interpretable or valued) features that were valued by head movement don't get interpreted–a parsimony problem results. Second, head-to-head adjunction violates the extension condition, in that a moving head adjoining to an already-merged head is not extending the structure itself. Third, head-to-head adjunction means that the moved head does not c-command its own trace; and finally, cyclic roll-up of heads through the functional spine isn't successive-cyclic (which would require excorporation from the head-adjunction complex). Chomsky concludes based on these issues that head movement is not part of the narrow syntax, but is perhaps better characterized as a morphological process (Chomsky 2001).

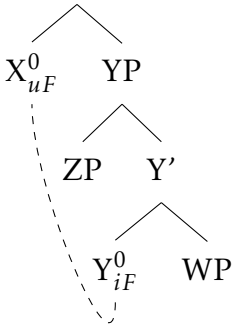
Alternative approaches to head movement outside the narrow syntax include PF movement (e.g. Chomsky 2001, Boeckx and Stjepanović 2001), remnant phrasal movement (Koopman and Szabolcsi 2000, Nilsen 2003, Bentzen 2007, 2009, i.a.), and reprojection or relabeling approaches (e.g. Donati 2006, Bury 2003, 2007, i.a.). I largely leave aside these alternatives for the important reason that, in order for my analysis of *n*-to-D movement to be legible, this type of head movement does in fact have to interface with the LF component, and thus cannot be banished from the narrow syntax entirely.

The strongest interpretation of the Strong Minimalist Thesis (SMT) (cf. Chomsky 1995, 2000, 2001, 2004, 2008) suggests that fewer, more generalizable operations in the core syntax reduce the cognitive load of language acquisition. Under this thesis it is

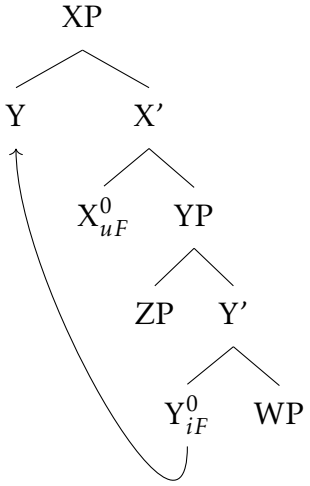
desirable for all movement to be included under the mechanism of Internal Merge(IM) (Chomsky 2004). If head movement is to be subsumed under Internal Merge–that is, if head movement can successfully be reduced to merging an element that has already been entered into the derivation –then there should not be a non-syntactic way of doing this. Rather than reinvent the existing mechanics of movement, Roberts (2011) suggests adopting a hybrid approach like Matushansky's (2006), where "head movement" is in fact a two-step operation, consisting of head-to-specifier movement and morphological rebracketing (as in Marantz 1984, 1988), which combine to obtain the effects of GB-style head-to-head movement.

Matushansky (2006) proposes that head movement is a two-step operation, consisting first of Internal Merge where a head moves to the specifier of its target head (thus obeying the Extension Condition), and then undergoes m(orphological)-merger with the target head. This captures the important insight that, when head movement has taken place, the moving head and the target head act morphologically as a single constituent after movement.

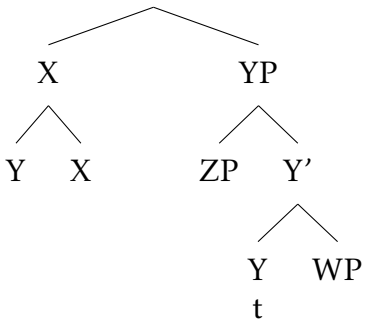
(10) a. C-select (Matushansky 2006:81)



b. Y moves to Spec,XP (*ibid.*)



c. m-merger combines X+Y (*ibid.*)



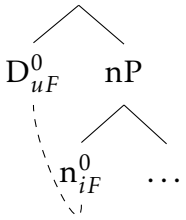
For Matushansky, m-merger is an operation in the morphological component, which cyclically follows syntactic movement; this means that the movement itself takes place in the narrow syntax, but the m-merger occurs at Spell-Out[2]. Thus, In order for the structure given in (10) to correctly derive iterative head movement, Matushansky suggests that a partial spell-out triggers m-merger in (10c), allowing the composite Y+X head to move as a single head later in the derivation. This partial spell-out must not, however, make the XP structure opaque to further operations; so long as the composite Y+X head is at the *edge* of the most recently combined structure, however, Y+X should be available

[2] Matushansky needs to assume a very strong version of cyclic spell-out to derive iterative head movement, along the lines of 'every phrase is a Phase.' This is a weakness of any syntactic account of head raising that separates a syntactic and post-syntactic operation, but it is surmountable for my purposes; I expand on this below.

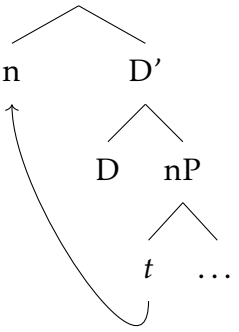
to higher probes. This is on analogy with other cyclic operations based on phasehood, and does successfully prevent head movement from 'skipping' intervening heads in a syntactic structure.

This is the form of head movement that I will undertake for the purposes of this chapter; it is compatible with the blocking effects shown in sections below as well as LF consequences. Adapting my derivation above to match Matushansky's two-step process, referential (free) pronouns involve the movements given in (10). C-select (in (11a) refers to Matushansky's formulation of External Merge controlled by a head's selection of its complement (in this case, D selcting *n*P).[3] (11b) shows movement to Spec,DP (which operates the same way any other Move/Internal Merge would), and (11c) is a representation of the effect of morphological rebracketing. Only Merge (internal and external) occur in the narrow syntax; the m-merger is a morphological operation and necessarily post-syntactic.

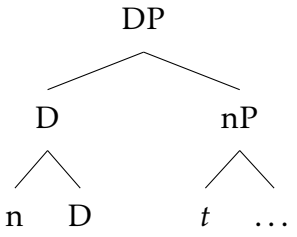
(11) a. (C-select = Matushansky's 10a)



b. (Move to Spec, DP = 10b)



[3] I abbreviate the relevant features in (11); their presence is primarily noted as the mechanism necessary to establish the Agree relationship that will result in movement.

c. 

(M-merger = 10c)

In addition to Matushansky's (2006) evidence for m-merger as an independent operation, this captures the intuitive understanding of pronouns as a "bundle of features" by composing them morphologically from individual ϕ-features (number, gender, person) separately, as well as providing a more generalizable understanding for pronouns that incorporate other features besides person, number, and gender.

The next section turns to semantic/pragmatic interpretation of pronouns, particularly focused on a formalization system for gender that is compatible with the fact that referential pronouns show a particular sensitivity to context that is weaker in variable pronouns and absent in predicative ones.

5.2.3 Use conditions

In this section I will review the semantic analysis that, combined with my syntactic approach, can correctly account for the sociopragmatic sensitivity of pronominal gender for referential pronouns in particular. In incorporating Gutzmann & McCready's (2014) semantic analysis into my syntactic proposal, my intention is to clarify how the semantic/pragmatic component is structurally composed along with the syntactic structure.

The mechanism for evaluating 'natural' gender that I propose resembles Sigurðsson's (2018 and others; cf Section 1.4.4 for discussion of D-gender) in that the D head is responsible for natural gender which is evaluated separately from grammatical gender. As I discuss in detail in the next section, Sigurðsson (2018) proposes that D, as a phase head, is crucially endowed with edge linking features that allow D to access discourse-sensitive context. I diverge from Sigurðsson and most others, however, in that I do not analyze English (and 'natural'-gender only languages) as having additional gender features on D

that must be related to the gender features on *n*. Instead, I use D as the locus of evaluation for the appropriateness of reference. This is an extension of Gutzmann & McCready's use-conditional semantics (2014; also Potts and Kawahara 2004, Potts 2007), which I will briefly review here before showing my own derivation incorporating their system into my existing analysis of head-raising. I discuss the more general application of use conditional semantics in Section 1.5.4 in the course of reviewing how these analyses have been applied to honorific marking.

The primary insight of use-conditional semantics that gives this system an advantage over traditional semantics is that, as I showed in Chapter 4, it is difficult and perhaps erroneous to attempt to discern the "truth" value of the gender features of referential pronouns, even if one is working on the assumption that the phi-features of referential pronouns are presuppositional. Instead of saying whether the use of *she* is "true" with respect to a particular referent, I show in Chapter 4 that the more pertinent question is whether that use is *appropriate*. It is not necessarily the case that instances of misgendering (whether malicious or benign) render an utterance infelicitous, and it is certainly not the case that the use of dsT to withold gender information can render a statement less semantically sound. Thus, the use condition program investigates whether the use condition *u* is valued as appropriate or inappropriate; this is entirely separate from whether a truth condition *t* is valued as true or false. This system has the additional advantage of being able to be imported into the syntactic structure, so that semantic structure can be composed merge-by-merge in a process that closely shadows the process in the narrow syntax.

The implementation of use conditional semantics that I adopt for referential pronouns is closely modeled after the closest non-pronominal parallel—referential descriptions. Referential descriptions are full DPs that are used to refer directly to a referent rather than pick one out of a set of possibilities. Gutzmann and McCready (2014) propose a treatment of referential descriptions where the appropriateness (in addition and orthogonal from truth value) is evaluated semantically. Their aim is to combine the two dimen-

sions of semantic content included in a statement with a referential description such as (12) below:

(12) The murderer of Jones is insane. (Gutzmann and McCready 2014: 58)

1st dimension: propositional content: Smith is insane.

2nd dimension: description content: Smith is the murderer of Jones.

Since the reference to Smith can be upheld in (12) even if the descriptive content (that Smith is the murderer of Jones) is not true, it is not necessarily appropriate to analyze descriptive content as an (uncomplicated) presupposition. Instead, Gutzmann and McCready show four possible values:

<1,1> Smith is insane and he is the murderer of Jones.

<0,1> Smith isn't insane, but he is the murderer.

<1,0> Smith is insane, but he's not the murderer.

<0,0> Smith is neither insane nor the murderer.

To account for this, Gutzmann and McCready (2014) propose a two-dimensional semantic model, where truth value (and conditions) and felicitousness (and felicity conditions) are evaluated on different dimensions; they give the example of expressive adjective uses:

(13) The damn dog howled (Gutzmann and McCready 2014)

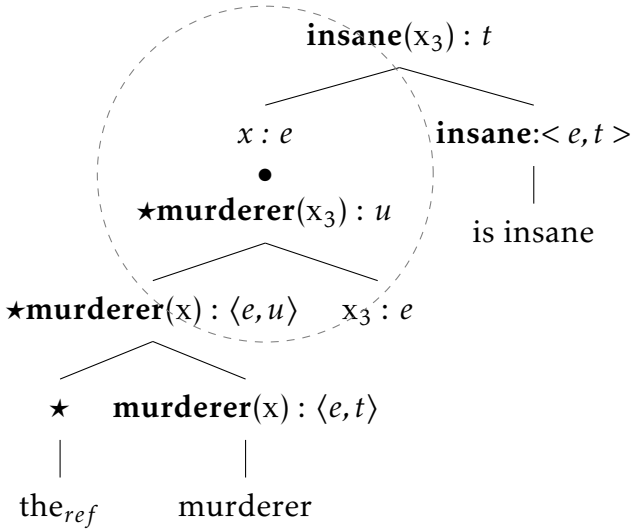
a. **true** if the dog howled

b. **felicitously used** if the speaker feels negatively about the dog

In applying the two-dimensional model to referential descriptions, Gutzmann and McCready (2014) locate the evaluation of felicitousness (a variable called *u* in their LF denotations) at the point in the derivation where the description combines with a referential index—that is, an index that is construed by way of a determiner. For Gutzmann and McCready (2014), the definite article *the* combines with *murderer* first, then an index is combined with the resulting DP at which point appropriateness can be evaluated (and

the referential description can then combine with the rest of the sentence). In Gutzmann and McCready's notation, the two dimensions of semantic meaning (truth and appropriateness) are shown separately and can be combined by an operator '•'. The *u* marker is the use-conditional equivalent of a *t* marker: *t* is a semantic type that can be valued {1,0} to indicate *truthfulness*, and *u* is a semantic type that can be valued {1,0} to indicate *use-appropriateness*. In defining referents, a numeral indexes a particular referent which is type *e* (entity).

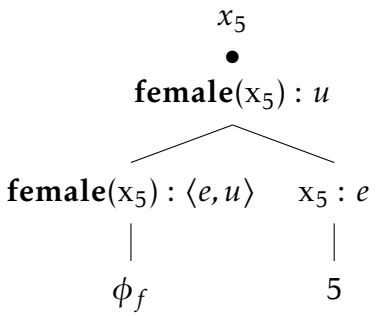
(14) **?**



Based on the derivation Gutzmann and McCready show for (14), the *truth* conditions will be evaluated at the root of the tree in (14), after the semantic type is resolved to *t*; the *use* conditions will be evaluated at the circled node, $x : e \bullet \text{murderer}(x_3) : u$, which can be paraphrased as '*the murderer, who is* x_3*, if it is appropriate to call* x_3 *the murderer.*'

Like I did in Chapter 2, Gutzmann and McCready show that pronouns can be semantically described using essentially the same mechanism as other R-expressions. I made the comparison between pronouns and proper names, while Gutzmann and McCready show an equivalent parallel between pronouns and referential descriptions. In the notation below, they replace *murderer* with ϕ, similarly to how Chapter 2 conceives of gender features.

(15) *pronouns* = $x + \phi$ (Gutzmann and McCready 2014: 63)
she$_5$ ≈



Gutzmann and McCready make this parallel quite explicit: "*[a]ccording to this analysis, free pronouns can be seen as minimal versions of [referential descriptions] that bring their own lexical content instead of incorporating an NP*" (2014:63). My adjustment to this analysis is the separability of the referential and predicative component, which I account for through head-raising. For pronouns, the gender component (what is denoted as ϕ above) can be forced to remain predicative when head-raising is blocked (by GEN or another external determiner, yielding predicative pronouns), and can become referential at the point that it combines with a referential index (yielding free pronouns). Thus, when a pronoun *does* successfully raise to D, it triggers a pragmatic evaluation of felicitousness *not* in general, but with direct respect to the particular referent—meaning that something like misgendering is analyzable not as a presuppositional failure, but instead as infelicity derived directly from the evaluation of **female**(x$_5$):u in (15) above.

For non-referential pronouns such as bound variable pronouns and predicative pronouns, I extend Gutzmann and McCready's analysis to more closely mirror Déchaine and Wiltschko (2002)'s: that is, bound variables are pronouns that do not combine with a determiner linked to a particular referent, but rather are bound by operators (such as quantifiers or indefinites)—this is instantiated by a failure of the pronoun to raise to D. Predicative pronouns likewise do not raise to D and cannot be interpreted referentially.[4]

[4]Certain bound pronouns such as anaphors can get interpreted referentially, and therefore under this analysis must combine with a D that is linked to a referent; however later on I will discuss how the conditions on binding and felicity of gender are much stricter than those on free pronouns, due in part

5.2.4 Phase edges

In this section I turn to the sensitivity of phase edges as the locus of context-sensitivity in discourse; the structures I proposed in (4) differ in the ultimate landing site of the pronoun, and my proposal hinges crucially on the fact that pronouns that land in D are pragmatically controlled in a way that other pronouns are not. The explanation for this is dependent upon an analysis of D as a phase head (cf Citko 2014 i.a. for independent discussion of the phase status of D). The merging of the phase head D in my analysis links the nominal structure to the DP-external context; this is based on Sigurðsson's (2018) proposal for the interaction between pronominal gender and the edge-linking ability of phase heads. While both my and Sigurðsson's proposals indicate that the discourse-sensitivity is a general property of phase heads and not particular to D, I will restrict my attention to pronouns because their discourse-sensitivity is dependent upon referentiality. This section (and this chapter in general) attempts to explain why referential pronouns can be affected by clause-external antecedents, while bound variable pronouns and predicative pronouns are much more restricted.

In his analysis, Sigurðsson (2018) identifies one of the serious problems of ϕ-features on free, referential pronouns; that free pronouns are free seriously limits the ability of the narrow syntax to force feature-matching even though empirically speaking feature-matching seems to be very robust, both in languages like English with only pronominal gender, and languages like Icelandic with pronominal and nominal (grammatical) gender features and rich concord DP-internally. Because free pronouns can be coreferential with antecedents several clauses (or utterances) removed, and indeed because free pronouns can be anteceded non-linguistically, CP-internal Agree under c-command would not be a sufficient mechanism for checking matching ϕ-features unless it was so unrestricted as to be descriptively useless. Instead, Sigurðsson suggests that the phase head D mediates context-scanning, analogous to the context-scanning by C to establish speaker-relative

---

to (I will argue) the *way* in which the D in anaphora gets linked to a referent.

tense or aspect.[5]

Sigurðsson starts this analysis from the premise that languages optionally express n-gender (his term for what is otherwise known as grammatical gender or noun-classes; cf Kramer 2016 i.a.) and/or D-gender (what would often be called "natural" gender). Because I have so far primarily been discussing gender as a social relationship mediated through pragmatics, his analysis of D-gender is the most similar to mine. For Sigurðsson, pronouns enter the derivation without valued ϕ-features, and the phase-head D must mediate their valuation through context-scanning; he gives a similar account for valuing the D-gender of articles in languages like French. The exact mechanism of valuation is not deeply explored—Sigurðsson does not explicate *how* a referent's "biological gender" (2018:12) is transformed into formal features valued on the D-head.

Once the DP is valued for gender, however, this can feed CP-internal syntactic agreement (such as subject-verb agreement). In his formalization, this means that CP-external gender features are transmitted to the DP and are available for what he calls "feature recycling." Under feature recycling, gender (or ϕ) features that are valued on D under context scanning are then available for agreement relationships internal to the DP (like concord) and its containing CP (like gender agreement on predicates). In this chapter I primarily focus on the context-scanning mechanism proposed by Sigurðsson (2018), since English does not have any DP-internal concord or gender marking on predicates; that is, my interest is primarily in how the gender on D is linked to context, rather than its later participation in local agreement.

Sigurðsson's analysis is in many ways similar to the proposal given by Kučerová (2018), in which Kučerová analyzes Italian nouns which may or may not carry implications of the referent's gender–i.e., *il chirurgo* (the.M surgeon) may, for some speakers, refer only

[5] The discourse-sensitivity of the D phase edge is something that is elsewhere supported by investigations into focus and the left periphery of the DP edge, as by Aboh (2004) and Ntelitheos (2004). While I do not include specific discussion of the left periphery in and around Spec,DP, these studies should be taken as independent evidence of the validity of locating discourse functions at the phase edge in the nominal domain, as well as supportive of the structural parallels between the clausal and nominal domain implied by both my and Sigurðsson's proposals.

to male surgeons, while for others it is possible to refer to female surgeons with the apparently-masculine form. For Kučerová, nominals can either bring gender features from the lexicon (in which case they are not interpreted to reflect social gender) or the DP can have its gender features valued at the syntax/semantic interface.

For the purposes of this chapter I will adopt Sigurðsson's version of gender valuation rather than Kučerová's, although the two both share important characteristics (including the dependence on phasehood for context-scanning). Before adopting Sigurðsson's analysis, I will make some modifications that move it closer to Kučerová's model of gender valuation: first, the suggestion that *interpretable* gender features (which he specifies D-gender to be) must be *valued* is somewhat unintuitive considering the inherently-social nature of these features; second, Sigurðsson's analysis does not and cannot possibly account for any of the variation or discourse-driven gender manipulations that I showed in Chapter 4. Taking the example of discourse-driven gender that is the least controversial (in my experience), (16) below shows an example where a pragmatic antecedent has a binary gender, but the speaker uses an ungendered pronoun because gender is simply not relevant to the conversation.

(16) **Context**: *The speaker is telling a colleague about a student's exceptional final paper. Speaker A has been working with the student, Sarah, all semester, and knows she prefers to be called* she.

They're my top student this semester. **Their** analysis of Yiddish diminutives was really inspired, I didn't know **they** had access to a native speaker!

The context, as given in (16), is sufficient for the speaker and hearer to construe reference—and, as mentioned, the speaker *knows* the referent personally and has a good idea of their social gender. The speaker can still opt to use *they* because they don't feel that gender is a relevant part of their conversational goals; or rather, for this speaker, the drive to minimize irrelevant information (Relevance) outranks the drive to maximize informa-

tion (Quantity). Sigurðsson does not discuss singular *they* in this article and therefore has nothing to say about genderlessness in English; but under the system he has laid out, gender valuation must apparently be obligatory when the gender of the referent is known.

Another problem with Sigurðsson's conception of gender marking as 'valuation' is the assumption that there is any set 'value' of the gender of a given referent which is objective, uncontroversial, and universally known. This account is difficult to adapt to gender relations around transgender people (including misgendering) without some adjustment. Moreover, in using the term "biological gender" Sigurðsson betrays a reliance on conflation between sex and gender that leaves his theory lacking in explanatory power for many social contexts. However, his analysis of edge linking through a phase head is still appropriate for the theoretical framework of syntax we are working in, and reliably explains the context-dependence of pronominal reference and features. Thus, I will make the following adjustment to his proposal before adopting it: phase heads are still responsible for scanning context and linking phase-external context with phase-internal features, but can only do so through providing anchoring indices against which already-valued features may be evaluated.

My adaptation of Sigurðsson's analysis therefore includes uGender on D (rather than his iGender) as well as an interpretable referential linking feature (iRef) which is a formal reflection of his edge-linker.

In terms of D-gender, what this means is that D still retains its edge linking feature and still scans the context—what it scans for, however, is only a link between the referential index with which it is endowed (x_3 in Gutzmann and McCready's terms above) and a viable entity for reference. Once the edge linker links the index with the referent, the appropriateness of the gender of a pronoun can be evaluated sociopragmatically. This is a hybrid approach, incorporating Gutzmann and McCready's use-conditional semantics into Sigurðsson's syntactic proposal to fill in the explanatory gaps. It also solves the problem of valuing interpretable features: in my adjustment, gender features are features that enter the derivation on *n* and raise to D in order to get evaluated. This means they

can still be interpretable, and potentially can still contribute to agreement operations with other local elements (as I will discuss in Section 5.4.1 below).

This hybrid analysis yields several predictions, which I will show to be borne out by the phenomena discussed in this dissertation so far. First, it predicts that the syntactic and semantic differences between referential and predicative pronouns are linked. Referential pronouns will be evaluated by felicitousness with respect to a particular referent, because the phase-edge anchor responsible for linking a referent to a DP in context is also the reference point against which appropriateness is judged. It also predicts that *non*-referential pronouns will not be judged for appropriateness in the same way, since they lack a referent to socially evaluate or relate to. Predicative pronouns therefore will, like other nominal predicates, be evaluated truth-conditionally and yield a function over a set of entities. This follows not only from Sigurðsson's analysis of edge linkers but also from the way Gutzmann and McCready conceive of use-conditions as determined with respect to referents.[6]

Second, the combined analysis predicts that felicity will be the deciding factor for referential pronouns with respect to acceptability of gender (mis)matching, rather than (necessarily) grammaticality—*but* that felicity constraints will yield an inability to use certain pronouns REFERENTIALLY by some speakers who can otherwise use them non-referentially (dsT being a prime example of this).

The prediction for dsT that this analysis produces is that, for speakers whose internal grammar of felicity conditions does not allow a particular person to lack gender, these speakers will report ill-formedness only when dsT is used referentially—a difference which is supported by the acceptability survey I have presented in Chapter 3, where

[6]Barbara Citko (p.c.) rightly points out that this predicts that non-referential pronouns can't be used to misgender someone; e.g. *"Every male professor$_i$ respects her$_i$ colleagues."*. In an informal poll conducted on social media, 38% of respondents could not resolve coreference, 11% reported total ungrammaticality of the sentence, 20% said that this sentence DOES constitute misgendering, and 32% said this construction was "rare but possible" (n = 408). These results are highly heterogenous, which I attribute at least partly to the lack of *formal* gender features on the antecedent. Further investigation is necessary as to the social sensitivity of bound anaphors such as these, particularly in light of the mixed responses to my poll.

participants rated non-referentially used singular *they* much more acceptable than the referential use. In the next section I will show a derivation using this analysis which demonstrates how this contrast is derived compositionally.

5.2.5 Deriving n-Raising

This section will step through the particulars of head raising as Matushansky gives them, then show precisely how they will apply in the nominal domain for the purposes of my analysis. At the end of the derivation I will briefly show how each step of the derivation interfaces with the semantics/pragmatics.

In order for Matushansky's (2006) conception of head movement to correctly derive other established facts (such as V to T raising), ordering of merge, movement, and Spell-Out are going to need to be exactly right.

(17) $[_{vP}$ v $[_{VP}$ V DO]]

Suppose the structure in progress is a vP with its head v and its complement VP merged. If you merge the external argument (EA) next, this causes a problem.

(18) $[_{vP}$ EA v $[_{VP}$ DO]]

If EA is the sister of v' then when you then try to move V to v, the V merges with the root node, meaning that the EA intervenes between V and v:

(19) $[_{vP}$ V [EA v $[_{VP}$ *t* DO]]]

In principle m-merger should not be able to differentiate between types of constituents, and should simply combine together two heads when they are immediately adjacent. If EA intervenes, then adjacency can't let m-merger proceed, meaning there must either be way to let m-merger skip over interveners or risk accidentally m-merging the external argument with the v, leaving the V stranded. This is not desirable.

Instead, the derivation *must* trigger head movement immediately when the next head is merged (meaning, you can't merge an external argument first). Call this 'head pulling;'

when a head with uninterpretable features is merged, it will preferentially agree with and trigger another head to IM rather than allow a new EM.

If we institute this strong constraint to derive the ordering, then the derivation proceeds correctly (at least at first):

(20) v_{uF} is merged with VP

(21) v 'pulls' / Probes V for iF

(22) V moves to spec,vP

(23) V+v m-merger

In this example, it's not unreasonable to say that V+v m-merger occurs at this point; additionally, as long as V moves before an external argument is merged in spec,vP, it is even possible to merge the EA before spelling out and triggering m-merge; because v is a phase head, sending the phase for spellout at this point is standard.

However, head movement to non-phase heads (or, heads typically not considered to be phase heads) makes this much trickier. Now moving v+V to T.

(24) $[_{TP}$ T $[_{vP}$ EA v+V $[_{VP} \ldots]\ ]\ ]$

T has, in languages where we expect to see overt verb raising at least, an uninterpretable v feature. Or v has an uninterpretable Tense feature. Either way, T pulls v+V up first (because head movement must take place immediately when a new head is merged):

(25) $[_{TP}$ v+V [T $[_{vP} \ldots]\ ]\ ]$

The external argument (EA) does not intervene in this search by T because, for Matushansky, the driving force of head movement is agreement between features; thus, T presumably must probe v+V for a particular uF that cannot be valued by the EA. If the EA were to raise first and the v+V complex raised second, m-merge would presumably fail and the derivation would crash. This ordering sensitivity is an existing problem for Matushansky's proposal that I do not attempt to solve here, particularly since there may not be a

parallel process in the nominal domain (i.e., I do not posit any external arguments which move from Spec,nP to Spec,DP). For the current purposes the ordering restriction is stipulated as the only way the derivation can converge from this juncture.

After the v+V complex raises to Spec,TP, either subject raising (movement to a second Spec,TP by a DP for case) must happen or m-merge. M-merge would be unexpected at this point because TP isn't a phase, so going to Spell-Out would be non-standard. If instead the subject raises first, TP gets another specifier, this one phrasal:

(26) $[_{TP}$ Subj $[_{T'}$ v+V $[_{T'}$ T $[_{vP} \ldots$]]]]

Presumably this also shouldn't go to Spell-Out, because it's still not a phase until you merge C. C is the next thing to get merged, and if there is any movement to Spec,CP, it needs to happen before the whole phase is sent to Spell-Out as well.

(27) $[_{CP}$ WH $[_{C'}$ C $[_{TP}$ Subj v+V T $[_{vP} \ldots$]]]]

At this point, there is no intervening material between the v+V complex and T, so if m-merge occurs "late" (after C is merged and TP is spelled out altogether) it should still meet the head adjacency requirement for m-merger to be successful. This derives V-v-T raising successfully.

However, if the head movement needed to proceed on to C, there would be a problem: when C is merged, v+V and T haven't yet been m-merged and so there is no reason to expect them to move together as a single unit. This is the problem that n-raising encounters with NumP (standardly a non-phase) intervening between *n* and D when trying to raise to D.

(28) $[_{NumP}$ Num$[_{nP}$ *n*]] — Numis merged with nP

(29) $[_{NumP}$ *n* $[_{Num'}$ Num$[_{nP}$ t]]] — *n* moves to spec,NumP

(30) [DP D $[_{NumP}$ *n* Num$[_{nP} \ldots$]]] — D is merged with NumP

(31) *n+Num isn't m-merged yet, so how can n+Num move to D?

Matushansky (2006:95) proposes a very strict version of Spell-Out in order to deal with the problem of non-phase head intervention: *"However, in order to preserve the intuition that a head created by m-merger (just like a simplex head inserted from the numeration)forms part of the input to the next syntactic cycle (merger of a specifier or merger of the next head), we need to assume a strongly cyclic view of syntax, where each newly merged node is a phase (Merge and Spell-Out)."* In a footnote (2006:n28) on the matter, Matushansky recognizes a middle-ground between the strong every-phrase-is-a-phase approach and the more traditional approach that differentiates phase and non-phase heads, suggesting that each Merge may only trigger a "partial PF and LF Spell-Out" which would depend on the absence of uninterpretable features. For the purposes of this proposal, I will take up the partial Spell-Out version for primarily theory-internal reasons: maintaining a distinction between phase and non-phase heads is still desirable in terms of explaining discourse sensitivity, and there appear to be meaningful differences when pronouns do or do not raise all the way to a phase head.

With the Merge=Spell-Out version, then, *n* raising can proceed uninterrupted:

(32) $[_{NumP}$ Num$[_{nP}$ *n*]] — Num is merged with nP

(33) $[_{NumP}$ *n* [Num' Num$[_{nP}$ t]]] — *n* moves to spec,NumP

(34) $[_{NumP}$ n+Num $[_{nP}$ …]] — n+Num m-merger

(35) $[_{DP}$ D $[_{NumP}$ n+Num $[_{nP}$ …]]] — D is merged with NumP

(36) $[_{DP}$ n+Num D $[_{NumP}$ t $[_{nP}$ …]]] — n+Num moves to Spec,DP

(37) $[_{DP}$ n+Num+D $[_{NumP}$ t $[_{nP}$ t]]] — n+Num+D m-merger

Revisiting the contributions from Sigurðsson (2018) and Gutzmann and McCready (2014), the merging of D triggers particular semantic and pragmatic effects.

(38) $[_{NumP}$ Num$[_{nP}$ *n*]] — *n* has gender features, nP is a predicate $\langle e, t \rangle$

(39) $[_{NumP}$ n+Num $[_{nP}$ …]] — Numhas number features, NumP is a predicate $\langle e, t \rangle$

(40) $[_{DP}$ D $[_{NumP}$ n+Num $[_{nP} \ldots]]]$ Merging D (referential)
triggers context-scanning – finds a referent; DP is an entity $\langle e \rangle$ and, once there is a referent, the entity is a point of evaluation for use appropriateness $\langle u \rangle$

In the next section I will show the desirability of a head-raising analysis for explaining the empirical phenomena around pronouns that I have shown in previous chapters of this dissertation; then I will discuss the predictions made by this analysis for a broader range of phenomena.

5.3 Support for head raising

This section consists of three parts: first, I discuss syntactic support for my proposal, showing in particular how the 'blocking' effect of depronominalizations is a phenomenon which occurs in head movement in the verbal domain–this parallel is most straightforwardly explained through head raising. Next, I discuss how the diachronic development of the current pronominal system in English has led to the current state of affairs, where dsT is undergoing the final stages of what has been a long-term reorganization since the OE-ME transitional period. Finally, I show how the featural specification of different D and Numheads can clearly account for the (micro)parametric variation within English such that speakers in close contact can report (and show) different grammaticality judgments for dsT.

5.3.1 Support in syntax: when raising is blocked

One straightforward syntactic diagnostic for whether a head typically raises is to block head-raising by having another element in the landing site that cannot be combined with the raising element for whatever reason. In looking at verb-raising in Romance languages, we can show that the presence of a modal influences the surface position of the verb:

(41) (Santorini and Kroch 2007[7])

[7] Online resource, no page number; https://www.ling.upenn.edu/~ beatrice/syntax-textbook/ch6.html

a. Elle **travaillait** à peine trois heures.
she work.imperf.3sg hardly three hours
'She used to hardly work three hours.'

b. Elle avait à peine **travaillé** trois heures.
she had hardly worked three hours
'She had hardly worked three hours.'

In simple, finite, declarative sentences, as shown above, French verbs appear before negation and before the adverb *a peine*. When the auxiliary verb *avait* appears, however, the verb appears after *a peine*. It is comfortably established (Emonds 1978 among many others) that this alternation can be accounted for by analyzing the verb as (usually) undergoing movement from the V position to the T position (or Infl, as well as moving through any other functional projections that exist on the sentential spine).

This blocking mechanism is directly analogous to the appearance of an article blocking a predicative pronoun from moving from *n* to D (in (42) below). However, the blocking alone is not sufficient, and the semantic differences between depronominalizations and referential pronouns may be argued to be unattributable to head-raising (if one is following the PF movement analysis). Instead, it is notable that predicative pronouns appear in a different position with respect to other modifiers compared to personal (1^{st} and 2^{nd}) pronouns bearing modifiers:

(42) a. the other she
b. a whole new he
c. we happy few

In previous literature (starting with Postal 1966), constructions like (42c) have been used as evidence for analyzing English pronouns uniformly as determiners. However, I suggest an alternative: (42c) is an example of successful (un-blocked) head-raising with modifiers,

which contrasts with the blocked depronominalized examples in (42a) and (42b). The difference between (42a,b) and (42c), besides the position of the pronoun, is the person feature. Person features have been traditionally analyzed as features of D or another high nominal projection; thus *we* can move up to D in order to get person features and get linked to a referent, whereas *she* and *he*, lacking person features, not only cannot move but are in fact blocked from doing so by the determiners *the* and *a*.

Crucially, the difference is that pronouns like *we* in (42c) do carry reference despite the presence of other nominal modifiers. In instances where third person pronouns appear with modifiers, however, that reference linking is mediated through *the*. In instances where third person pronouns are both modified AND referential, the surface form that results more overtly resembles an appositive relative clause.

(43) **He** who must not be named walked into the room.

What differentiates the underlined portion in (43) from the restrictive pronominal relative clauses discussed in Chapter 2 is that in (43), *he* has moved all the way to D and does directly carry reference; in this way non-restrictive pronominal relative clauses are more comparable to (42c). The pattern here is that pronouns with modifiers can only appear in a D position if they are referential; if they are not, they will appear in a more noun-like position with respect to the modifiers.

This does beg the question: what about the cases where personal pronouns appear to be depronominalized Melchin (2015), as well as Cowper and Hall (2009), included examples like (44) below, where personal pronouns like *you* are in (what I have argued to be) the low, *n* position.

(44) the new you

What I assume is that, instead of person features being attached strictly to D, *reference* is attached to D; this is an easily-conflated distinction, but (45) can help to untangle it. In (45) the inclusion of a definite determiner *the* is only licensed if it can take a function as

an argument—the function in common NPs is the nominal predicate, while the function I proposed in Chapter 2 was a predicative form of pronominal gender. The possibility of (44) can be contrasted with (45) below, which is not possible except under the particular context that gives a construal where *the* combines with a set (and thus a function).

(45) a. *the you (i.e. *The you went to the store.)

b. Speaker A: I have a bunch of little figurines of everyone in the reading group, including myself.
Speaker B: I want the you!

In (45b), context necessarily sets up the possibility of picking *the you* out of a set of various others (*the me*, *the Karen*, *the Barbara*), which allows the personal pronoun *you* to be depronominalized without an explicit modifier; thus it is not apparently the nominal modifiers that are licensing depronominalization, but rather the non-referential (predicative!) nature of the depronominalized thing. This follows smoothly from the analysis that I gave in Chapter 2, but does suggest that person features need to enter the derivation lower than D; the fact that referential personal pronouns manifest person features is therefore a consequence of successful head-raising.

This allows us to revisit (42c), repeated here:

(42c) we happy few

In (42c) the meaning is referential, which is why it is distinct from (45) above. Revisiting the difference between (42c) and (42a,b), this means that an external determiner combined with a predicative pronoun can block head-raising even if that pronoun has 1^{st} or 2^{nd} person features. The meaningfulness of a depronominalized personal pronoun of the type shown in (45b) is still essentially predicative, but takes on a *de re* meaning (see Anand 2004 a.o.). Crucially, the threshold for my purposes is referentiality; depronominalized personal pronouns as in (45) can only 'refer' to an entity by way of picking the entity out of a set of other equally-feasible options. In *the new you*, the pronoun *you* is one among several *you*'s that might be referred to, which is why the article *the* is licensed.

The distinction I am making—that depronominalized pronouns are predicative because they are blocked from raising far enough to become referential—is not circular because they can be analyzed essentially in the same way that lexical nouns are (only differing in that they lack a lexical root as complement). The important takeaway is that not only can one block head-raising in pronouns, but one can detect blocked head-raising in the nominal domain by the position of the pronoun relative to modifiers (in the same way that verb-raising can be detected relative to adverbs). This blocking mechanism is one of the strongest pieces of syntactic evidence for head-raising, and its relationship with the semantics of the pronoun and reference are inextricably connected to the mechanism of blocking. In Section 5.4, when I explore the consequences of attributing edge linking power to the phase head D, I will return to the referentiality of the pronoun with respect to modifiers.

The next two sections deal with further predictions and consequences of the head-raising analysis, but both will focus much more on variation. Interspeaker variation (synchronic and historic) and intraspeaker variation (mediated by discursive forces such as stance, politeness, and conversational maxims) are both highly dependent on underspecification. For the purposes of these sections, I will be following Bjorkman (2017) and subsequent works (Konnelly and Cowper f.c.) in analyzing singular *they* as lacking gender features and number features; the radical underspecification of *they* in this analysis will yield surface ambiguity that can drive variation and change.

5.3.2 Support from syntax and pragmatics: number features and variable pronouns

The previous section focuses on what differentiates predicative pronouns from the other two types (referential and variable) that are part of my three-way proposal. In this section, I focus on variable pronouns as distinct from referential or variable ones. What sets predicative pronouns apart is their indication of acting more similar to a nominal predicate than a true "pronoun," and likewise what makes referential pronouns distinct is their direct link to the discourse context. Variable pronouns show some properties of

discourse-dependence for their reference (like referential ones), but they are not linked directly to a referent. In my proposal I have suggested that variable pronouns are a result of partial head movement to an intermediate projection, Num; this is similar in some ways to the proposal Déchaine and Wiltschko (2002) give for phi-P pronouns.

In order to motivate the intermediate projection Num as the target of head movement for bound variable pronouns, I first discuss their relationship with gender features, especially looking at ways in which the gender requirements are different for (bound) variable pronouns versus (free) referential pronouns. I then turn my attention to the Num(ber) head in particular in order to explain why the Numhead (and not another intermediate projection) should be the target of movement. I end this section with a discussion of the nature of the Number feature on *n* and Num, and how it fits into the wider assumptions about Agree operations in this analysis of head movement.

Among the questions brought up in an analysis of underspecification is how meaning is derived from a *lack* of features. Following Bjorkman (2017) in analyzing singular *they* as radically underspecified, I have in some ways set singular *they* up as a 'nothing' pronoun; Bjorkman's morphosyntactic analysis calls for its insertion under 'Elsewhere' conditions. I have argued thus far that head-raising to a referential D invokes the evaluation of felicity conditions; so how is felicity evaluated in terms of an 'Elsewhere' pronoun?

I have previously compared the use of gendered pronouns with the use of honorific or sociopragmatically-determined pronouns (in Romance, Japanese, and other languages); however languages maintaining a T/V distinction do not have an 'Elsewhere' form of address, one that is totally unspecified for social relationship—in systems where the honorific paradigm is a binary choice, neither of the choices can be called an 'Elsewhere' pronoun. German's inventory of *du* and *Sie* does not (to my knowledge) provide a third option where "neither familiar nor formal" can be invoked—yet the inclusion of singular *they* gives English speakers this type of choice.[8] *They* can only gain meaning through

[8]One possible exception to this is the retention of the *thee* and *thou* forms in certain literary registers of English well after the collapse of the T/V distinction in vernacular English. Because *you* in contemporary

pragmatic implicatures and inference, because it does not carry features even when it succeeds in head-raising to a referential D head. For the conservative speakers who report ungrammaticality or infelicity of singular referential *they*, metalinguistic commentary suggests that the problem is in their willingness to refer to a person without relying on gender—meaning that conservative (non-dsT) speakers are still using the forced binary choice of *he* and *she*.

For speakers whose felicity conditions do allow dsT referentially, the featurelessness of *they* is shown in the ways it can be exploited pragmatically. In Chapter 4 I showed that, depending on variable relative rankings of constraints of Relevance, Quantity, politeness constraints, and stance sensitivity, speakers may choose to use *they* to refer to an entity who does have a binary gender identity in certain contexts. The pronoun alternations shown in Chapter 4 also reinforce that the choice of 3^{rd} person singular pronoun (*he/she/they*) is much more of a pragmatic question than a syntactic one—this is true of referential pronouns but *not* bound ones (including anaphors). (46-47) show the contrast between a maybe-infelicitous free pronoun (46) and a definitely-unacceptable bound reflexive (47).

(46) RRA*: His partner at the time was also dating this other person that was in our group. Um, and* ***they*** *have a very, um,* ***he's*** *a very strong and kind of controlling personality, and so* ***he*** *had kind of taken over like the whole thing,* [. . .] *Ha. Yes. it kind of, that was kind of one of those things where it just- and that same person, I would see* ***them*** *more often than I would see [RRB] and* ***they*** *were trying to like convince me of these like negative things* [...]

(46) is an attested example from Experiment One (Chapter 3) where pronouns alter-

spoken English is the sole remaining $2^{n}d$ person pronoun, but *thee* is preserved in a literary register, the *thou/you* distinction in English becomes one of markedness. That is, *you* is the unmarked ("Elsewhere") form of address, while *thou* is marked. This has the paradoxical consequence of giving *thou* a "fancier" sound to it, since it is preserved primarily in works of literature considered to be high-prestige despite *thou*'s history as an INformal pronoun. It may be the case that investigating the sociopragmatic patterns around the collapse of the T/V alternation in English reveal similar patterns to the weakening of gender that I discuss in this dissertation.

nate between *he* and *they*; importantly, none of these are bound.

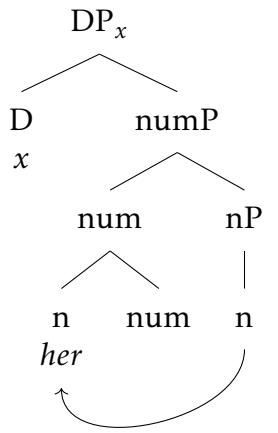
(47) *He$_i$ should behave themself$_i$. / *They$_i$ should behave himself$_i$.

The unacceptability of this alternation for a bound anaphor in(47) where it is possible for free pronouns in (46) suggests that there is a syntactic or semantic difference between anaphor feature valuation and feature evaluation of free pronouns. So far I have argued that free, referential pronouns are instances where a n has raised to a referential D and that felicity of the gender features is evaluated relative to the referent of the D head. If this were the process for anaphors as well, then alternations should be equally possible for pronouns as for anaphors—yet the intuitions on (47) are very strong, and not at all comparable to (46). The sociopragmatic sensitivity of free pronouns is, I have argued, a direct consequence of their internal syntax. Thus, the apparent *lack* of flexibility in anaphors should also be a consequence of their syntax.

I here revisit Déchaine and Wiltschko (2002): in their analysis, bound pronouns (including anaphors) were structurally smaller than free pronouns—for them, English reflexives qualify as ϕ-Ps, while free pronouns in English are full DPs. In my 'translation' of their structure, I suggest instead that pronouns start at n and raise to different levels: for predicative pronouns, pronouns simply do not raise, and stay in n; but for variable-bound pronouns (including anaphors) pronouns raise from n to an intermediate projection—NUM has the advantage of explaining number agreement—but do not raise to D, and combine with a null D that is a variable x. (49) below repeats the schematic representation I showed for this structure at the beginning of the chapter:

(48) Every senator$_x$ appreciates the people who voted for her$_x$.

(49) **Bound variable pronouns:**



In (49), the D is neither referential nor a determiner. Instead, I use the x variable placeholder to notate that anaphoric or bound Ds are *not yet* linked to referents or able to pick out individuals from sets on their own. Variable Ds are instead dependent upon their local binder to glean meaning. In the example I showed in (47) above, the anaphor was bound by a free (referential) pronoun. In this case, the anaphor (*themself* or *himself*) would not be evaluated for felicity until after the antecedent *he* or *they* was merged in a local configuration. The advantage of the intermediate raising analysis for variable pronouns is that it correctly derives the semantic effects and captures Déchaine and Wiltschko's insight that variable pronouns do not behave like either nouns or Ds. Variable pronouns are distinct from predicative pronouns in that they are generally not modifiable, and that they participate in a certain kind of coreference (binding) rather than predicative pronouns which cannot participate in binding/coreference.

It is also worth noting that the number morphology of predicative pronouns is irregular and does not behave at all like the expected number morphology found on either variable or referential pronouns. First, they do not directly control verb agreement; secondly, predicative pronouns appear to be pluralized more like regular nouns (which is not the case for variable pronouns). (50a) shows the predicative use of singular *they* does not cause plural verb agreement, while (50b-c) show that both variable and referential singular *they* do cause plural verb agreement.

(50) a. The other they **is/*are** the one you're looking for.

b. Any linguist$_i$ should know that they$_i$ ***is/are** a member of a society.

c. Kirby$_i$ knows that they$_i$ ***is/are** doing their best.

The evidence in (50) supports the differentiation between predicative and variable pronouns generally. This also shows why predicative pronouns apparently do not control number agreement, while variable and referential ones do—movement to or through Numis how the pronoun *they* in (50) accumulates (syntactic) plural marking, even though all senses of *they* are semantically singular.

This proposal is comparable to many previous analyses of anaphor ϕ-valuation (cf Kratzer 2009 i.a.). The difference is that rather than *valuing* the gender or number features, in my analysis the *x* operator is an edge linker that tests for a *match*, where reference is calculated external to the DP of the anaphor itself. In a case like (51) below, reference for *themself* is dependent upon reference of *Jayden*, which means that felicity and feature matching can't be evaluated until *Jayden* enters the derivation and, as an R-expression, brings along a referential index.

(51) a. *He likes themself.

b. Jayden likes themself.

Because the conditions on *themself* are only evaluated with relation to the antecedent *Jayden*, the threshold for feature matching is much stricter in that the anaphor must have the exact same feature set as its antecedent. Thus, *themself* is only licensed with an antecedent that has no gender features. In the case of referential proper names as antecedents, this is dependent on the speaker's understanding of the referent's gender (or the social appropriateness with regard to the referent).

In the case of a free pronoun anteceding an anaphor as in (51a), however, the free pronoun already introduces gender features that must then be exactly matched by the anaphor. The fact that *themself* is unspecified for gender features means that the match is not exact.

The discussion in this section thusfar has been focused on the differences between treatments of gender and number features between variable pronouns and other types. However the descriptive patterns that I show here do not necessarily solve the question of where formal features appear in the syntactic structure. In the proposal I gave (4) at the beginning of this chapter shows that Num carries a iNum feature, and *n* carries an uNum feature. This is the opposite pattern of that found in the gender features: (4) shows D (the higher head) with uGender and *n* with iGender.

The reason that I propose that *n* pronouns carry interpretable gender but do not carry interpretable number is that the number in predicative pronouns is not determined by the form of the pronoun, but rather is determined by an independent Num element. (52) shows that a "singular" pronoun can be coerced to show regularized plural marking (suffixed *-s*); furthermore, even *them* can apparently be doubly marked as plural when it is a predicative pronoun preceded by a numeral.

(52) a. The two shes that I was talking about **are/*is** ...

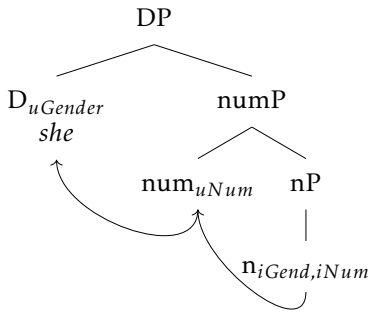
b. the two thems that I was talking about

The data in (52) suggest that the Num head is in fact the locus of interpretable number, not *n*. This poses a slight problem for the internal parsimony if I try to adopt Matushansky's conception of head movement exactly as-is. In that proposal, movement is driven exclusively by downward Agree, which is triggered by the merge of an uninterpretable feature to serve as a probe. In the case of Num features, if uNum is merged *before* iNum, then Agree probing downward will not find a goal and uNum will not be valued.

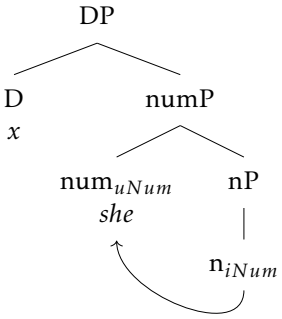
The possible solutions to this conflict are that either Agree can operate bidirectionally (meaning the probe can merge lower or higher than a goal and still be able to establish an Agree relationship that can drive movement); or else there is an additional unidentified feature on Num that probes *n* independently. The third possibility is that interpretability of the Num feature on Num is linked to the difference between triggering head raising (as will happen in variable and referential pronous) versus triggering concord (as happens in

the hyperregularized plurals in (52)). This would mean that uNum on Num triggers head raising, while an iNum feature on overt numeral Num heads instead results in concord. If this is the case, then the three structures would have slightly different arrangements of features, shown in a revision of (9) here.

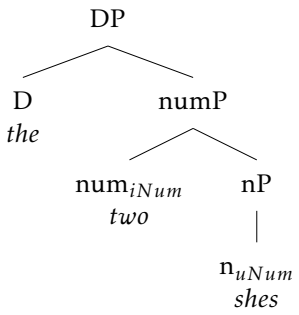
(9') a. Referential pronouns:



b. Bound variables:



c. Predicative prns:



In the structure I show in (9'c), the predicative pronoun *she* can spell out gender features independently of number features (*two*). This tentative solution correctly predicts that predicative pronouns alone should show concord (instead of raising); however this solution may be overfitted to English. Future work may uncover clearer diagnostic tests for distinguishing between the logical possibilities that I have listed here.

In this section so far, I have shown evidence for differentiating between predicative and variable pronouns on the grounds of number agreement, which supports my proposal for predicative pronouns remaining in *n* and variable pronouns raising to (at least) Num. However, it is also an open question whether variable pronouns should move even higher to D. If variable pronouns did not land at Num but did continue to D, this might suggest that bound variable pronouns would potentially be subject to the same gender requirements as referential pronouns—this would depend on whether a variable D carried a strong uGender feature. If variable D did carry uGender, this would imply that singular *they* would be equally ungrammatical for non-dsT speakers as its referential counterpart. This does not appear to be the case: bound variable singular *they* was not rated lower relative to bound variable *he* or *she*, as I showed in Chapter 3.

It is also not apparent that a uGender feature would need to be valued on variable D even if variable pronouns did not raise to D, since singular *they* is widely accepted in its variable use and has been in regular use since at least the Middle English period (Curzan 2003 i.a.). This diagnostic is therefore ultimately not a definitive way to show whether

variable pronouns do or do not raise to D. It is worth noting that bound anaphora can be apparently bound by null arguments, as in (53) below—and these null arguments apparently will license featureless pronouns like *oneself*, generic $2^{n}d$ person pronoun *yourself*, but NOT an obviously-gendered *herself* (and marked "literary"/informal *thyself* is also marginal to my ear).

(53) a. to love oneself is divine
b. to love yourself is divine
c. * to love herself is divine
d. ? to love thyself is divine

What this suggests is that gender features in bound positions must be licensed relative to an actual referent, and a null argument does not sufficiently license the gender (which shows that the constraint demonstrated in (51) goes both ways). This licensing requirement may also be present for socially-mediated features like honorifics. Future work investigating this matter may probe cross-linguistic patterns in languages with more overt politeness or honorific marking in the pronominal system. For the purposes of this analysis I will continue to show variable pronouns only raising as high as Num.

This section has shown more evidence for the validity of the three subtypes of pronouns that I propose, as well as clarity on why head movement correctly predicts the different patterns of behavior for each type. In this section I have focused primarily on syntactic evidence, especially around number and gender matching and agreement phenomena. In the next section I focus more on the advantage of the head movement analysis as a parameterizable structure, which correctly predicts synchronic variation and gives insight into what diachronic development may have looked like.

5.3.3 Support from sociopragmatic variation: parameterization

In this section I propose that a single parameter—the presence of an uninterpretable Gender feature on referential D—can account for the existing variation in speakers who judge

and use singular *they* differently. I focus primarily on synchronic variation because the data in Chapters 3 and 4 are synchronic; at the end of this section I will suggest how the parameter I proposed can potentially give insight into possible paths for diachronic development which may have led to the current state of affairs. In (4) I detailed two different feature sets responsible for head movement in referential pronouns, where the difference between these feature sets represents the different parameter settings for grammars that do or do not include dsT.

For speakers who have uGender on referential D heads, singular *they* fails *only* when the pronoun moves all the way to D — that is, singular *they* is only ungrammatical if it is referential. This feature robustly explains metalinguistic comments about the grammaticality of singular *they* anteceded by a proper name, which I showed in Chapter 3 was notably different from other antecedents. If a referential D has a uGender feature that can only be valued by head movement, and *they* is the pronoun that moves to D, then there is no gender to value uGender and the feature is unvalued when the DP phase proceeds to Spell-Out — this causes a crash. The crash does not occur for epicene or generic singular *they* because either the D head is definite (not referential) and therefore doesn't have or need uGender, or because the D head is a variable *x* and likewise doesn't have or need uGender. Additionally, in my analysis so far I have suggested that variable pronouns do not raise as far as D anyways. Thus, speakers will readily produce singular *they* in generic or variable contexts but report serious ungrammaticality in referential ones; this is a natural consequence of the proposal that I have given so far.

The apparent link between referentiality and gender obligatoriness suggests that for the non-dsT speakers, (social) gender and reference are intrinsically linked (which is quite in keeping with Pullum's (2003) insights). However, for speakers who do accept and produce dsT (even when anteceded by proper names), it must be possible for singular *they* to move to a referential D without causing a crash. What this suggests is that either these speakers are somehow valuing uGender in some other way or, as I propose here, uGender is absent and a different feature is responsible for movement to D. The total

absence of uGender means that the derivation will not crash due to uGender remaining unchecked, which explains the use of dsT to refer to nonbinary referents, and which gives this analysis an advantage over alternative explanations.

Given the possibility of dsT, referentiality and gender cannot be a single unified feature, or else the loss of uGender would cause a total collapse of pronominal reference for dsT-users. It is clearly the case that dsT can be used referentially by many speakers, however, which means that an independent feature (which I have called Ref/Reference) must still be present to trigger head movement and mediate context evaluation. The independence of Gender and Reference also allows for a more cross-linguistically sound featural system which can account for the many languages with no pronominal gender features whatsoever (Corbett 2006). However, it does seem to be the case that dsT-speakers still retain and use *he* and *she*—thus gender marking must be possible but non-obligatory. I adopt Bjorkman's (2017) analysis of the featural innovation around singular *they* to explain how dsT-users can include gender but not suffer a crash when it is unvalued.

As I reviewed in Chapter 1, Bjorkman (2017) proposes that singular *they* lacks gender features completely; this is made possible by a different gender featural system for the dsT-dialect, where *masculine* and *feminine* are optional extra ("adjunct") features rather than possible values of an obligatory formal Gender feature. This is the most compatible analysis with the deployments in discourse-context I showed in Chapter 4, and most readily explaines speakers' strong preference for *they* in contexts where gender isn't otherwise necessary (including epicene and generic uses). The extension of Bjorkman's proposal for singular *they* is that gender features are (newly) optional in English, and that lack of gender is no longer sufficient to cause a crash. The optionality of gender features generally is compatible with my proposal here that the parameter rests on whether or not speakers need a uGender feature valued. For speakers who don't need to value uGender (because they lack the formal feature on referential D), gender features become non-contrastive and are included primarily to serve social and discursive goals.

To return to the link between gender and referentiality, I note that even for non-dsT

speakers, gender does not need to be valued for *all* pronouns. This parameter is specific to referential pronouns because Chapter 3 showed a notable difference in ratings of singular *they* with proper names, but not for bound variables or generic antecedents. The narrowness of this parametric difference leads to a significant amount of overlap in the use of singular *they* by speakers of the two different dialects. Speakers who do not accept dsT will still readily produce or accept epicene singular *they*, and there are definite epicene antecedents that result in uses of singular *they* that are ambiguous between an epicene meaning and a referential one. This ambiguity can be reproduced only with sufficient context for both construals:

(54) **Utterance**: My math teacher always gives me a lower grade for doodling, they are so unfair!

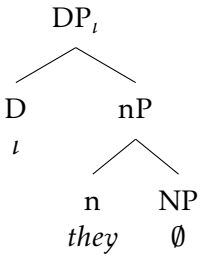
No-dsT meaning: *No matter what math teacher I get, each math teacher grades me down for doodling – they (epicene) are unfair!*

Possible-dsT meaning: *This quarter my math teacher (whom you've never met, and don't know the gender of) graded me down every time I doodled – they (specific) are unfair!*

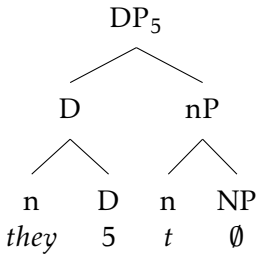
The dsT meaning in (54) differs from the non-dsT meaning not in syntactic number, or even in gender information, but in specificity of reference. The more conservative dialect's meaning, an epicene, varies across a set of math teachers, while the more innovative dialect's meaning can refer to a particular math teacher directly. The underlying difference between these two meanings is also, under the analysis I have proposed, structural: in the epicene meaning, *they* has not raised to D to combine with a referring edge linker, while in the construed meaning, it has. I use ι to represent a definite determiner selecting from a set (following Elbourne 2013) and numeric indices to represent reference to particular individuals without picking from a set in formalizing (54) below.

(55)

a. Epicene with definite antecedent



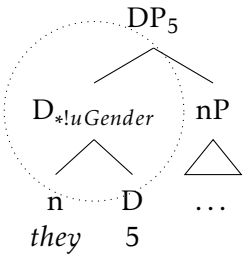
b. Referential pronoun



The structural difference shown in (55) is not immediately apparent based on the surface string, and it is quite possible for two speakers of different parameter settings to share in a conversation without running into significant confusion despite their slightly different interpretations. The surface ambiguity that masks a structural difference only becomes clear with additional context (as is given in (54)). For non-dsT speakers, however, the construal in (55b) is either highly unlikely or indeed impossible, because the uGender feature on D *must* be valued. An alternate structure given in (56) below shows the non-dsT structure that results in a crash.

(56) Referential pronouns for non-dsT speakers

No value for uGender causes crash



The crash due to the unvalued uGender feature in (56) shows why speakers who are able to accept singular *they* in variable contexts will reject or misunderstand referential uses.

Another distinction remains, however—gender-nonspecific dsT is still not equivalent to non-binary use of dsT. For speakers who *do* accept a referential reading in (54)/(55b), it is still not necessarily the case that non-binary contexts will be within their repertoire.

The construal in (54)/(55b) is referential and without gender, but the lack of gender features can only give a non-binary reading when supported by the discourse context; this is because dsT does not constitute a third gender for English pronouns, but rather is underspecified and only gives meanings through pragmatic implicatures. For dsT speakers to accept or produce dsT referentially in gender-nonspecific contexts like (54), they only need to lack uGender on referential D in their grammar; but crucially an additional social factor is necessary for non-binary uses. Namely, speakers must not only lack uGender on D, but they also must internally accept the possibility that a person can be neither a man nor a woman—this is the difference between a gender-nonspecific reading and a non-binary reading. However, that difference is purely a pragmatic question, as both uses are referential.

Due to this social/pragmatic difference, speakers who use dsT for gender-irrelevant purposes may still reject dsT anteceded by a proper name—however, this is mediated by social evaluation of appropriateness, rather than a problem with the grammar itself. In Chapter 4 I showed how dsT can be produced by differently-ranked sociopragmatic constraints which give rise to implicatures of either gender-irrelevance or gender-nonbinaryness. The two can be de-conflated through additional context to show the difference between dsT users who do and do not use dsT for non-binary referents.

(57) *My math teacher . . .* ***They*** *are unfair*

Referential (but not nonbinary) meaning: The math teacher I have this year, whose gender you don't know and it's not really relevant – they (gender vague and irrelevant) are unfair!

Referential and nonbinary meaning: The math teacher I have this year, whose class you are also taking, and whom we know can't appropriately be referred to as *he* or *she* – they (gender-neutral on purpose) are unfair

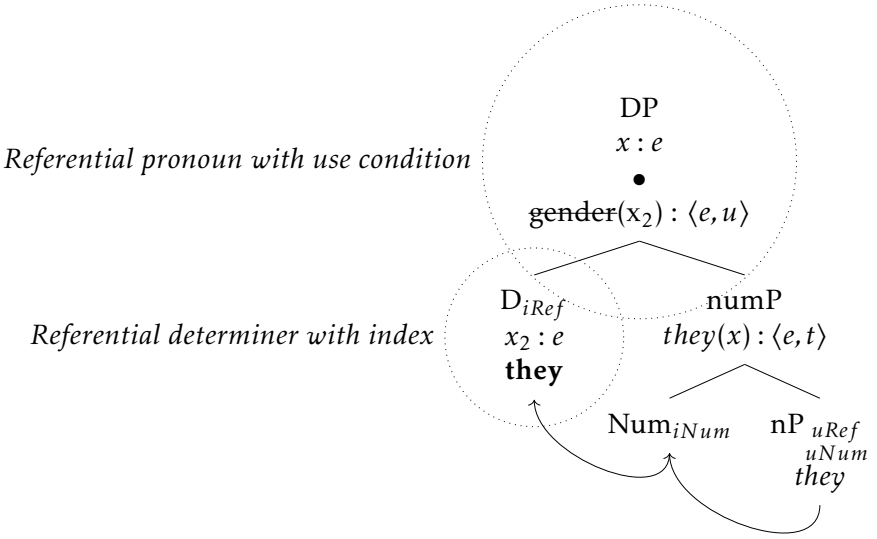
The difference between the interpretations in (57) is in whether the speaker assumes the hearer to be sufficiently socially familiar with the referent in order to evaluate the felici-

tousness. This also explains why non-dsT users can more easily identify the ungrammaticality of dsT when anteceded by a proper name, since non-proper DPs may be ambiguous between referential descriptions and definite descriptions.[9]

The evaluation of social appropriateness with respect to a referent, as I showed in (4a), occurs upon the convergence of the entire DP so that a referent is directly linked to the pronoun. (58) below shows the resulting structure of referential dsT having moved to D, including the use conditions from Gutzmann and McCready (2014) extended to include dsT, in order to show why the referent-dependent sociopragmatic sensitivity appears primarily in referential pronouns.

(58) **Use conditional structure of dsT**

a.



[9]Because proper names (when they, too, have raised to D; cf Matushansky 2015) not only pick out a specific referent but also assume a familiarity with at least the naming convention associated with the referent, it is more likely that hearers will hear a proper name and assume that they are supposed to know some social information about the referent. In fact, this is a guiding principle in conventions about newspaper headlines and lede-writing:

i. ***Who (and what) is whom (or what)?*** *Make it clear: Don't use proper names in headlines unless the name is well-known enough to be recognized immediately. The same is true for abbreviations.*
Source: https://web.ku.edu/~edit/heads.html

Languages can in fact conventionalize the difference between a definite and referential use with different pronouns, as well; Japanese *kare/kanojo* are only felicitous when the speaker and addressee share enough context to identify the referent directly (pointed out by Edith Aldridge, p.c.). This may suggest that referential indexes and definite determiners (ι) are available for separate grammatical uses.

b. **Paraphrase**: an entity x such that x is the entity indexed as 2 AND such that it is appropriate to refer to 2 without gender features

In the paraphrase I give in (58b), the use condition is evaluated not based on whether it is appropriate to refer to x_2 as *they*, but rather whether it is appropriate to refer to x_2 without gender features; this is a reflection of the fact that it is possible to omit gender features for a variety of reasons, and the use condition value will be contextually determined.

For this reason, the use of a proper name in asking participants for acceptability judgments does illuminate the difference in readings between a discourse-driven, gender-nonspecific *they* and a referential, socially-supported (nonbinary) *they* in (57) above. It is also for this reason that metalinguistic commenters, including trained linguists, have pointed out that proper names render singular *they* unacceptable despite accepting its use in every other situation. Definite epicene antecedents are still picking individuals out of sets and—which I distinguish from referential DP antecedents, essentially R-expressions (including proper names and referential descriptions). I argued that the completion of head-raising to a referential D is what initiates the evaluation of appropriateness with respect to a particular referent. This means that for speakers who *do* have dsT in their grammar, the use conditions do allow underspecification for any given referent, while speakers who *do not* have dsT in their grammar only lack singular *they* when it is referential—thus meaning these speakers' felicity conditions prevent felicitous head-raising to a referential D. For these speakers, it is simply not possible to refer to an animate, human referent without gendering them. I revisit Pullum's quote:

> "*I would now say that although* *Chris left their pen *still sounds dreadful for some reason (perhaps because* ***whoever Chris is, he or she really does have a gender****), nonetheless it is possible to have a singular they with a singular proper name antecedent*." [emphasis added] (Pullum 2003)

Pullum's report places him in the group lacking dsT—which is linked to his inability to

felicitously refer to a person without recourse to *he* or *she*. This is additional support that not only is referential D a particular locus of evaluation, but that the presence or absence of gender (and, relatedly, the uGender feature) is a primary deciding factor in who accepts and uses dsT.

The synchronic variation shown in this dissertation—including both the grammaticality of dsT, as well as its different uses—can be explained by the presence or absence of uGender on referential D. This parameter correctly predicts that reports of ungrammaticality will hinge specifically on referential uses of dsT when gender is 'expected' (for conservative speakers who need uGender valued), since *they* lacks gender features and leaves uGender unchecked, leading to a crash. The change in apparent time reported in Chapter 3 also implies a possible shift in the grammar since the 1970s (based on the birth years of participants who were more likely to rate dsT lower), but I have not presented data to confirm how or precisely when this change may have taken place. Diachronic examples of referential pronouns are exceedingly difficult to collect and analyze in great volume, since automatic coreference resolution technology is not accurate enough to detect very low-frequency variation such as nonbinary and trans uses of dsT. Future studies looking to confirm the grammatical change around dsT through corpus investigation will need to be hand-annotated for coreference, and careful attention to pragmatic context will be necessary for differentiating specific uses from epicene definite uses. Because of this, any conjecture I provide on the diachrony of dsT based on this proposal is based primarily on extrapolation from synchronic data.

However, I note here two trends in the history of English pronouns that may be of interest in further study of the history of dsT and the emergence of its nonbinary use. The first is that Curzan (2003) identifies that the shift from pronouns agreeing with grammatical gender towards instead agreeing with 'natural' gender occurs around the transition period from Old English to Middle English, in concert with the more general loss of grammatical gender marking on nouns. Secondly, Van Gelderen (2013) discusses the diachronic shifts in the English pronominal system around the introduction of *she* and

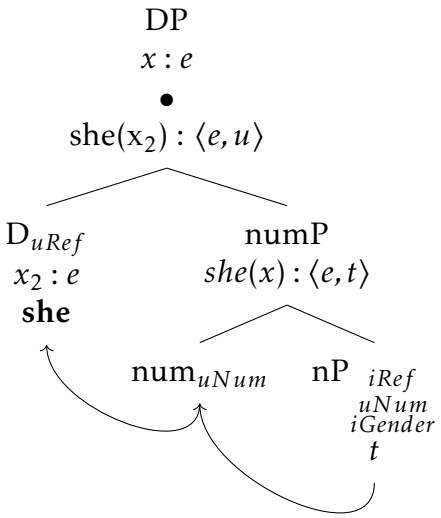
they, dated around the same period. It may be the case that the emergence of dsT is a continuation of a general shift away from gender features altogether, given that the resulting grammar renders gender non-obligatory on the only remaining locus of gender features in English—pronouns. Future investigations can therefore follow the loss of gender morphology and the expansion of the domain of singular *they* as potentially related.

In this section I have showed how a single parameter (uGender on D) can robustly explain synchronic variation, including both interspeaker variation (via the presence or absence of uGender) and intraspeaker variation (via the pragmatic implicatures that arise in the absence of gender features). I have also suggested a possible route for further investigation into the historical development of dsT. In the final section of this chapter I will detail the direct consequences of the head raising proposal, with attention to how pronouns, phases, and head movement are affected.

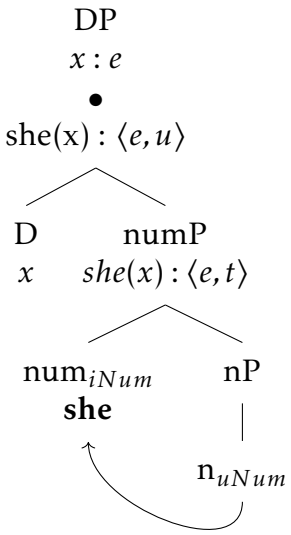
5.4 Consequences and Predictions

The structures that I have proposed result in (at least) three identifiable subtypes of pronouns, which are differentiated not only by their syntactic properties but also by the way that they interface with pragmatic and social components. I repeat the structures here for reference.

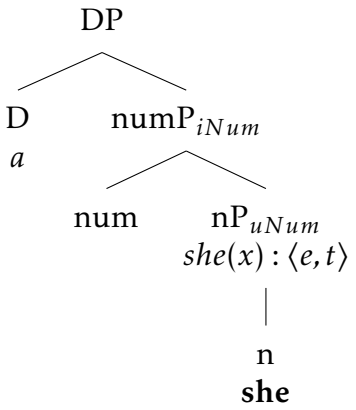
(4) a. Referential pronoun



b. Variable pronoun



c. Predicative pronoun



The structures in (4) differentiate between referential, bound variable, and predicative pronouns as a consequence of different landing sites for head movement. The semantic type and pragmatic use conditions are composed along with syntactic structure, which as a consequence means that referential pronouns are subject to sociopragmatic evaluation relative to a particular referent while variable and predicative pronouns are evaluated primarily on semantic grounds.

This proposal incorporates a particular view of head movement where the *movement* itself occurs in the narrow syntax, and does effect the semantic/pragmatic composition and interpretation of the structure. I also draw heavily from the concept of use conditional semantics, particularly as a way of explaining why using a certain gendered pronoun may be considered *inappropriate* for a particular referent as a separate consideration from whether the gender feature is "true" or not. I also give a phase based explanation for why use conditions for referential pronouns are sensitive to the particular referent, and why these conditions don't appear to be evaluated in the same way for variable or predicative pronouns. In this section I discuss the implications of this proposal, including how its core components constrain one another. I first address the implications for pronouns (generally), then for phases and phasehood, and finally for head movement.

5.4.1 Predictions for pronouns

As I have previously alluded to, this proposal flattens the distinction between social gender features and honorific features. In some ways, the direct formalization of the syntax/pragmatics interface (through use conditions) makes these highly-social, dynamic features much more accessible to operations in the narrow syntax than previous analyses of gender on pronouns. This is not without precedent: analyses of Japanese honorific morphological marking on verbs have taken honorification to be a feature accessible in the syntax and subject to syntactic processes such as Agree (Niinuma 2003, Boeckx and Niinuma 2004, Boeckx 2006, among many others). Since it is generally the case that mostly-pragmatic features (like honorific marking) are available for strictly-syntactic processes like Agree, I extend this to cover social gender features.

However, my proposal in some ways strengthens the presence of social gender in the syntax, and in other ways weakens the strength of gender agreement (when based upon social features) as an aspect of grammatical structure. Where previous analyses even of singular *they* (such as Bjorkman 2017) have suggested that proper names or gendered lexical nouns carry the same type of gender features as pronouns, I instead claim that ONLY pronouns in English carry these features–and that the features are optional even on pronouns (for speakers with dsT). This claim more thoroughly explains the kinds of gendered pronoun alternations that I show in Chapter 4, especially the observation that local binding between a pronoun and an anaphor is structurally subject to different agreement requirements than pragmatic coreference between a free pronoun and an antecedent. This apparent asymmetry in the 'strength' of agreement is readily explained by my analysis where variable pronouns are structurally dependent on binders, while referential pronouns' gender features are evaluated socially (regardless of whether there is a linguistic antecedent vs. a pragmatic one).

Generally speaking, my analysis predicts that pronouns reflecting social (and not grammatical) gender should be subject to the same or very similar constraints as the ones

I showed for English. Specifically, referential free pronouns should show more flexibility with gender features, such that they can be influenced by stance, affect, and politeness considerations. These pragmatic considerations will be culturally specific, and future work investigating alternations in these pronouns should include careful sociolinguistic and ethnographic investigation into how particular speakers and communities decide what consitutes an "appropriate" use of a gendered pronoun.

Because some of English's flexiblity is specifically enabled by the availability of an 'elsewhere' pronoun (singular *they*), this analysis also predicts that certain pragmatic constraints will be more obvious in languages with a widely-accepted gender neutral singular pronoun in addition to however many gendered pronouns. Swedish and Spanish are both ripe areas for further study–both are in the process of introducing gender neutral pronouns, though through apparently very different social and linguistic mechanisms. Swedish gender-neutral *hen*, borrowed from Finnish *hän*, may therefore show very different sociolinguistic uses and diffusion than Spanish gender-neutral *elle*, which is a morphological innovation based on combining *ella* and *él*. Gustafsson Sendén, Bäck, and Lindqvist (2015) discusse *hen* in Swedish, and Papadopoulos (2019) investigates Spanish-speaking gender marking strategies outside the binary(including pronoun innovations).

Additionally, my proposal for how singular *they* is developing in English partially rests on the fact that English does not have grammatical gender; for languages that DO have grammatical gender, it may be the case that pronominal innovation is dependent on (or indicative of) a much larger shift in the overall system. If it is the case that dsT is a natural consequence of the loss of grammatical gender from OE, this begs the question of whether Spanish, for example, must lose grammatical gender in order to take up the same gender-optional status.

Finally, since the point of social sensitivity I propose is at the D phase head, this analysis predicts that pronouns should be cross-linguistically socially sensitive with respect to a specific referent ONLY when they show evidence of being full DPs and/or moving all the way to D through head movement. Honorific pronouns, for example, should apply

honorific properties to the referent only when they are referential (rather than variable or predicative); this also predicts that deictic personal pronouns (1st or 2nd person singular, especially) should be cross-linguistically more likely to have honorific properties.

5.4.2 Predictions for phases

The proposal in this chapter also depends on a crucial contribution from Sigurðsson (2018): that the locus of evaluation for appropriateness in context must occur at the edge of a phase. Sigurðsson draws a parallel between the evaluation of gender at D with the evaluation of utterance-relative tense at C; if it is the case that phase heads are generally the point at which context is available to the syntax, then it should also be the case that phase heads/edges should be the locus of evaluating social appropriateness as well. The structural explanation that Sigurðsson gives for this is based upon the presence of edge linkers at phase edges; in my analysis I additionally connect this context sensitivity with the timing of Spell-Out.

It seems to be a reasonable generalization to say that phase edges are socially and pragmatically sensitive; if, for example, the left periphery around C includes Force, then the matter of "when is it appropriate to phrase something as a question or a statement" is more readily subject to pragmatic considerations. Certainly it is the case that questions and statements can have the same or very similar truth conditions, but very different use conditions—evaluating this at Force/C is a good way of formalizing that distinction. It is also worth noting that head movement in the clausal domain (for example, T to C in English questions) also is connected to that same discourse-sensitivity.

If this generalization holds for all phases, then sociopragmatic sensitivity may be a useful diagnostic for phasehood. If a particular head is proposed to be a phase head (in, for example, every-phrase-a-phase proposals) then pragmatic sensitivity can be detected by defeasibility, attested synchronic variation, and context manipulation. This generalization also appears to be support for the explanatory power of phases more generally.

5.4.3 Predictions for head movement

This proposal relies heavily on the mechanism of head movement as a vehicle for differentiating different apparent category-linked behavior within a single domain without resorting to category-switching. While such a differentiation may be more transparent in 'larger' domains (like the verbal), pronouns are relatively speaking microscopic, and detecting head movement is based partly on syntactic markers and partly on inference from semantic/pragmatic effects. Firstly, the semantic/pragmatic effects are strong evidence that head movement can and should take place at least partly in the narrow syntax.

Secondly, and related to the previous section, there is a significant difference between head movement that proceeds all the way to a phase head, and movement that does not. My analysis strongly predicts that when head movement makes it as far up the functional hierarchy as a phase head, the head-complex is enabled to interface with pragmatic context more readily. This is an extension of Sigurðsson's analysis, but my differentiation between different types of D heads also suggests that other phase heads may have different instantiations that mediate the relationship between the discourse and the syntax. In investigating the predictions further, linguists may turn to to the clausal or verbal domains to investigate how and when C and *v* (or any other purported phase edges) are a landing site for head raising, and whether this correlates directly with discourse-sensitivity. In languages where T raises to C, for example, the instances of T-to-C raising should (if they are parallel) be directly affected by discourse context. As mentioned above, the presence of T to C in questions may be an example of this.

Chapter 6

CONCLUSION

In this thesis, I incorporated syntactic, sociolinguistic, and pragmatic data into an analysis of English third person singular pronouns, with particular attention to how 'natural' gender is conventionalized in functional categories in the grammar. I argued that pronominal relative clauses and depronominalizations (1) provide evidence that English pronouns must optionally appear low in the nominal domain, rather than exclusively in D.

(1) a. he who is without sin
b. That person is a she.

In Chapter 2 I compared predicative pronouns to predicative versions of proper names as in (2), and proposed that the categories of names and pronouns are lexical and functional counterparts; this is reflected in the structure by presence or absence of a $\sqrt{}$ (root) complement to a functional head n (3).

(2) a. the two Brents that live in Seattle
b. the two shes that I was talking about

(3) a. $[_{DP}$ the $[_{NumP}$ two $[_{nP}$ n^0 [$\sqrt{(\text{"BRENT"})}$]]]]
b. $[_{DP}$ the $[_{NumP}$ two $[_{nP}$ *she*]]]

One of the major contributions of Chapter 2 is a direct syntactic analysis of names and pronouns that accounts for the structural differences between their referential and predicative uses, as well as a clearer understanding of the semantic role played by the n head. The analysis of n as a functional projection that makes a root into a 'name' parallels analyses of categorizing projections in other domains (such as v). Chapter 2 also solves one of

the ongoing paradoxes of pronominal category: pronouns that are used as predicates are not lexicalized, but rather are still functional elements that are serving a different syntactic/semantic role. The incorporation of the *n* head as the origin point for pronouns also relieves much of the tension of debates between D-like and N-like attributes, intra- and cross-linguistically. In Chapter 5 I went on to show how a movement analysis can further capture the nuances of different pronoun behaviors.

In Chapter 3 I turned to focus on definite, specific singular *they*, taking data from interviews and acceptability ratings to show that there is a difference in ratings of *they* with different antecedents depending upon speaker age. Specifically, singular *they* showed lower ratings by older participants when it was paired with proper names (regardless of the common gender of the name). Older speakers also used *they* for specific (singular) referents less than younger speakers did.

In analyzing the influence of social factors on the use and acceptance of dsT, I proposed that the age-related variability suggests a change over time, based on the Apparent Time Hypothesis. The age variability had a noticeable cut-off point where speakers born before 1983 and after rated singular they significantly lower when used with a proper name. While this study is suggestive of a change in apparent time, future longitudinal studies (modeled after Sankoff and Blondeau 2007) should be able to show more conclusively what variability is due to grammatical shift in the population, and what variability is explained by individual speakers shifting their behaviors over time. Based on metalinguistic comments collected in Experiment Two, dsT is a salient variable and speakers have the sense that its use is changing over time, though some expressed difficulty in keeping up with that change. I predict that a longitudinal panel study of dsT will find similar results to those of Sankoff and Blondeau (2007): speakers who already rate dsT at either the floor or ceiling are likely to show similar behaviors at different data collection times, while speakers who rated dsT somewhere in between (neither consistently low nor consistently high) may rate its use differently at a later data collection time. Such a study would be informative both as a validation of the ATH even around individual variation,

but also would provide important support for the applicability of the same models of change for both sound-based and grammar-based variables.

The other important conclusion in Chapter 3 was the relation between use/acceptance of dsT and proximity to nonbinary genders; the data from both studies in the chapter suggested that nonbinary speakers use dsT more frequently, and rate it more highly. Metalinguistic comments also indicated that, aside from a speaker's own nonbinary identity, social proximity or membership in a wider LGBT+ community was a factor that speakers identified as contributing to their higher ratings of dsT. I analyzed these tendencies through the lens of orders of indexicality, where indexical meaning of the use of dsT could be abstracted from direct nonbinary identity to progressively further-removed associations or allegiances around that identity. The findings in this chapter indicate that LGBT+ identification (or speaker identity) is a meaningful social category that in fact can contribute to syntactic change (where previous research has mostly focused on phonetic or lexical variation).

Additionally, I extended the model of indexical orders to show how syntactic or pragmatic sociolinguistic variables must have, 'underneath' any abstracted indexical meaning, a direct denotative meaning that is produceable within the confines of the grammar, and which will directly contribute to the emergence of the first indexical order as the variant gains its initial connotative meaning. This expansion of how indexical orders can be conceptualized is only possible by examining syntactic/pragmatic variables (of the size of at least a morpheme)—such an inference cannot be made about sound-based variables in that sub-morphemic elements definitionally do not have direct, denotative meanings. The parallel, however, still exists between sub-morphemic and super-morphemic units in the emergence of first-order indexical meaning. Namely, either type of change must be a natural possibility within the structure of the language; in Chapter 5 I revisited this 'inherent possibility' as a necessary element in deducing what diachronic changes could have lead to the emergence of dsT in the first place.

In Chapter 4 I directly investigated the logical possibilities resulting from the avail-

ability of dsT in the grammar, as well as investigating related phenomena where the choice of third person singular pronouns is largely pragmatic. In order to explain the contextual dependency of pronoun choice, I synthesized a model of pragmatics where different pragmatic principles—including those informed by Politeness Theory, Gricean maxims, and stance relations—can be formulated as constraints that apply generally and must be ranked relative to one another in order to resolve conflicts. This constraint-ranking model (OT-like but without the related theoretical assumptions) was robust in not only explaining attested pronoun alternations (between *they* and other pronouns) but also in predicting possible behaviors (including alternations between *he* and .

The constraints I proposed were slight reconfigurations of existing pragmatic principles. These principles (and their sources) can be summarized by each constraint by Table 6.1.

Quantity (Grice 1968)	*Maximize Quantity!* *Minimize Quantity!*
Quality (Grice 1968)	*Don't lie!* *Don't guess!*
Saving referent's positive face (P. Brown and Levinson 1987)	*Don't misgender!* *Don't ungender!*
Saving speaker's negative face (P. Brown and Levinson 1987)	*Don't correct me!*

Table 6.1: Pragmatic constraints influencing gender of pronouns

In order to incorporate stance relations into the OT-like pragmatic model, I used DuBois' 2007 configuration of the Stance Triangle to track the three-way relationship between speaker, addressee, and referent that necessarily arises around third person reference. DuBois' model conceives of any stance-taking act as consisting of a Stance Subject, Stance Object, and interlocutor; the Stance Subject relates to Stance Object by positioning themself in relation to the object, and evaluating the object, and the Stance Subject relates to their interlocutor by aligning themself relative to the interlocutor's own stance relation with the (shared) object. In the case of gendered third person pronouns, the 'con-

straints' that arise are necessarily contextual and specific to the speakers and situation. These constraints will only be in conflict if it is the case that two interlocutors have different assumptions or ideas about the appropriate pronoun to use about a single referent—I gave an example in Chapter 4 to illustrate this conflict, where a trans person may be 'out' to her friends but not to her parents. In this instance, a friend will have the conflicting constraints of their own evaluation of the referent's gender, and the pull to try and align with or anticipate the stance of the interlocutor (parent).

In this very contextually-specific situation, contextually-specific constraints compete with the more general constraints at play. For the example of the trans woman in the closet, the constraints in competition are these:

(4) Don't misgender my friend (a specific application of the general *Don't misgender!* constraint

(5) Don't out my friend to her parents (NOT an application of any of the general constraints—instead, a constraint specific to the social situation)

I note that the constraint in (5) is not a contextually-bound application of a general constraint; this is because it is utterly dependent on the social knowledge of how parents can be anticipated to behave if they find out their offspring is transgender. This is not only culturally bound, but specific on an individual level—the behavior of a specific parent contributes significantly to the existence of this constraint. These individual social constraints are still, however, built off the more general architecture of anticipating and aligning (or deigning not to align) with the stance of an interlocutor, and are a necessary part of the equation.

The inclusion of socially-specific constraints along with more general constraints is further indication that the gender of any given pronoun is decided by discourse context, including world knowledge about the social landscape around the immediate context of the conversation. Chapter 4 gives ample data in support of a discourse-pragmatic analysis of 'natural' gender that is divorced from supposedly-inherent or biological sexed

categories, especially around pronouns. Chapter 4 also gives a model of pragmatics that is highly flexible and generalizable, and the OT-like pragmatic model should be able to translate equally well across different datasets and languages so long as the relevant constraints are included.

In Chapter 5 I proposed an analysis of *n* to D head movement that differentiated three sub-types of pronouns, which are distinguished by how far *n* raises in the nominal spine. I adopted Matushansky's 2006 split analysis of head movement where movement proper occurs in the narrow syntax, and morphological merger occurs at or after Spell-Out. In instances where the pronoun raises all the way to D, the phase head, use conditions are evaluated with respect to a particular referent. This phase-based approach of use conditional semantics accounts for the instances where the gender features of a pronoun do not cause ungrammaticality or semantic incoherence, but rather constitute an inappropriate way to refer to a specific person.

This analysis divided pronouns into three sub-types. Predicative pronouns, which in English manifest as depronominalizations or PRCs, are pronouns that merge in *n* and combine with Num and D but do not undergo head movement; for predicative pronouns, the associated D can either be an article or a GEN(eric) determiner, neither of which have strong uninterpretable features that would necessitate head movement. Additionally, predicative pronouns with uNum combine with a Num head that has a iNum feature which is not strong–thus, number morphology on predicative pronouns is regularized, and predicative pronouns when plural receive a plural -s suffix in English (rather than the syncretic plural morphology usually expected in pronouns).

The second type of pronouns, variable pronouns, are constituted of reflexives and bound anaphora. The D for variable pronouns has no strong features, and consists of a variable *x* that must be bound locally in order to evaluate any use or truth conditions. The uNum feature in variable pronouns is strong, meaning that pronouns merge in *n* and raise to Num (but no further) and receive syncretic plural morphology–which differentiates variable pronouns from predicative ones.

The third type of pronouns I discussed is referential pronouns, which are obligatorily free and refer directly to a referent with or without a linguistic antecedent. For referential pronouns, D is a direct referential index and has a strong uninterpretable feature uRef. The uRef and uNum features trigger head movement from *n* to Num to D, and upon combining with D the pronoun is evaluated with respect to discourse context for appropriateness.

The features needed for this analysis include 'Ref,' which I take to be an edge linking feature that can only appear on phase heads and which is the syntactic reflex of direct indexation with a referent, and Num, which is responsible for number morphology. In Chapter 5 I also discussed a third feature, uGender, but for the purposes of my proposal the Gender feature is not present for all speakers. The presence or absence of uGender on D for referential pronouns determines whether individual speakers can grammatically produce dsT, where singular *they* is used referentially for a specific person. For speakers who maintain the uGender feature, dsT causes a crash because *they*, lacking gender features, cannot check off uGender even if it raises to D. For speakers who can use dsT, gender is an optional feature and there is no uninterpretable uGender probe on D, so dsT does not cause a crash when used referentially.

Chapter 5 draws several conclusions related to the narrow syntax. First, head movement to phase heads results in a Spell-Out process that is qualitatively different than head movement that does not move to phase heads; phase edges are the locus of discourse sensitivity through edge linking features.

The parameterization of uGender I proposed in Chapter 5 also suggests that previously-formal features (such as gender in English, historically grammatical) can be repurposed from formal to pragmatic uses; however, this repurposing is related to a general syntactic weakening of these features, such that gender has become non-obligatory in English even in pronouns. This weakening and resulting optionality also means that gender features (or any other features which get repurposed for pragmatic uses) are more discourse-sensitive and subject to less strict agreement requirements. The structural difference be-

tween variable pronouns and referential pronouns that I propose is helpful in explaining why the apparent agreement requirements for bound pronouns are stricter than those for referential pronouns. Singular *they* is a possibly-unique case in that it is radically underspecified, meaning that under looser agreement requirements it fails to cause conflict even when (non-locally) coreferential with a pronoun of different gender features.

Finally, in Chapter 5 I discussed the cross-linguistic predictions made by this proposal. Future work investigating pronouns should be careful to look for exceptional use of social gender, especially with an eye towards how speakers use gender features for sociopragmatic purposes. The proposal I have given predicts that the sociopragmatic behaviors of pronouns should be constrained by their syntactic structure; referential pronouns under this analysis are freer to coopt towards discourse goals, while variable (bound) pronouns are predicted to be much more constrained.

This dissertation has been an effort to show how pronouns 'emerge,' both out of syntactic word-formation and out of sociolinguistic variation and innovation. While I have shown a wide range of previously under-documented empirical phenomena, the practical constraints of research have necessitated a certain amount of extrapolation; my intent has been to put forward falsifiable hypotheses that future scholars can take up for further testing.

APPENDIX

Appendix materials for Chapter 3

Additional graphs showing non-significant results from Experiments 1 and 2

Figure 6.1 shows the production of dsT by each participant in Experiment 1, and includes token counts of all third person singular pronouns for comparison. For most speakers dsT was a small minority of all the third person singular pronouns they used.

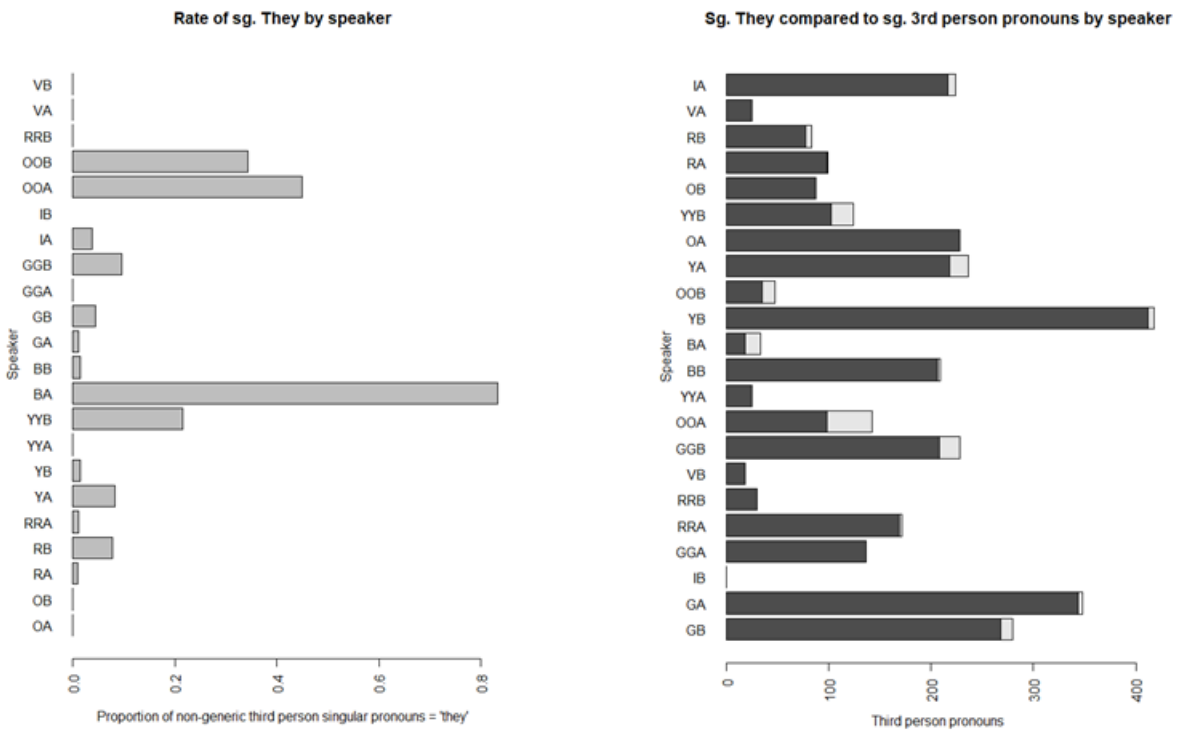


Figure 6.1: Proportion of dsT for each speaker

Figure 6.2 shows that rates of production of dsT did not significantly vary by speaker

income in Experiment 1.

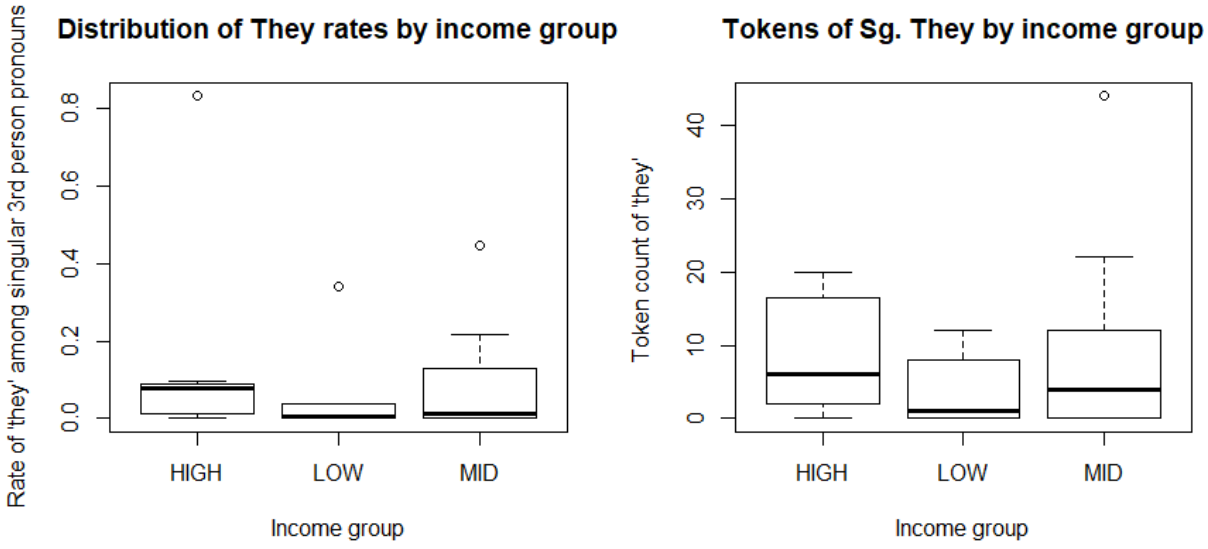


Figure 6.2: Use of dsT by income group of speaker

Figure 6.3 below shows the rate and token counts by age group in Experiment 1.

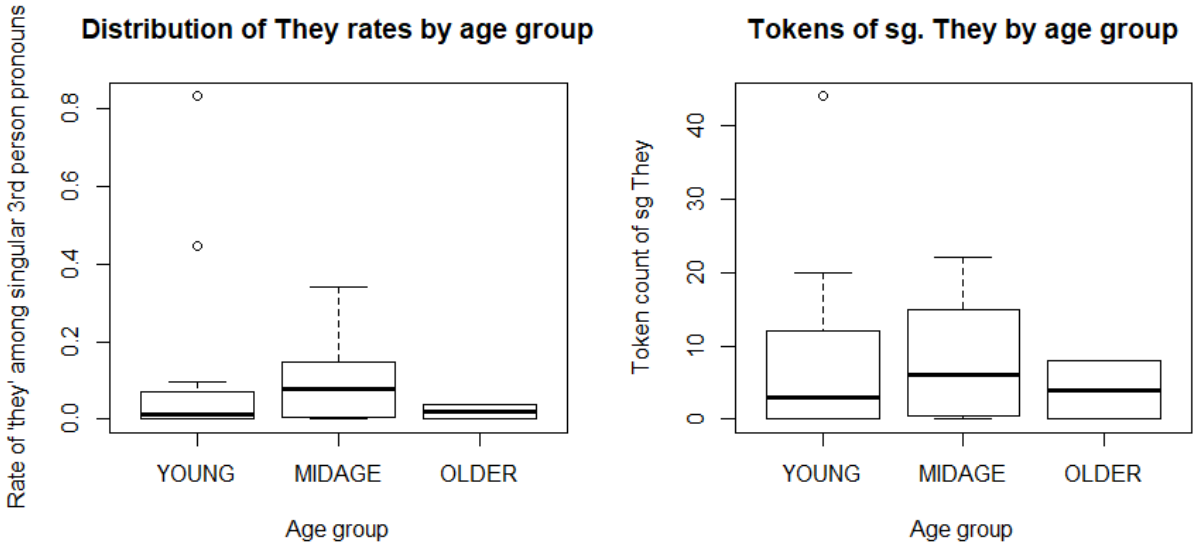


Figure 6.3: Use of dsT by age group of speaker

Figure 6.4 shows the main effect of antecedent type only on singular *they* in Experiment Two, leaving aside *he/she*.

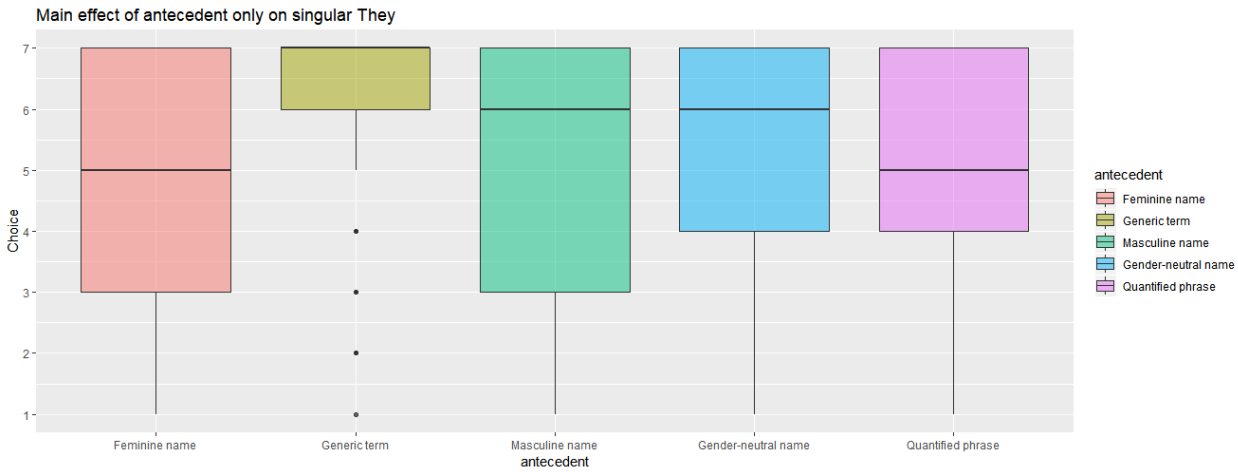


Figure 6.4: Rating of all pronouns by antecedent type

Figure 6.5 shows how participants of different genders rated dsT in Experiment 2.

The box width in Figure 6.5 indicates how many participants were in each gender group (where N/A includes participants who gave no response to the gender question).

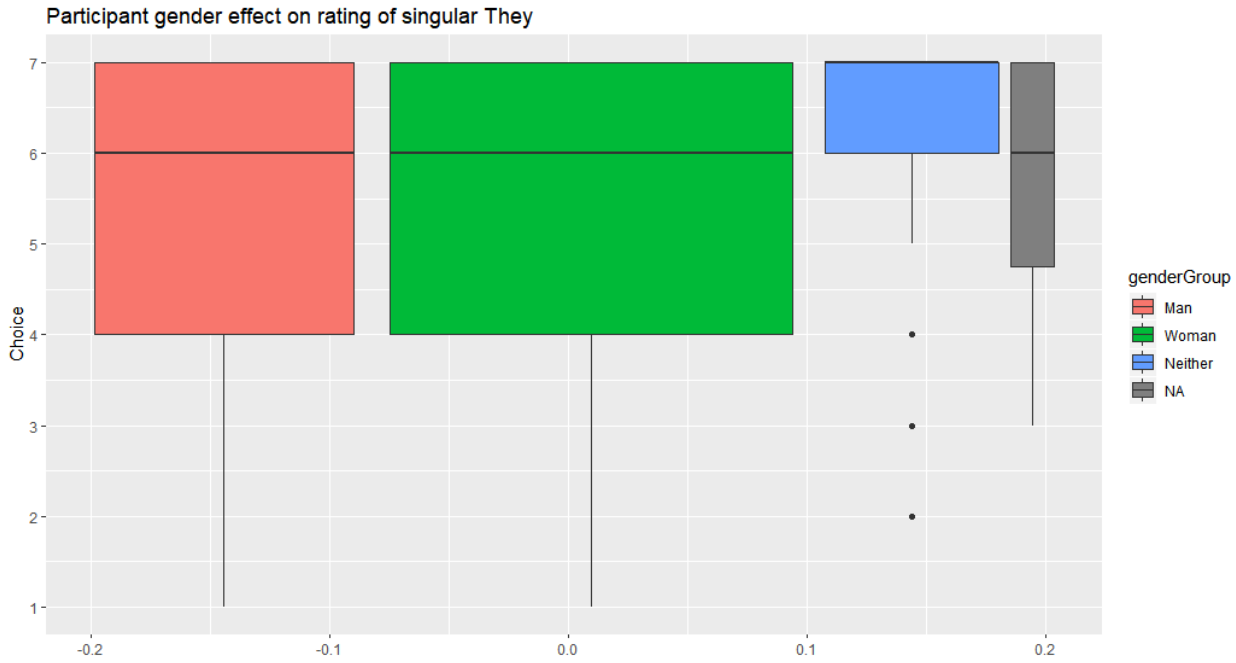


Figure 6.5: Rating of singular *they* by speaker gender

Appendix 1: Experiment One Materials

Survey materials

(6) **Demographics survey: Please fill out each section with how you most closely identify:**

Age:

Ethnicity:

Gender:

Preferred pronouns:

Are there any other pronouns you're okay with people using?

Have you discussed your pronoun preferences with your partner?

Sexual orientation:

Where are you from?

About how much is your yearly household income?

Interview questions:

(7) **Strangers:**

a. **Pair interview:**

b. What's your first impression of each other? Can you describe each other?

c. Can you find a topic to disagree about? How would you describe each others' perspective?

d. What about something you both agree on? What's something you both have in common?

e. Do you remind each other of anyone? Who does X remind you of and why?

f. **Solo interview:**

g. How would you describe X when you're talking to your other friends? Personality, looks, stories?

h. What's one thing you really like or admire about X? What about something about X that irrituates you?

i. Is there anything like a book or TV show that makes you think of X? Can you talk about why?

j. How do you think the first interview went? How were you feeling, and how do you think X was feeling? Is this what you expected?

(8) **Acquaintances:**

a. **Pair interview:**

b. How do the two of you know each other? What was your first meeting like? Is there a story there?

c. Can you tell me about a time that the two of you disagreed about something? Do the two of you have any differing opinions?

d. Can you tell me about one of your happiest memories together? What about a not-so-happy memory?

e. Do you have any mutual friends that know both of you? How did you meet them?

(9) **Film clip response survey:**

(10) **About each character in each clip:**

(11) How did you feel about this clip?

(12) How did you feel about (Character 1)?

(13) How did you feel about (Character 2)?

These are statements about the clip you just watched. Each of the following statements is followed by a scale. Please rate the statement from 0-'Strongly disagree' to 6-'Strongly agree.' Circle one number to show your agreement or disagreement with each statement. Answer with your first impression or instinct.

(14) I like this clip.

(Strongly disagree) 0 1 2 3 4 5 6 (Strongly agree)

(15) (Character) is a good person.

(16) (Character) is educated.

(17) (Character) is a typical woman.

(18) (Character) is a typical transgender person.

(19) (Character) is a typical man.

(20) (Character) is friendly.

(21) (Character) is intelligent.

(22) (Character) is attractive.

(23) (Character) is offensive.

(24) (Character) is funny.

(25) I could be friends with (Character).

(26) The character '(Character)' is realistic.

About the Movie:

Have you seen this movie (*Boy Meets Girl*, 2014) before? If so, did you like it? What did you think about it? If not, what do you think of what you've seen?

What do you think about the actors playing Ricky (Michelle Hendley), Francesca (Alexandra Turshen), David (Michael Galante), and Robby (Michael Welch)?

Do you have anything else you want to add?

(27) (27) **Attitude survey:**

Section 1: Feelings Thermometer

On a scale from 0 to 100, what are your personal feelings towards the following groups? As you do this task, think of a thermometer: 0 is the 'coldest,' least favorable rating, and 100 is the 'warmest,' most favorable rating.

(28) Men in general: _____

(29) Women in general: _____

(30) Gay men: _____

(31) Lesbians: _____

(32) Bisexual men: _____

(33) Bisexual women: _____

(34) Transgender people in general: _____

(35) Transgender men: _____

(36) Transgender women: _____

Section 2: Scalar Questions

Each of the following statements is followed by a scale. Please rate the statement from 0-'Strongly disagree' to 6-'Strongly agree.' Circle one number to show your agreement or disagreement with each statement. Answer with your first impression or instinct.

(37) Sex between two women is wrong.

(Strongly disagree) 0 1 2 3 4 5 6 (Strongly agree)

(38) I think lesbians are disgusting.

(39) Female homosexuality is a natural expression of sexuality.

(40) Sex between two men is wrong.

(41) I think gay men are disgusting.

(42) Male homosexuality is a natural expression of sexuality.

(43) I think transgender people are unnatural.

(44) Transgender women should not be allowed in women's bathrooms.

(45) Transgender men should not be allowed in men's bathrooms.

(46) Transgender women are real women.

(47) Transgender men are real men.

(48) I have more than a few gay male friends.

(49) I have more than a few lesbian friends.

(50) I have more than a few transgender friends.

(51) It is valid for people to identify as neither men nor women.

People who identify as neither men nor women are confused.

***Boy Meets Girl* Clip Transcripts**

Clip 1

Hey, how's it goin'?

What can I get ya?

You have the prettiest hair. I always wanted to try those straight-across bangs like that.

Yeah, and you know I've always wanted to try wearin' a big ole sparkly engagement ring like that.

Yeah, well, you know the grass is always greener, I guess.

- So...?

- Oh! Right. May I please have a double latte caramel frappuccino, please.

Sweetheart, Starbucks is on Montague Street. We sell coffee here.

Oh, right. Um,

- well, may I have...

- Tread lightly now.

- Um, a ca- cappa...

- Cap...?

Chino?

Yes, you may.

So, you've never been married or engaged, even?

Mm-mm. No.

I mean, I do, but most of 'em I'm not really attracted to and the ones I have been always seem to get cold feet.

Oh, God, men. The very thing they want most is the very thing they're most afraid of.

Ain't that the truth.

- Commitment.

- Dick.

Oh, well, I mean that, too, but it's a distant second.

Did you say what I think you just said?

Mmm, that commitment's a distant second?

No. Before that.

That men want dick but are afraid of it.

Okay. Apparently you did.

Two-fifty, please.

Thank you.

Do you- do you have a lot of guy friends confide this in you? I-I mean how do you know that they're telling you the truth and not just messin' with you?

'Cause if I weren't, I'd have a big ole gorgeous ring on my finger by now.

I don't follow. That's my mom. We're late for church.

Oh, I understand.

It was real nice talkin' with you.

And you.

I'm Francesca, by the way.

Ricky.

Clip 2

Okay, so there is one main reason why, even if he did like me like that, we still couldn't be anything more than friends.

Mm-hmm, and what's that?

- Well...

- Tell me.

Okay.

- What's your number?

- Why?

Just, please.

Uh, eight-five-nine-five-five-five-seven-six-two-four.

Oh.

Who the heck is...?

It's me.

Oh!

How excitin'.

No.

But...

No!

Wow!

May I ask you a question?

Yeah, sure.

So where did you- like where did you get it? I mean who gave it to you?

What?

Y-your, um...

Oh!

Well, um, God, I guess?

When I was born.

Oh, my- okay,

so biologically you were born -

A boy.

Well, um,

so, uh, do you like it?

My...?

- Yeah.

- Oh! Yeah, I mean, I do, it's just I wish I'd been born a genetic girl. I do plan on gettin' the full surgery someday. It's just so expensive. But, for now, I might as well just dance with the one that brung me, right? And it's not really about hatin' my body, so I've learned to live with it.

And do you- do you like, keep it a secret that you..?

Oh, no, no, no. I am completely comfortable with who I am, and I like to make sure everybody else in my life is, too. So, you know, whether it's a boy, or like you, a new friend, um, if for any reason it should or does come up, I mean, I just get it right out in the open.

Well, we are new friends, aren't we?

I think so.

As long as you don't go fallin' in love with me or anything.

Well, right back at ya!

Clip 3

Hello, Baby.

Ah, I miss you, Baby Doll. So how's everythin' goin' at home?

Well, I am gettin' my dress tomorrow.

Oh, you're gonna look just like an angel. You know, that's probably the one thing civilized about these Arabs, saving themselves for marriage.

Just like you, Baby. My virgin angel.

God, stop! Oh! Hey, uh, I made some new friends today, and they say they know you.

- No way.

- Yeah.

Well, who? Uh, Robby Riley? Uh, and his friend Ricky?

The tranny?

- Don't call her that.

- You hung out with them?

Well, she's just like any other normal girl. I mean she's just like me.

Okay, seriously, you're gonna make me puke now, you keep talkin' like that.

Well, she said since you're fightin' for America and all, that you're all right in her book.

Yeah, well, I ain't fightin' for the America that fucking thing is part of!

I just miss you, Baby. I can't wait 'til you're home. Ninety-three more days.

Yeah.

Well, I gu-I guess I should be goin'.

Oh, yeah.

I love you.

Be careful.

I love you too, Baby.

Clip 4

Really, Ricky, don't worry about it.

I'm not worried, I'm just annoyed. They should be growin' faster by now. I've been on hormones for, like, seven years

Yeah, but guy-guys don't care about that shit.

Yeah, well as soon as I can afford it I'm gonna get implants.

No, but-but it's, Ricky, it's not the size, okay? It's the- it's the shape, it's the feel, it's the- it's the buoyancy.

You don't know what buoyancy means.

The cute little tits are just so sexy, okay? They feel great, they're wonderful. I- those-those big ole fake balloons fuckin' thing -

I wouldn't get anything like that.

Just nice, small elegant C-cups.

I-I wouldn't.

Really.

Don't you peek, now.

Yuck. As if.

Okay, Sam, you can come back in.

Appendix 2: Experiment Two Materials

Target stimuli (3 pronouns x 5 antecedents = 15 x 2 versions each = 30 items)

target 1 masc-he1

John is very forgetful. He never remembers library due dates.

target 1 masc-he2

Bill is very forgetful. He never remembers library due dates.

target 2 fem-he

Susan is very studious. He sets aside two hours every night for homework.

target 2 fem-he

Jessica is very studious. He sets aside two hours every night for homework.

target 3 neu-he

Lee is very creative. He is always learning new ways to make art.

target 3 neu-he

CJ is very creative. He is always learning new ways to make art.

target 4 masc-she

Rob is very sensitive. She can always tell how people are feeling.

target 4 masc-she

Jack is very sensitive. She can always tell how people are feeling.

target 5 fem-she

Jenny is very outgoing. She always makes friends with strangers at cafes.

target 5 fem-she

Darla is very outgoing. She always makes friends with strangers at cafes.

target 6 neu-she

Riley is very curious. She loves to check out new museums and bookstores.

target 6 neu-she

Reese is very curious. She loves to check out new museums and bookstores.

target 7 masc-they

Ivan is very generous. They usually have a spare dollar to give to a good cause.

target 7 masc-they

Mark is very generous. They usually have a spare dollar to give to a good cause.

target 8 fem-they

Reba is very clumsy. They can't try a new sport without getting injured.

target 8 fem-they

Sarah is very clumsy. They can't try a new sport without getting injured.

target 9 neu-they

Hayden is very funny. They can make a pun out of almost any word.

target 9 neu-they

Jordan is very funny. They can make a pun out of almost any word.

target 10 quant-he

Concerts are very popular. Every music fan tries to buy his ticket early.

target 10 quant-he

Concerts are very popular. Each music fan tries to buy his ticket early.

target 11 quant-she

Students are very ambitious. Every student tries to write her essay perfectly.

target 11 quant-she

Students are very ambitious. Each student tries to write her essay perfectly.

target 12 quant-they

Dogs are very cute. Every dog owner tries to take their dog's picture constantly.

target 12 quant-they

Dogs are very cute. Each dog owner tries to take their dog's picture constantly.

target 13 gen-he

The perfect spouse is very thoughtful. He will always try to remember birthdays and anniversaries.

target 13 gen-he

The perfect spouse is very considerate. He will always try to remember birthdays and anniversaries.

target 14 gen-she

The average driver is very cautious. She will always come to a full stop at a stop sign.

target 14 gen-she

The average driver is very careful. She will always come to a full stop at a stop sign.

target 15 gen-they

The ideal barista is very attentive. They will always make drinks carefully and quickly.

target 15 gen-they

The ideal barista is very efficient. They will always make drinks carefully and quickly.

Filler stimuli (30 items)

filler 1 filler

Tractors are very heavy. They often has computer software to help with farming.

filler 2 filler

That dog is very friendly. It likes run around with other dogs at the dog park.

filler 3 filler

The dean is very mysterious. The dean is a man who I don't know what looks like.

filler 4 filler

Many investors are very wealthy. They a lot of assets and real estate own.

filler 5 filler

Some websites are very complicated. They have been usually built by teams of people.

filler 6 filler

My cat is very sneaky. It always pounces on my feet when I walk down the hall.

filler 7 filler

The penguin is very majestic. It can grow up to five feet tall.

filler 8 filler

Wedding cakes are very detailed. They often take hours to bake and decorate.

filler 9 filler

Artists are very particular. Many will only buy a certain brand of paint or canvas.

filler 10 filler

Programmers are very fastidious. Some of refuse to use a computer with a dirty keyboard.

filler 11 filler

Most politicians is very honest. Most want to serve all of their constituents equally.

filler 12 filler

The smallest teacup is very delicate. It was made by a master artisan in the country.

filler 13 filler

This book is very long. The author wrote it in only one hundred days of work.

filler 14 filler

Cell phones are very expensive. Many people don't like to buy a new one every year.

filler 15 filler

Some sticky notes are very cheap. The adhesive is weak and doesn't stick at all.

filler2 1 filler2

My oven is completely broken. I my landlord need to ask to fix it.

filler2 2 filler2

The biggest airplanes are not fuel efficient. Airlines shouldn't do nothing to fix the problem.

filler2 3 filler2

Sunflower stems are tough and fibrous. What and fibrous are daisy stems?

filler2 4 filler2

All shower curtains are transparent. The opaque one is cheapest on the online store.

filler2 5 filler2

Ben's sister is very smart. Dave's is even smarter sister than Ben's.

filler2 6 filler2

International trade is very important. Many people do not understand no details of the issues.

filler2 7 filler2

Pure coal is very expensive. It might should be subsidized by the government in some states.

filler2 8 filler2

Small snakes are very cute. Some people are very afraid of green small cute snakes.

filler2 9 filler2

Captive zebras are very intelligent. They are having more puzzle-solving abilities than donkeys.

filler2 10 filler2

Long books are very heavy. Many believe that electronic files are better than.

filler2 11 filler2

Strong smells are very annoying. If there is a strong then people will want to leave the room.

filler2 12 filler2

Lion cubs are incredibly cute. Mother lions will pick them up in their mouths and carry them.

filler2 13 filler2

Sharp edges are very dangerous. If you do not pay attention you can cut yourself on them.

filler2 14 filler2

Chronic conditions are very exhausting. Having a chronic illness drains your bank account and your energy.

filler2 15 filler2

Church bells are very loud. Living next to a building with them can be annoying if you are.

Comment prompt

What kinds of things did you notice in these sentences? what information about the sentences did you use to inform your ratings? Do you have any tips for us to improve these sentences?

Demographic questions and answer options

Are you a native speaker of English? Yes No

Do you speak a language other than English? Yes NoWhat language(s) do you speak?:

How old are you? Under 20 20-25 26-30 31-35 36-40 41-45 46-50 51-55 56-60 61-65 66-70 Over 70

How would you best describe your gender? Man Woman Other

Do you identify as transgender? Yes No

How would you best describe your ethnicity? (Categories reflect the U.S. Census) White Black or African American American Indian or Alaska Native Asian Native Hawaiian or Other Pacific IslanderOther:

Which dialect area of the United States do you identify most closely with? New England Atlantic South Delta South MidwestCentral WestSouthwestNorthwestOutside U.S.If outside U.S., where are you from?

BIBLIOGRAPHY

Aarts, B., & McMahon, A. (2006). *The Handbook of English Linguistics*. Wiley Online Library.

Abney, S. P. (1987). *The English noun phrase in its sentential aspect* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Aboh, E. O. (2004). Topic and focus within D. *Linguistics in the Netherlands*, *21*(1), 1–12.

Ackerman, L. (2017). *Reevaluating how gender is represented and used during sentence processing.* Newcastle University: Paper presented at the 12th North East Language and Linguistics Research Away Day (NELLi), Newcastle University. Retrieved from https://drive.google.com/file/d/0B5blL1tcy5_hV1ZLWlpkRlptVWM/view

Ackerman, L. (2019). Syntactic and cognitive issues in investigating gendered coreference. *Glossa*. Retrieved from https://psyarxiv.com/6cf57/

Ackerman, L., Riches, N., & Wallenberg, J. (2018). *Coreference dependency formation is modulated by experience with variation of human gender.* Paper presented at the 92nd Annual Meeting of the Linguistic Society of America, Salt Lake City, UT. Retrieved from https://drive.google.com/file/d/1yEl1eJRelU-c7u3y7kBGyUWkHEvTwdc7/view

Adger, D., & Smith, J. (2010). Variation in agreement: A lexical feature-based approach. *Lingua*, *120*(5), 1109–1134.

Anand, P., & Nevins, A. (2004). Shifty operators in changing contexts. In *Semantics and Linguistic Theory* (Vol. 14, pp. 20–37).

Ansara, Y. G., & Hegarty, P. (2014). Methodologies of misgendering: Recommendations for reducing cisgenderism in psychological research. *Feminism & Psychology*, *24*(2), 259–270.

Armoskaite, S., & Wiltschko, M. (2012). There are many ways to be gendered. In *Proceedings of the 2012 annual conference of the canadian linguistic association*.

Audring, J. (2008). Gender assignment and gender agreement: Evidence from pronominal gender languages. *Morphology*, *18*(2), 93–116.

Austin, J. L. (1975). *How to do things with words*. Oxford University Press.

Bailey, G., Wikle, T., Tillery, J., & Sand, L. (1991). The apparent time construct. *Language Variation and Change*, *3*(3), 241–264.

Bell, A. (1984). Language style as audience design. *Language in Society*, *13*(2), 145–204.

Bernstein, J. B. (2001). The DP hypothesis: Identifying clausal properties in the nominal domain. In M. R. Baltin & C. Collins (Eds.), *The Handbook of Contemporary Syntactic Theory* (p. 536). Blackwell.

Bhatt, R. (2002). The raising analysis of relative clauses: Evidence from adjectival modification. *Natural Language Semantics*, *10*(1), 43–90.

Bianchi, V. (1999). *Consequences of antisymmetry: Headed relative clauses* (Vol. 46). Walter de Gruyter.

Biggs, S. (2007). Thinking about generations: Conceptual positions and policy implications. *Journal of Social Issues*, *63*(4), 695–711.

Bjorkman, B. M. (2017). Singular they and the syntactic representation of gender in English. *Glossa*, *2*(1), 80.

Bobaljik, J. D., & Yatsushiro, K. (2006). Problems with honorification-as-agreement in Japanese: A reply to Boeckx & Niinuma. *Natural Language & Linguistic Theory*, *24*(2), 355–384.

Bodine, A. (1975). Androcentrism in prescriptive grammar: singular 'they', sex-indefinite 'he', and 'he or she'. *Language in Society*, *4*(2), 129–146.

Boeckx, C. (2006). Honorification as agreement. *Natural Language & Linguistic Theory*, *24*(2), 385–398.

Boeckx, C., & Niinuma, F. (2004). Conditions on agreement in japanese. *Natural Language & Linguistic Theory*, *22*(3), 453–480.

Borer, H. (2005). *In name only* (Vol. 1). Oxford University Press.

Bresnan, J., & Grimshaw, J. (1978). The syntax of free relatives in English. *Linguistic Inquiry*, *9*(3), 331–391.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage* (Vol. 4). Cambridge University Press.

Brown, R., Gilman, A., et al. (1960). The pronouns of power and solidarity. In T. A. Sebeok (Ed.), *Style in language* (pp. 253–76). MIT Press.

Butler, J. (2011). *Gender trouble: Feminism and the subversion of identity*. Routledge.

Cardinaletti, A. (1994). On the internal structure of pronominal DPs. *The Linguistic Review*, *11*(3-4), 195–220.

Carlson, G. N., & Pelletier, F. J. (1995). *The Generic Book*. University of Chicago Press.

Cecchetto, C., & Donati, C. (2015). *(Re) labeling* (Vol. 70). MIT Press.

Chomsky, N. (1965). Aspects of the theory of syntax. *Multilingual Matters*.

Chomsky, N. (1995). *The Minimalist Program* (Vol. 28). Cambridge University Press.

Chomsky, N. (2000). *Minimalist inquiries: The framework.* Cambridge, Massachussets: MIT Press.

Chomsky, N. (2001). *Beyond Explanatory Adequacy* (No. 20). MIT working papers in linguistics.

Chomsky, N. (2004). Beyond explanatory adequacy. In A. Belletti (Ed.), *Structures and beyond* (Vol. 3, pp. 104–131). Oxford: Oxford University Press.

Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry*, *36*(1), 1–22.

Chomsky, N. (2008). On phases. *Current Studies in Linguistics Series*, *45*, 133.

Chomsky, N. (2015). Problems of projection. *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*, *223*, 3.

Cinque, G. (1982). On the theory of relative clauses and markedness. *The Linguistic Review*, *1*(3), 247–296.

Citko, B. (2004). On headed, headless, and light-headed relatives. *Natural Language and Linguistic Theory*, *22*(1), 95–126.

Citko, B. (2014). *Phase theory: An introduction*. Cambridge University Press.

Conrod, K. (2017a). *Names before pronouns: Variation in pronominal reference and gender.* Vancouver, BC: Paper presented at Cascadia Workshop in Sociolinguistics, University of British Columbia.

Conrod, K. (2017b). *Pronominal and quantificational relative clauses.* Unpublished manuscript.

Conrod, K. (2017c). *Pronouns and misgendering.* Unpublished manuscript.

Conrod, K. (2018a). *Changes in singular they.* Portland, OR: Paper presented at Cascadia Workshop in Sociolinguistics, Reed College.

Conrod, K. (2018b). *Depronominalization and gender ideology.* Paper presented at New Ways Of Analyzing Variation, University of Wisconsin.

Conrod, K., Tatman, R., & Koncel-Kedziorski, R. (2016). Pronominal relative clauses on Twitter. In *Proceedings of the Corpus Linguistics Fest* (pp. 17–21). Bloomington, Indiana. Retrieved from http://ceur-ws.org/Vol-1607/conrod.pdf

Conrod, K., & Woo, B. (2018). *Hydras: Split heads and light heads.*

Corbett, G. G. (2006). *Agreement* (Vol. 109). Cambridge University Press.

Cowper, E., & Hall, D. C. (2009). Argumenthood, pronouns, and nominal feature geometry. *Determiners: Universals and variation*, *147*, 97–120.

Cukor-Avila, P., & Bailey, G. (2013). Real time and apparent time. *The Handbook of Language Variation and Change*, 239–262.

Curme, G. O. (1912). A History of English Relative Constructions. *The Journal of English and Germanic Philology*, *11*(3), 355–380.

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

D'Arcy, A., Haddican, B., Richards, H., Tagliamonte, S. A., & Taylor, A. (2013). Asymmetrical trajectories: The past and present of –body/–one. *Language Variation and Change*, *25*(3), 287–310.

De Vries, M. (2002). *The syntax of relativization*. LOT.

Déchaine, R.-M., & Wiltschko, M. (2002). Decomposing pronouns. *Linguistic Inquiry*,

33(3), 409–442.

Delorme, E., & Dougherty, R. C. (1972). Appositive NP constructions: we, the men; we men; I, a man; etc. *Foundations of Language*, *8*, 2–29.

De Saussure, F. (2011). *Course in general linguistics*. Columbia University Press.

De Vries, M. (2010). Apparent nonlocality. *Linguistics in the Netherlands*, *27*(1), 129–140.

Donati, C., & Cecchetto, C. (2011). Relabeling heads: A unified account for relativization structures. *Linguistic Inquiry*, *42*(4), 519–560.

Du Bois, J. W. (2007). The stance triangle. *Stancetaking in discourse: Subjectivity, evaluation, interaction*, *164*(3), 139–182.

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin*, *23*(3), 316–326.

Eckert, P. (1998). Gender and sociolinguistic variation. In *Language and gender: a reader* (pp. 64–75).

Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, *12*(4), 453–476.

Eckert, P., & McConnell-Ginet, S. (2013). *Language and Gender*. Cambridge University Press.

Elbourne, P. (2013). *Definite Descriptions*. Oxford University Press.

Erlewine, M. Y., & Kotek, H. (2016). A streamlined approach to online linguistic surveys. *Natural Language & Linguistic Theory*, *34*(2), 481–495.

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.

Gastil, J. (1990). Generic pronouns and sexist language: The oxymoronic character of masculine generics. *Sex Roles*, *23*(11-12), 629–643.

Goffman, E. (1967). On face-work. *Interaction Ritual*, 5–45.

Grant, J. M., Mottet, L., Tanis, J. E., Harrison, J., Herman, J., & Keisling, M. (2011). *Injustice at every turn: A report of the national transgender discrimination survey*. National

Center for Transgender Equality.

Grice, H. P. (1968). Utterer's meaning, sentence-meaning, and word-meaning. In *Philosophy, Language, and Artificial Intelligence* (pp. 49–66). Springer.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics* (Vol. 3, pp. 41–58). Academic Press.

Groos, A., & Van Riemsdijk, H. (1981). Matching effects in free relatives: A parameter of core grammar in theory of markedness in generative grammar. In L. B. Belletti & L. Rizzi (Eds.), *Theory of markedness in generative grammar.* Scuola Normale Superiore.

Gumperz, J. J. (1982). *Discourse strategies* (Vol. 1). Cambridge University Press.

Gustafsson Sendén, M., Bäck, E. A., & Lindqvist, A. (2015). Introducing a gender-neutral pronoun in a natural gender language: the influence of time on attitudes and behavior. *Frontiers in Psychology*, *6*, 893.

Gutzmann, D., & McCready, E. (2014). Using descriptions. *Empirical Issues in Syntax and Semantics*, *10*, 55–72.

Hall, K. (2003). Exceptional speakers: Contested and problematized gender identities. In M. M. Holmes J. (Ed.), *The handbook of language and gender.* Oxford-Carlton, Victoria: Blackwell Publishing.

Hankamer, J., & Sag, I. (1976). Deep and Surface Anaphora. *Linguistic Inquiry*, *7*(3), 391–428.

Heim, I. (1991). Artikel und definitheit. *Semantik: ein internationales Handbuch der zeitgenössischen Forschung*, 487–535.

Herek, G. M. (2009). Sexual stigma and sexual prejudice in the united states: A conceptual framework. In *Contemporary perspectives on lesbian, gay, and bisexual identities* (pp. 65–111). Springer.

Hughes, D. L., & Casey, P. L. (1986). Pronoun choice for gender-unspecified agent words: developmental differences. *Language and Speech*, *29*(1), 59–68.

Hulsey, S., & Sauerland, U. (2006, June). Sorting out relative clauses. *Natural Language*

Semantics, *14*(2), 111–137. doi: 10.1007/s11050-005-3799-3

Hyde, J. S. (1984). Children's understanding of sexist language. *Developmental Psychology*, *20*(4), 697.

Johnstone, B., & Kiesling, S. F. (2008). Indexicality and experience: Exploring the meanings of /aw/-monophthongization in Pittsburgh. *Journal of Sociolinguistics*, *12*(1), 5–33.

Kayne, R. S. (1994). *The antisymmetry of syntax* (No. 25). MIT Press.

Kempsell Jacinto, W. C. (2015). *Language attitudes and bilingualism in turkish-german popular film* (Doctoral dissertation, University of Washington). Retrieved from https://digital.lib.washington.edu:443/researchworks/handle/1773/33995

Kerr, J. S., & Underwood, G. (1984). Fixation time on anaphoric pronouns decreases with congruity of reference. In *Advances in Psychology* (Vol. 22, pp. 195–202). Elsevier.

King, K. E. (2016). *Mixed gender agreement in Russian DPs* (Master's thesis, University of Washington). Retrieved from https://digital.lib.washington.edu/researchworks/handle/1773/35231

Konnelly, L., & Cowper, E. (2017). *The future is "they": The morphosyntax of an English epicene pronoun.* Manuscript.

Kramer, R. (2016). The location of gender features in the syntax. *Language and Linguistics Compass*, *10*(11), 661–677.

Kratzer, A. (2009). Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry*, *40*(2), 187–237.

Kratzer, A., & Heim, I. (1998). *Semantics in generative grammar*. Blackwell Oxford.

Kučerová, I. (2018). ϕ-features at the syntax-semantics interface: Evidence from nominal inflection. *Linguistic Inquiry*, *49*(4), 813–845.

Labov, W. (1966). The social stratification of English in New York City.

Labov, W. (1994). *Principles of linguistic change, internal factors*. Wiley.

Labov, W. (2018). The role of the Avant Garde in linguistic diffusion. *Language Variation*

and Change, *30*(1), 1–21.

Lees, R. B. (1960). The grammar of English nominalizations. *Folklore and Linguistics*.

Lees, R. B. (1961). The constituent structure of Noun Phrases. *American Speech*, *36*(3), 159–168.

Levinson, S. C. (1983). Pragmatics. *Cambridge/New York*.

Lobeck, A. (1995). *Ellipsis: Functional heads, licensing, and identification*. Oxford University Press.

Lobeck, A. (2006). Ellipsis in DP. In M. Everaert & H. van Riemsdijk (Eds.), *The Blackwell Companion to Syntax* (pp. 145–173). Blackwell Publishing.

Longobardi, G. (1994). Reference and proper names: A theory of N-movement in syntax and logical form. *Linguistic Inquiry*, 609–665.

MacKay, D. G. (1980). On the goals, principles, and procedures for prescriptive grammar: Singular they. *Language in Society*, *9*(3), 349–367. doi: https://doi.org/10.1017/S0047404500008253

MacKay, D. G., & Fulkerson, D. C. (1979). On the comprehension and production of pronouns. *Journal of verbal learning and verbal behavior*, *18*(6), 661–673.

Mair, C., & Leech, G. (2006). Current changes in English syntax. *The Handbook of English Linguistics*, *36*, 318. doi: https://doi.org/10.1002/9780470753002.ch14

Manlove, K. (2016). *How peripheral a phenomenon is it?: On the interaction of DP-internal ellipsis, degree inversion, and islands* (Unpublished doctoral dissertation). University of Washington.

Marantz, A. (1985). Lexical decomposition vs. affixes as syntactic constituents. In *Cls 21: Papers from the parasession on causatives and agentivity, university of chicago, chicago.*

Marantz, A. (1988). Clitics, morphological merger, and the mapping to phonological structure. In *Theoretical Morphology.* Academic Press.

Marantz, A. (1999). *Creating words above and below little v.* Manuscript.

Marantz, A. (2000). Roots: the universality of root and pattern morphology. In *Conference on Afro-Asiatic languages, University of Paris VII* (Vol. 3, p. 14).

Matushansky, O. (2006). Head movement in linguistic theory. *Linguistic Inquiry*, *37*(1), 69–109.

Matushansky, O. (2015). The other Francis Bacon: on non-bare proper names. *Erkenntnis*, *80*(2), 335–362.

McConnell-Ginet, S. (2013). Gender and its relation to sex: the myth of 'natural' gender. In *The expression of gender* (pp. 3–38). De Gruyter Mouton Berlin.

McIntyre, J. (2009). Revealing and revolting: Homosexual and transgender panic in two Australian films. *Intersections: Women's and Gender Studies in Review Across Disciplines*, *7*, 40–52.

McKinney-Bock, K. (2013). Deriving split-antecedent relative clauses. *University of Pennsylvania Working Papers in Linguistics*, *19*(1), 14.

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

McLemore, K. A. (2018). A minority stress perspective on transgender individuals' experiences with misgendering. *Stigma and Health*, *3*(1), 53.

Meerwijk, E. L., & Sevelius, J. M. (2017). Transgender population size in the united states: a meta-regression of population-based probability samples. *American Journal of Public Health*, *107*(2), e1–e8.

Melchin, P. B. (2015). On the modifiability and syntactic category of pronouns in Japanese and Mandarin. In *MOTH Syntax Workshop*.

Mencken, H. L. (1945). *The American language: An inquiry into the development of English in the United States* (Vol. 2). Alfred a Knopf Incorporated.

Milroy, L., & Margrain, S. (1980). Vernacular language loyalty and social network. *Language in Society*, *9*(1), 43–70.

Moulton, J., Robinson, G. M., & Elias, C. (1978). Sex bias in language use: "Neutral" pronouns that aren't. *American Psychologist*, *33*(11), 1032. doi: https://doi.org/10.1037/0003-066X.33.11.1032

Mühlhäusler, P., & Harré, R. (1990). *Pronouns and people: The linguistic construction of*

social and personal identity. Blackwell.

Muysken, P. (2008). *Functional categories* (Vol. 117). Cambridge University Press.

Newman, M. (1997). *Epicene pronouns: The linguistics of a prescriptive problem*. Garland.

Niinuma, F. (2003). *The syntax of honorification* (Doctoral dissertation, University of Connecticut). Retrieved from http://search.proquest.com/docview/305334022/

Norton, A. T., & Herek, G. M. (2013). Heterosexuals' attitudes toward transgender people: Findings from a national probability sample of US adults. *Sex Roles*, *68*(11-12), 738–753. doi: https://doi.org/10.1007/s11199-011-0110-6

Ntelitheos, D. (2004). *Syntax of elliptical and discontinuous nominals* (Unpublished master's thesis). University of California Los Angeles.

Palakornkul, A. (1975). A socio-linguistic study of pronominal usage in spoken Bangkok Thai. *International Journal of the Sociology of Language*, *1975*(5), 11–42.

Papadopoulos, B. (2019). *Innovaciones al género morfológico en el Español de hablantes genderqueer (morphological gender innovations in Spanish of genderqueer speakers)* (Unpublished master's thesis). University of California Berkeley.

Partee, B. (1975). Montague grammar and transformational grammar. *Linguistic Inquiry*, 203–300.

Pesetsky, D. (1978). Category switching and so-called pronouns. In *Papers from the 14th regional meeting chicago linguistic society* (Vol. 14, pp. 350–361). Chicago, IL.

Pesetsky, D. (2013). *Russian case morphology and the syntactic categories* (Vol. 66). MIT Press.

Pike, K. L. (1967). Etic and emic standpoints for the description of behavior.

Postal, P. (1966). On so-called pronouns in English. *Monograph series on language and linguistics*, *19*, 177–206.

Potts, C. (2007). The expressive dimension. *Theoretical Linguistics*, *33*(2), 165–198.

Potts, C., & Kawahara, S. (2004). Japanese honorifics as emotive definite descriptions. In *Semantics and Linguistic Theory* (Vol. 14, pp. 253–270).

Prasad, G., Morris, J., & Feinstein, M. (2018). *The P600 for singular 'they': How the brain*

reacts when John decides to treat themselves to sushi. Poster presented at the 31st Annual CUNY Conference on Human Sentence Processing, Davis, CA. Retrieved from https://grushaprasad.github.io/files/CUNY2018_poster_draft2.pdf

Prewitt-Freilino, J. L., Caswell, T. A., & Laakso, E. K. (2012). The gendering of language: A comparison of gender equality in countries with gendered, natural gender, and genderless languages. *Sex Roles*, *66*(3-4), 268–281.

Pullum, G. (2003). *Postcard from Vegas, 2: Syntactic data collection on the strip.* Retrieved 2018, from http://itre.cis.upenn.edu/~myl/languagelog/archives/000172.html

Raymond, C. W. (2016). Linguistic reference in the negotiation of identity and action: Revisiting the T/V distinction. *Language*, *92*(3), 636–670.

Renouf, A., & Kehoe, A. (2006). *The changing face of corpus linguistics* (No. 55). Rodopi.

Ritter, E. (1995). On the syntactic category of pronouns and agreement. *Natural Language & Linguistic Theory*, *13*(3), 405–443.

Roberts, I. (2011). Head movement and the Minimalist Program. In *The Oxford Handbook of Linguistic Minimalism.*

Saab, A. (2014). Nominal Ellipses. In J. van Craenenbroeck & T. Temmerman (Eds.), *The Oxford Handbook of Ellipsis.* Oxford University Press.

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

Sankoff, G., & Blondeau, H. (2007). Language change across the lifespan: /r/ in montreal french. *Language*, *83*(3), 560–588. Retrieved from http://www.jstor.org/stable/40070902

Sauerland, U. (1998). *The meaning of chains* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Sauerland, U. (2009). *The presuppositional approach to phi-features.* Center for General Linguistics, Berlin. Retrieved from https://ling.auf.net/lingbuzz/003877

Sigurðsson, H. Á. (2018). Gender at the Edge. *Linguistic Inquiry*, 1–48.

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

Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language & Communication*, *23*(3-4), 193–229.

Simpson, R. C. (1997). Metapragmatic discourse and the ideology of impolite pronouns in Thai. *Journal of Linguistic Anthropology*, *7*(1), 38–62.

Smolensky, P., & Prince, A. (n.d.). Optimality Theory: Constraint interaction in generative grammar. *Optimality*.

Sudo, Y. (2012). *On the semantics of phi features on pronouns* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Szabolcsi, A. (1987). Functional categories in the noun phrase. In *Approaches to Hungarian 2: Theories and analyses* (pp. 167–191). József Attila University.

Trager, E. C. (1956). Superfix and sememe: English verbal compounds. *General Linguistics*, *2*(1), 1.

Travis, L. (1984). *Parameters and effects of word order variation* (Unpublished doctoral dissertation). MIT.

van Gelderen, E. (2013). The diachrony of pronouns and demonstratives. *In search of universal grammar: From Old Norse to Zoque*, *202*, 195.

Van Riemsdijk, H. (2007). Free Relatives. In *The Blackwell Companion to Syntax* (Vol. 2, pp. 338–382).

Vergnaud, J.-R. (1974). *French relative clauses.* (Unpublished doctoral dissertation). Massachusetts Institute of technology.

Weinreich, U., Labov, W., & Herzog, M. (1968). Empirical foundations for a theory of language change.

Wiltschko, M. (2012). What does it take to host a (restrictive) relative clause? *Working Papers of the Linguistics Circle of the University of Victoria*, *21*(2), 100.

Wiltschko, M. (2013). Descriptive relative clauses in Austro-Bavarian German. *Canadian*

Journal of Linguistics/Revue Canadienne de Linguistique, *58*(2), 157–189.

Zimman, L. (2018). Transgender voices: Insights on identity, embodiment, and the gender of the voice. *Language and Linguistics Compass*, *12*(8), e12284.

Zimman, L., & Hayworth, W. (2018). *Lexical change as sociopolitical change in talk about transgender bodies: New methods for the corpus analysis of internet data.* Paper presented at New Ways of Analyzing Variation 47, New York University.

Zobel, S. (2015). Voldemort phrases in generic sentences. *Grazer Linguistische Studien*, *83*, 107–123.