

**RESEARCH**

# Syntactic and cognitive issues in investigating gendered coreference

Lauren Ackerman

Newcastle University, Newcastle upon Tyne, GB  
lauren.ackerman@ncl.ac.uk

---

This paper presents a framework for how the multifaceted nature of “gender” (human and linguistic) interacts with grammatical operations such as coreference dependency formation. It frames the question through the lens of English, in which it focuses on how personal names and referents who identify as nonbinary can provide insight into the conceptual representations of gender. Additional data from a variety of modern languages supports a model of how gender might be cognitively represented such that the observed linguistic patterns are available. I propose a three-tiered model of gender that unites grammatical, cognitive, social, and biological aspects and describes how implications of this model might be tested in future work.

---

**Keywords:** gender; features; coreference; pronouns; personal names

---

## 1 Introduction

### 1.1 Preface

The inspiration for this paper comes from the observation that gender agreement can sometimes follow different criteria cross-linguistically, and more crucially, in different contexts within a language and between individuals. This is not a new observation; this paper elaborates on it by examining typological variation and contextual variation to propose a system for discussing different types of “gender” that are linguistically encoded and how they affect the form that agreement takes. The novel contribution is a framework for integrating general and linguistic cognition as related to gender, broadly construed. This framework will allow theoretical and experimental work in this area to more clearly identify and navigate issues relating to human gender as both a categorical and gradient phenomenon.

To begin, I lay out my proposed terminology for discussing gender in a principled way. This sets the stage for examining data from English, which does not overtly mark gender agreement outside of third person singular pronouns, and comparing it to observations from a variety of other languages that have richer gender inflection systems. I also examine how some lexical innovations which encode nonbinary gender fit into the wider picture of coreference.

Finally, these observations provide the foundation for a proposal which places languages (or, potentially individual speakers of those languages) along a gradient of permissiveness in gender agreement and relates this to how different types of gender, including nonbinary identities, are conceptualized and learned. The intention of this structure is to organize formal, empirical, and philosophical evidence to support the claim that gender is represented and accessed at different levels and to different degrees during the process of coreference resolution.

## 1.2 Gender as a complex phenomenon

The term *gender* is fraught in part because definitions given in the linguistics literature can vary dramatically across subfields or even specific works and are sometimes left as tacit assumptions, even within contexts like coreference resolution. This paper aims to clarify what kinds of gender might be relevant for real-time processing of syntactic agreement and coreference between a pronoun and a referring expression, noting proper names and genders outside of the ‘masculine-feminine’ (or ‘male-female’) binary. It develops the hypothesis that the type of gender involved in coreference checking in English, and possibly other languages, is primarily a domain-general categorical representation of the referent which a formal syntactic or semantic feature can draw upon during agreement and checking operations. That is, the mechanism for categorization of gender which is used to check gender congruency between a pronoun and the expression with which it corefers relies fundamentally on a general cognitive mechanism for classifying and checking congruency of gender rather than relying on a mechanism specific to linguistic processing, in line with the ‘mental model’ framework (e.g., Garnham & Oakhill 1990; Garrod & Terras 2000). Finally, I suggest some lines of research into individual variation that would be able to inform the questions brought up herein.

I will explicitly and precisely define several types of gender in order to provide consistent and unambiguous terminology to the study of coreference and pronouns. These definitions of “gender” include grammatical gender, conceptual gender, gender identity, gender expression, and biosocial gender. I iteratively develop a criterion for checking gender congruency (whether or not two lexical items ‘match’ in gender), then suggest a gradient way in which languages might employ the final formulation of the criterion to result in the typological variation observed. I also describe a three-tiered schema for formalizing the process of gender checking during coreference resolution. While English is the primary focus of this paper, I will demonstrate that motivation for these three categories can be found cross-linguistically. I draw on biological, social, cognitive, and grammatical evidence for how gender is conceptualized and used in human interaction in order to argue that coreference resolution (in English) relies primarily on a non-syntactic property, conceptual gender, for determining whether or not a pronoun and coreferring expression match or mismatch, which is domain-general in origin.

The relative difference in acceptability between sentences (1-a) and (1-b) (indicated by a #) illustrates that English coreference is influenced by discourse-level information and world knowledge. In order to develop a felicitous context for (1-a), one almost must assume the speaker is communicating their disapproval of the referent through misgendering. That is, although the referent’s gender remains ambiguous without further context, a salient interpretation would be that the speaker is intentionally discussing the referent using gendered words (either pronouns or definitional nouns) that are incongruent with the gender identity (defined in Section 2) and wishes of the referent.<sup>1</sup> In contrast, (1-b) provides a context that immediately allows for a felicitous and not necessarily transphobic interpretation since the gender of the costume-wearer and the gender of the pronoun may ‘mismatch’ without qualifying as misgendering the referent, as the costume is intended to mask their gender identity.

- (1) a. #At the farmhouse, the cowgirl<sub>i</sub> left his<sub>i</sub> lasso in the kitchen.  
 b. At the Halloween party, the cowgirl<sub>i</sub> left his<sub>i</sub> lasso in the kitchen.

<sup>1</sup> ‘Misgendering’, or referring to someone in a way that invalidates and devalues their identity, is known to cause mental, emotional and social distress, negatively impacting health and well-being, particularly in adolescents (McLemore 2015; K. Johnson et al. 2019).

The difference in apparent acceptability between these two sentences indicates that the property of gender relevant for coreference is, at the very least, more complex than a formal syntactic feature. This observation by itself is not novel (e.g., Joseph 1979; Hess, Foss & Carroll 1995; Garnham, Oakhill & Reynolds 2002; Duffy & Keir 2004; Nieuwland & Van Berkum 2006; Gygax et al. 2008; Pyykkönen, Hyönä & van Gompel 2010; Collins & Postal 2012; Frazier et al. 2015). Thus this paper develops a formal treatment of how certain “types” of gender can match or mismatch during coreference dependency resolution and what this means for the linguistic encoding of gender identity across languages.

One possible model to explain how gender is conceived and applied to linguistic referents is described in the following sections. It represents a self-consistent, comprehensive model that can be tested empirically. Furthermore, it provides a starting point for interdisciplinary research into the many linguistic facets of gender. In particular, I anticipate this approach will benefit linguistic work which examines phenomena where an individual’s gender identity and/or gender expression is relevant, as well as work which makes use of biosocial gender, including phenotype and hormonal profiles. I especially hope to encourage linguists who make use of psycholinguistic properties of gendered pronouns in their research to be aware of the issues surrounding the various ways in which gender broadly construed and cognition may interface.

## 2 Defining gender

In order to precisely distinguish different types of gender, the following section briefly defines the types of gender relevant to this proposal. These types have been derived from syntactic, semantic, typological, sociological, anthropological, and neuro-biological work on gender. They are not intended to be all-encompassing; rather they are a terminological starting point for a coherent and precise discussion across fields and subfields in which the word *gender* may be used for multiple distinct concepts. The following definitions are elaborated upon in this section.

**Grammatical gender:** The formal syntactic and/or semantic feature that is morpho-syntactically defined. (e.g., Ritter 1993; Comrie 1999; Schriefers & Jescheniak 1999; Harley & Ritter 2002; Kratzer 2009)

**Conceptual gender:** The gender that is expressed, inferred, and used by a perceiver to classify a referent (typically human, but can be extended to anthropomorphized non-humans). (e.g., Newman 1992; Bussey & Bandura 1999; Gygax et al. 2008; Irmen & Kurovskaja 2010; Armann & Bühlhoff 2012; Ansara & Hegarty 2013; McConnell-Ginet 2015)

**Biosocial gender:** The gender of a person based on phenotype, socialization, cultural norms, gender expression, and gender identity. These attributes may conspire to influence conceptual gender and gender expression, but this is an ongoing debate in the field. (e.g., Taylor & J. A. Hall 1982; Waxman 2010; Ansara & Hegarty 2013; Eckert 2014)

**Gender role:** A set of norms conventionalized by society which are associated with clothing or appearance, behavior, preferences, and social expectations. (e.g., Gabriel et al. 2008; Brutt-Griffler & Kim 2018)

**Gender expression:** The way a person appears and behaves, as relating to cultural norms for distinct gender roles. This type of gender can feed into others’ perception, thus into conceptual gender as well. (e.g., Rubin & Greene 1991; Garnham, Oakhill & Reynolds 2002)

**Gender identity:** The mental state of a person regarding that individual’s association with conceptual gender, gender role, gender expression, and biosocial

gender. When grammatical gender referring to a person and the gender identity of that person mismatch, this is likely to be considered ‘misgendering’. (e.g., Ansara & Hegarty 2013; Zimman 2017; K. Johnson et al. 2019)

## 2.1 Grammatical gender

*Grammatical gender* comprises formal morphosyntactic features. They are the properties of words that allows the formal grammatical process of agreement to be carried out. This includes agreement of grammatical gender categories such as *masculine*, *feminine*, *neuter*, *common*, etc.<sup>2</sup> These features are properties of the morphemes themselves, and may be independent from the real-world biosocial genders associated with the referents. However, Corbett (1991) notes that there is a tendency for languages to correlate grammatical gender with the gender of the referent, particularly if human. Moreover, Comrie (2005) adds that there is a tendency for personification of animals and inanimate objects in languages with grammatical gender to correlate with the grammatical gender of the noun phrase. This is further supported experimentally by Konishi (1993), who suggests that perception of inanimate referents are semantically influenced by grammatical gender cross-linguistically. Finally, it may be noted that languages that use different noun classes for subdividing humans almost always divide along a male-female category line independently of how many other noun classes are present or what other types of nouns are included in those two classes. Subdivision of humans across noun classes is a crucial point here, as noun classes that use animacy as a distinction will group humans in the animate category, independent of human gender. I am not aware of any language that encodes more than two human genders grammatically.<sup>3</sup> Even languages of people whose culture encodes more than two human genders do not seem to encode those genders grammatically, as illustrated in Section 3.3.

In (2), the Dagestani language Tsez places animals in a noun class that is distinct from ones that include humans, and Comrie (2005) reports that grammatical gender does not change to reflect the gender roles of personified animals. This contrasts with languages like English, in which personification or anthropomorphization can result in the use of gendered third person pronouns to refer to non-human animals and inanimate objects that would otherwise be referred to with *it*, the inanimate/non-human pronoun. I argue that the variation in use of grammatical gender points to a deeper, more complex system of gender categorization both grammatically and conceptually. That is, grammatical gender is in principle independent from other types of gender but the way it is deployed and the way it influences non-grammatical interpretation suggests it is not entirely decoupled.

The extracts in (2) come from a story in which a rooster (definitionally male) and a hen (definitionally female) are married, but the rooster has another romantic partner (a frog, no specified gender explicitly or grammatically) thus causing strife in the rooster and hen’s relationship (Comrie 2005). Although all animals fall into the third noun class (III) in Tsez, the words for rooster (*mamalay*) and hen (*onoču*) still have defined conceptual or semantic genders despite this not being reflected in the grammatical features. That is, the grammatical gender of the frog, the hen, and the rooster are all obligatorily noun class III, with agreement marked on the verb, which is not used for humans of any gender.

<sup>2</sup> Grammatical gender may include other noun classes as well, although a detailed discussion of noun classes is beyond the scope of this paper.

<sup>3</sup> Kirby Conrod, *p.c.*, suggests that examination of how honorifics are encoded, conceptualized, and learned may provide insight into how gender categories adapt and change over time. Although this is outside the purview of this paper, I suspect that this line of research could potentially be very fruitful. However, it is important to note that honorific systems are much more variable cross-linguistically and also seem to be more susceptible to change over time than gender systems. Still, this comparison warrants further investigation.

- (2) Tsez:
- a. b - oʎix - no ʎohr - ā eʎi - n wit'-wiš ʎin  
 III - appear - PST + CVB frog - ERG say - PST + UNW wit'wish QUOT  
 'The frog appeared and said "witwish".'
- b. onoč - ā b - egir - xo zew - č'ey mamalay  
 hen - ERG III - send - PRS + CVB be - NEG + PST + UNW rooster  
 neɬ - de - r - tow b - ik'i mi yaq'uɬ - no ʎin  
 it- APUD - LAT - EMPH III - go.IMP you today- and QUOT  
 'The hen wouldn't let the rooster in, saying, "Go to her<sup>4</sup> again today".'

In (2), the gender roles of the three characters are inferred through cultural norms, e.g., marriage, and expectations, e.g., housework and romantic liaisons, rather than solely through grammatical gender such as noun class morphology. In the case of *ʎohro* (frog) there is no lexical distinction between the males or females of the species. Thus, the interpretation that the frog is a female interloper in the birds' marriage is not linguistically encoded. Comrie (2005) reports that Tsez speakers uniformly interpret the frog to be female and not male, although it would not be ungrammatical for the frog to be male. Thus, the interpretation of the frog as female must come from the cultural expectations of the speakers rather than from their language.

Compare *rooster* and *hen* in English and Tsez to languages like German (masculine *Hahn* and feminine *Henne*, respectively) and Russian (masculine *petux* and feminine *kurica*, respectively), in which the grammatical gender of the words and the real-world sex of the animals is congruent. In Russian, the word for *frog* (*ljaguška*) happens to be grammatically feminine, thus congruent with the anthropomorphic gender role of the frog character. However, in German the word for *frog* (*Frosch*) is grammatically masculine. Comrie (2005) reports that this makes it difficult, potentially bordering on ungrammatical, to use *Frosch* in translation, since the grammatical gender is incongruent with the anthropomorphic gender role of the frog character. According to him, the way to translate this story without indicating a homosexual relationship between the rooster and the frog would be to change the species of the interloping character to a feminine word like *toad* (*Kröte*). This suggests that the grammatical gender of a word and the gender role of the character are conceptually connected, even though this need not be the case (Konishi 1993; Irmen & Kurovskaja 2010). On the other hand, what might be called grammatical gender in English, which is restricted almost entirely to third person pronouns, appears to be fully coupled to conceptual gender since the pronoun used would determine how the character's gender role is interpreted. This leads us to the question: what role does grammatical gender play in English, if any?

It is unclear whether or not grammatical gender plays a role in English syntactic operations or psycholinguistic processes. It has been argued that English has completely lost grammatical gender, based on historical changes and loss of productive gender morphology (Baron 1971). Certainly, there is no overt gender agreement between nouns, adjectives and articles. However, Bjorkman's recent treatment of gender agreement between names and pronouns makes a case for a limited grammatical gender system in English, in which sentences like (3) display a contrast in acceptability (Bjorkman 2017).

- (3) a. That surgeon<sub>i</sub> operated on three of their<sub>i</sub> patients today.  
 b. ?\*Jonathan<sub>i</sub> operated on three of their<sub>i</sub> patients today.

<sup>4</sup> Here, *her* refers to the frog because in translation to English, *it* would be ambiguous and unnatural.



Bjorkman observes that sentences like (3-a) are more acceptable than (3-b), even when the surgeon is known to all parties, and suggests this is due to names having grammatical gender (i.e., a  $\phi$ -feature) in English, which must then agree with the pronoun, at least for some speakers. A reviewer points out that (3-b)'s acceptability is contextually dependent, as Johnathan's gender identity and the interlocutors' knowledge of this will affect the acceptability of the sentence. For instance, however, consider people like anti-bullying activist Jeffrey Marsh who is nonbinary and whose pronouns are *they/them*, but whose forename is strongly biased as masculine. In this case, it is unlikely that speakers will have a lexical entry for *Jeffrey* that doesn't have a masculine  $\phi$ -feature, but this does not change that Jeffrey Marsh's pronouns are *they/them* and using other pronouns would be misgendering. Speakers would then need to have explicitly acquired the knowledge of which pronouns are appropriate in order to avoid misgendering a person whose gender identity is not immediately inferred from culturally specific cues in gender expression and gender role.

Whether or not English makes use of grammatical gender to determine gender congruency between coreferring elements, an argument for  $\phi$ -features on names must account for how gender (conceptual and/or grammatical) is associated with their referents, since gender bias of names is wildly variable and mutable, more akin to cultural shifts than language change (Barry & Harper 1982; 1993; Van Fleet & Atwater 1997; Lieberman, Dumais & Baumann 2000; Hahn & Bentley 2003; Barry & Harper 2014). Thus, for grammatical gender to play a role in English, it would need to be the case that names and a limited number of nouns have  $\phi$ -features for gender, but that agreement with a coreferring pronoun is optional in cases where the antecedent does not have a  $\phi$ -feature for gender. To this end, I will set aside the status of grammatical gender in English for the time being and return to it in Section 4.1.3.

## 2.2 Conceptual gender

*Conceptual gender* encompasses a large number of closely related terms currently in use in the literature. This includes semantic gender (e.g., Asarina 2009), definitional gender (Kreiner, Sturt & Garrod 2008) and notional gender (i.e. natural gender, but see McConnell-Ginet (2015) for why the term 'natural' is inappropriate), which are ways of associating lexical items with masculine or feminine properties, but without necessarily attributing formal features to them.

This may be illustrated by the strong gender biases of many English occupational terms (e.g., Garnham, Oakhill & Reynolds 2002; Kennison & Trofe 2003; Duffy & Keir 2004; Gygas et al. 2008; Kreiner, Sturt & Garrod 2008). These biases, although in principle mutable, seem to hold consistently and for large swathes of the population. This bias underpins the confusion caused by the "riddle" cited in Reynolds, Garnham & Oakhill (2006) (originally from Sanford 1985: 311):

*A man and his son were away for a trip. They were driving along the highway when they had a terrible accident. The man was killed outright but the son was alive, although badly injured. The son was rushed to the hospital and was to have an emergency operation. On entering the operating theatre, the surgeon looked at the boy, and said, "I cant do this operation. This boy is my son." How can this be?*

The difficulty of interpreting the surgeon as being either the son's mother or any other parental figure besides the previously mentioned father is reflected in the enduring nature of this riddle. In either case, *surgeon* is demonstrated to have a strong male bias despite there being no definitional requirement for surgeons to be men. While gender is not overtly morphologically or grammatically marked in English, there is still some sort of conceptual bias that can be difficult to override.

In Russian, conceptual gender and grammatical gender sometimes clash. Asarina (2009; 2011) observes that *doctor* (*vrach*) is in the first noun class (I), which typically includes human male nouns, among other things. However, when referring to a doctor who is a woman, there are a few strategies that may be employed in different registers.<sup>5</sup> See also King (2015) for another detailed account of mixed agreement in Russian. This is a particularly clear case of a clash between grammatical and conceptual gender because there are two loci that agreement could target and the different structural positions each target a different locus.

The explanation Asarina gives for how Russian can have mixed case agreement is that there is a structural representation of the grammatical feature in the syntax (as opposed to in the semantic representation). This means that an unpronounced functional projection encodes something about conceptual gender. For example, in Russian, there is a functional projection in sentences like (4), i.e.  $\langle wmn \rangle$ , and the agreement is triggered by the closest class feature in the tree, i.e. noun class II. Thus the adjective agrees with the grammatical gender of the noun (masculine/noun class I, because ‘doctor’ is in the first/masculine noun class), but the verb agrees with the conceptual gender of the noun phrase (feminine/noun class II, because the doctor is a woman).

(4) Mixed agreement in Russian where *vrach* (*m*) refers to a woman and possible structural representation, adapted from Asarina (2009):

- a. Zubn- oj vrach prishl- a.  
 dental- M doctor(I) came- F  
 ‘The (female) dentist has come.’
- b.
- 
- $wmn(II)$ 
*prishl-a*  
came-F  
*zubn-oj*
*vrach*  
dental-M
doctor(I)

In this representation, it’s argued that ‘dental’ agrees with ‘doctor’ because the masculine  $\phi$ -feature from *vrach* is the closest target of agreement in the tree, whereas the verb agrees with the (unpronounced) functional head  $\langle wmn \rangle$  as it is the closer target of agreement. This requires the functional head be tied to the discourse context, thus is more flexible and potentially more defeasible than if such a functional head were absent or unavailable in the language. In fact, this type of functional head only seems to be available for human referents and not animals, even when the animals are anthropomorphic (Comrie 2005). This suggests that there is some super-level of categorization in Russian that distinguishes animals and humans even in contexts where animals are filling human-like gender roles. I will set aside the question of distinguishing animals and humans grammatically, but I will also suggest that the categories could be cognitively structured in a manner similar to gender.

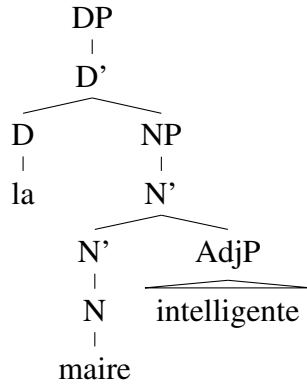
On the other hand, this is not also the case in formal registers of European French.<sup>6</sup> In (5), the form of the noun (masculine) does not change to match the gender of the referent, although this is at least partly for orthographic reasons.

<sup>5</sup> While Asarina does not address how nonbinary conceptual gender could be encoded in Russian, this is an issue which is being explored by nonbinary users of Russian (Wilson 2018).

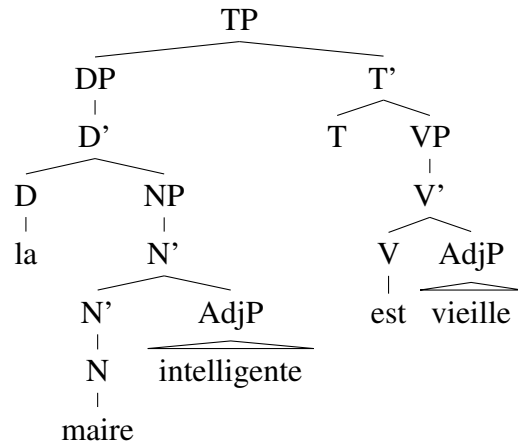
<sup>6</sup> Speakers of Canadian French report the best solution is to use the feminine word *mairresse*. This is purportedly unavailable in formal registers of European French, as it means *the wife of the mayor* rather than the mayor herself. This is also attested as an older definition in Québécois French (Office québécois de la langue française 2017).

(5) Mixed agreement in French where *mayor (m)* refers to a woman and possible structural representation

- a. la maire intelligente  
 det.FEM mayor.MASC intelligent.FEM  
 ‘The intelligent mayor’



- b. la maire intelligente est vieille  
 det.FEM mayor.MASC intelligent.FEM is old.FEM  
 ‘The intelligent (female) mayor is old.’



In formal European French the form of the noun does not change. All gender agreement must match either the grammatical gender of the head noun or the conceptual gender of the referent. Thus, any mixed agreement should only occur when the conceptual gender of the referent mismatches the grammatical gender of the head noun. In this case, the *<wmn>* features Asarina proposed would be located above N but below any of its projections, which is *prima facie* counter-evidence for a syntactic head that governs gender agreement in French.

Responses by Francophone colleagues to my informal queries indicate that mixed agreement in formal French is marginal in some speakers, since there is often an alternative form of the noun that would match the conceptual gender of the referent. Thus, further investigation into the nuances of mixed agreement in French is warranted. Further investigation into agreement with nonbinary conceptual gender will also become a viable line of research, as users of French (much like Russian) are in the early stages of developing gender-neutral or nonbinary grammatical solutions to conceptual gender (Shroy 2016).

Returning to sentence (1) for instance, *cowgirl* is definitionally female, but can be used for a male/masculine referent in certain circumstances. The feminine definition associated with *cowgirl* is thus defeasible, since gender agreement between *cowgirl* and *his* should be impossible if the property being checked is a morpho-syntactically defined  $\phi$ -feature.



This is not incompatible with English having formal gender features for some words, but I argue that it is strong evidence that what is primarily relevant for coreference resolution is not the morphosyntactic feature. This argument will be elaborated upon in Section 4.1, below.

Furthermore, there is evidence from developmental psychology and language acquisition that young children acquire labels for gender categories before they are able to consistently sort people into those categories (Fagot & Leinbach 1993; Welch-Ross & Schmidt 1996; Bussey & Bandura 1999; O'Brien et al. 2000; Zosuls et al. 2009; Waxman 2010; Fausto-Sterling 2012). At this point in development, (at least) two gender categories are present but not enough input has been received to develop a consistent rubric for evaluating the massive variation present in the population. For instance, children may be able to use the proper pronouns for common and canonically gendered referents (e.g., “mommies” and “daddies”) but fail to generalize identification criteria to novel referents that deviate in one or more ways (e.g., men with long hair, women wearing collared shirts) (Taylor & J. A. Hall 1982; Fagot & Leinbach 1993; Armann & Bühlhoff 2012; Ansara & Hegarty 2013). This may indicate that gender categories are developed and refined by repeated exposure to exemplars and top-down societal reinforcement. The acquisition of gender category labels could conceivably support the acquisition of the conceptual categories. I am unaware of any cross-linguistic differences in age of acquisition of gender categories, but should such differences exist, this would support my claim that linguistic labels feed into non-linguistic categorization behaviors.

### **2.3 Biosocial gender**

*Biosocial gender* is, fundamentally, an individual's gender as it is experienced internally. In addressing this type of gender, a few terminological clarifications are necessary. I will assert a distinction between sex and gender, which are widely confounded terms in linguistics and psychology (Cheshire 2002; Ansara & Hegarty 2013). Herein, *sex* refers to biological properties such as karyotype (XX, XY, etc.) and phenotype (e.g., internal and external anatomy, circulating hormonal milieu). Even in biological terms, sex is not a binary property since the physical traits contributing to an organism's sex can vary along multiple dimensions. See Fausto-Sterling (2019) for a recent review. As an example of an edge case, people with Complete Androgen Insensitivity Syndrome (CAIS) may have XY chromosomes but a predominantly female phenotype (e.g., Hughes et al. 2012). However, *sex* is still often used as a shorthand for distinguishing the bimodal nature of the male-female spectrum (Lorber 1996; J. L. Johnson & Repta 2012).

This definition of sex overlaps with biosocial gender. More precisely, biosocial gender is the multidimensional property of an individual as determined by their biology and cultural norms of identity expression. What distinguishes biosocial gender from other types of gender is that, as an external observer, one's accuracy of categorization is impossible to assess without input from the individual's introspection and medical history. That is, biosocial gender may not be something that can be doubtlessly determined without detailed anthropological, introspective and potentially invasive medical analyses. This is because social pressures and societal norms can contribute to an individual representing themselves in a way that is inconsistent with the way they categorize themselves (Fausto-Sterling 2012; Ansara & Hegarty 2013; Zimman 2017). One clear illustration is the case of transgender people who are “in the closet” or otherwise representing themselves as the binary gender category to which they were assigned at birth, despite not identifying as this gender. Here, an individual's biosocial gender might be in direct conflict with the gender with which other people would categorize them, that is, the conceptual gender other people attribute to them.

Our current census data suggests that the majority of people have a gender identity that fall into a bimodal distribution of biosocial genders (0.4% of respondents in a UK survey reported thinking of themselves as a way other than ‘male’ or ‘female’; Glen & Hurrell 2012). But many individuals do not categorize themselves with a discrete binary label, and it would do the science and the individuals a disservice to gloss over the often subtle and diverse variations in gender identity present in the population at large, even within male and female categories (J. L. Johnson & Repta 2012). Despite the potential complications in identifying the precise biosocial gender of an individual, it is still an important factor for phenomena involving social identity and certain physiology relevant to linguistic processes such as auditory brainstem responses (Liu et al. 2017). One’s biosocial gender can affect mental, emotional, and social well-being outcomes, indexical properties of speech, and perception of in-group versus out-group (Rubin & Greene 1991; Zimman 2017; K. Johnson et al. 2019). Therefore, it is important to explicitly define biosocial gender as distinct to ensure it is not confounded during investigation of phenomena associated with either grammatical or conceptual genders.

### 3 Further evidence for distinguishing gender types

#### 3.1 Personal names as antecedents

Personal names comprise a large portion of antecedents used in empirical investigations and syntactic judgments of English coreference, presumably due to their intuitive gender-specificity, although this has been identified as an issue in stimulus design (Kasof 1993; Merritt & Kok 1995; Van Fleet & Atwater 1997; Lieberman, Dumais & Baumann 2000; Gabriel et al. 2008). However, English lacks overt morphological marking on names to unambiguously distinguish a correct assessment of the gender identity of the referent, where a ‘correct assessment’ would result in a conceptual gender that is congruent with the referent’s gender identity. A clear example of this problem is illustrated in (6-a), in which the two given pronouns can corefer with the name equally well in the absence of disambiguating context (such as whether the Taylor in question is Taylor Swift, a woman, or Taylor Lautner, a man. As for Taylor Mason, a nonbinary character played by the nonbinary actor Asia Kate Dillon, (6-b) is the appropriate construction (Dillon 2017), although the processing cost and intuitive acceptability of this linguistic structure, in terms of linguistic judgments, is currently a subject of investigation and may vary in reported ‘acceptability’ (Konnolly & Cowper 2017; Ackerman 2018; Conrod 2018; Prasad, Morris & Feinstein 2018).

- (6) a. On the red carpet, Taylor<sub>i</sub>’s fans screamed to get [his<sub>i</sub>/her<sub>i</sub>] attention.  
 b. On the red carpet, Taylor<sub>i</sub>’s fans screamed to get [their<sub>i</sub>] attention.

One possibility is that the name *Taylor* is stored in the lexicon as discrete entries (e.g., Taylor<sub><masc></sub>, Taylor<sub><fem></sub>). The possibility of the lexicon containing Taylor<sub><nonbinary></sub> is a logical possibility but cannot be discussed in much more detail at this point without introducing speculation because of the current dearth of empirical studies on nonbinary gender perception and its influence on lexical categories. If we consider the two binary grammatical genders, a comprehender may retrieve one of the two entries initially, but have to revise the selection if conflicting information is received at a later time during comprehension. The presence of different lexical entries for each string-identical name, each with a distinct valuation of a gender  $\phi$ -feature, makes testable predictions regarding the learning and application of new lexical entries. One can quickly learn a new name or a new use of a common name, but if extensive previous experience with a common name

(e.g., Michael<sub><masc></sub>) influences the processing of a newly encountered and rare version of the name (Michael<sub><fem></sub>), this might be observable in behavioral or psychophysical measures. If this is the case, it would need to be determined how names most often used by nonbinary people are stored in the lexicon, and if these entries are associated with a specifically nonbinary feature or another configuration of grammatical gender. If names are stored generically with some gender label determined by stereotypicality or statistical probability, then by familiarizing a naïve participant to an uncommon or novel pairing between a name and gender (e.g., a woman named Michael or a boy named Sue), there should still be a detectable processing cost to forming a coreference dependency between the pronoun and name. However, if names instead receive gendered properties from domain-general or world knowledge, then retrieval of the uncommon entry should be facilitated more by the context and less processing cost should be observed (Pyykkönen, Hyönä & van Gompel 2010). See Cai et al. (2017) for examples of how long- and short-term learning can be tested.

Another possibility is that “unisex” or names which are not strongly associated with a particular gender category are morpho-syntactically underspecified for gender (e.g., Taylor<sub><0></sub>), and whatever gender assumptions are made about the referent are done so without reference to the lexicon or morphosyntactic features. However, it is not immediately clear what the implications of this configuration would be or how this could be tested. At the very least, it would be necessary to conduct extensive evaluation of each individual participant’s experience with the target names and gender nonconformity and examine effects from the perspective individual differences (Barry & Harper 1982; 1993; Van Fleet & Atwater 1997; Lieberman, Dumais & Baumann 2000; Barry & Harper 2014).

### 3.2 English as a leader of change

More than just a language of convenience, English has certain properties that allow dissociation of the three proposed types of gender. English marks gender (broadly construed) on its third person pronouns (*she*, *he*), but it does not have consistently overt or productive morphological agreement for gender. Numerous studies demonstrate strong gender biases of certain noun phrases (e.g., *surgeon*, *pilot*, *nurse*, *babysitter*), but these are defeasible which indicates the biases are tied to conceptual gender rather than grammatical gender (Garnham, Oakhill & Reynolds 2002; Kennison & Trofe 2003; Duffy & Keir 2004; Kreiner, Sturt & Garrod 2008; Pyykkönen, Hyönä & van Gompel 2010). Furthermore, English has some remnants of gendered morphology (*actor/actress*, *aviator/aviatrix*) and definitionally gendered nouns (*mother*, *father*, *cowgirl*, *bellboy*). It is conceivable that the morphologically gender-marked words *do* have grammatical gender. At least those marked as *<fem>* are the most likely to have retained grammatical gender in English, as those are distinctly non-default and definitional. As for the definitionally gendered words, I have already demonstrated that it is possible to find contexts where the gender is defeasible. This suggests that these words are not grammatically gendered, or at least the relevant type of gender is conceptual gender and not grammatical gender.

Finally, in cultural terms, English has been at the international forefront of informal, community-based development of nonbinary language and so-called “neopronouns”. Examples of neopronouns include Spivak pronouns introduced by Spivak (1990: xv) and gender variant neologisms described in Centauri (2013), Hord (2016) and Bradley et al. (2019), among others. The combination of linguistic innovation, on-going sociological research, and prominence of media exposure makes the English language uniquely situated (in the present moment) to development and linguistic change regarding gender categories inclusive of nonbinary gender(s) and gender neutrality (Page 2013; Brutt-Griffler & Kim 2018).

### 3.3 Other gender paradigms

Many cultures around the world have established and traditional nonbinary, queer, and third-gender categories. Navajo people called *nádleehí* are traditionally characterized as participating in gendered behaviors of the “opposite sex” (Epple 1998). However, the Western concepts of being ‘transgender’, ‘queer’, or ‘homosexual’ do not quite capture the Navajo cultural concept. To this end, the terms ‘alternate gender’ and ‘two spirit’ have been used to describe *nádleehí*. While these cultural concepts seem to provide potential for investigating concepts of gender categories and language, the Navajo language does not mark grammatical gender on human pronouns. Furthermore, the strategy for speaking about *nádleehí* in English is to use standard binary pronouns in a similar manner to how binary trans men and trans women use English pronouns, and “not neuter pronouns or pronouns specific to *nádleehí*” (Epple 1998: 279).

This seems to be very similar to how Māori culture and language encodes gender outside of the binary (Murray 2003). The terms *whakawāhine* and *whakatāne* are “terms which translate roughly to ‘becoming’ or ‘making’ woman or man, indicating a transcendent or permeable gendered identification” (Murray 2003: 240). However, as in Navajo, Māori grammatical gender does not distinguish conceptual gender on pronouns.

The *hijras* of India are similarly difficult to quantify in Western terms, considering themselves to be “‘deficiently’ masculine and ‘incompletely’ feminine” (K. Hall & O’Donovan 1996: 229). Linguistically, they use the grammatical (and conceptual) gender system of Hindi to express their relationship to their gender identities and their affiliation to the community with a mix of grammatical gender and fluid interaction with binary gender roles.

In Buginese, a language spoken on Sulawesi, Indonesia, by approximately five million people, there are distinct lexical items for each of the five recognized genders, but gender is not otherwise encoded grammatically (Graham 2004). The five genders can roughly be translated into Western concepts as feminine woman (*makkunrai*), masculine man (*oroani*), feminine man (*calabai*), masculine woman (*calalai*), and nonbinary (*bissu*).<sup>7</sup> Importantly, people who identify as *calalai* and *calabai* do not wish to conform to feminine or masculine standards for the *makkunrai* or *oroani*, respectively, but rather have their own standards for gender expression. Furthermore, people who identify as *bissu* are considered to have both masculine and feminine elements in their souls and thus serve spiritual roles in the community (Graham 2004). Much like in Navajo and Māori languages, Buginese does not distinguish pronouns for these five conceptual gender categories.

Generally, pronouns are more likely to mark animacy as a  $\phi$ -feature in these example languages. When a language does mark grammatical gender, nonbinary gender categories can be indicated through shifting use of standard binary gender agreement (e.g., K. Hall & O’Donovan 1996). Investigation of gender perception, category acquisition, and development in other cultural paradigms will bring crucial supplementary information to our understanding of how different types of gender are mentally represented and how they influence each other during linguistic and non-linguistic cognitive behaviors.

## 4 Gender in coreference resolution

Coreference resolution is said to compare the grammatical features of the pronominal element and its candidate antecedent in cases where the parser checks for coreference (Garnham & Oakhill 1990; Garnham, Oakhill & Reynolds 2002). Thus, there must be criteria for what counts as ‘matching’ or ‘mismatching’ in order for a coreference dependency to be resolved

<sup>7</sup> I have taken the liberty of adapting these rough translations away from including terminology such as “female-bodied man” or “masculine female” as these terms can carry negative connotations in English and are more likely to describe gender expressions rather than gender identities.



or rejected. In a case such as (1), restated below in (7), where coreference is resolvable but is not *a priori* congruent, one might expect the apparent mismatch in gender between *cowgirl* and *his* to create a processing slowdown in contexts that do not include clues or information about the referents ahead of time.

- (7) a. #At the farmhouse, the cowgirl<sub>i</sub> left his<sub>i</sub> lasso in the kitchen.  
 b. At the Halloween party, the cowgirl<sub>i</sub> left his<sub>i</sub> lasso in the kitchen.

In (7-a), without knowledge of the context, the conceptual gender of *the cowgirl* and *his* mismatch until a suitable alternative context is imagined. In (7-b), the context of a Halloween party (in which gender roles, expression, and possibly even conceptual categories are expected to be challenged) easily provides the alternative context. The difference, therefore, between (7-a) and (7-b) in terms of acceptability comes from the readers' ability to find a suitable situation in which the conceptual genders match. However, the underlying mechanism for such a prediction is not transparently derivable from syntax-first models of real-time coreference resolution without incorporation of discourse-level knowledge. In what follows, I will set out and incrementally refine a criterion used to evaluate gender congruency in coreference resolution. A strict criterion for matching might look something like this, loosely adapted from definitions of agreement by, e.g., Lasnik & Uriagereka (1988); Payne & Huddleston (2002); Carnie (2007):

**Strict matching criterion:** Matching gender requires the formal grammatical feature ( $\phi$ -feature) of the pronoun to be identical to the candidate antecedent. If the features are not identical, the coreference dependency is rejected.

This strict version of a matching criterion can be rejected immediately because it is insufficient to account for some common, well-described types of coreference. Looking briefly at (7-b), *cowgirl* must either have no  $\phi$ -feature for gender or the  $\phi$ -feature is <fem>, both of which necessarily mismatch with *him*<sub><masc></sub>. Another example of how the strict matching criterion fails is when the antecedent is not explicitly or overtly present in the syntax, e.g., the 'statue rule' (Jackendoff 1992) and "impostors" which are superficially 3<sup>rd</sup> person but conceptually 2<sup>nd</sup> or 1<sup>st</sup> (Collins & Postal 2012), thus the  $\phi$ -features do not directly match:

- (8) a. Regarding a customer (Jackendoff 1992):  
 [The ham sandwich in the corner]<sub>i</sub> needs his<sub>i</sub> bill.  
 b. Spoken to a king (Collins & Postal 2012):  
 [Your majesty]<sub>i</sub> must protect yourself<sub>i</sub>/himself<sub>i</sub>/\*herself<sub>i</sub>/\*themselves<sub>i</sub>.

Even still, in these cases of apparent feature mismatch, some formal level of representation could contain formal features that can be checked during coreference resolution, i.e. what Collins & Postal (2012) term a 'source'. These formal features could be located in either (or both) the semantic and syntactic representations, but the strict definition can only account for the apparent gender mismatch in (7) if we posit that a masculine  $\phi$ -feature is attributed to *cowgirl* only after it is identified as the candidate antecedent of *his*. In order to account for more data, a slightly less strict criterion might be formulated as such, adapted for coreference processing from Collins & Postal (2012: 182):<sup>8</sup>

<sup>8</sup> I have taken liberties in adapting this condition in order to present it in a relatively theory-agnostic manner.



**Less strict matching criterion:** The act of resolving a coreference dependency requires an identity relation between the  $\phi$ -features of a pronoun and either (a)  $\phi$ -features of the antecedent, or (b)  $\phi$ -features of the antecedent as determined by the semantic properties of the notional ‘source’. If the features are not identical, the coreference is rejected.

One reviewer noted that this less strict criterion might account for sentences like in (9-a), where the source of *person*<sub><0></sub> could be *woman*<sub><fem></sub>. If so, it is fairly acceptable due to the reduction in ambiguity from the antecedent to its source (Foraker & McElree 2007). Compare this to (9-b), in which *woman*<sub><fem></sub> could have a source of *person*<sub><0></sub>, creating a noticeable reduction in acceptability, presumably because *woman* is a proper subset of *person* and thus increases ambiguity unnecessarily.<sup>9</sup>

- (9) a. One person<sub>i</sub> said she<sub>i</sub> lost her<sub>i</sub> sunglasses.  
b. ?One woman<sub>i</sub> said they<sub>i</sub> lost their<sub>i</sub> sunglasses.

Yet, this next formulation still might not quite cover the case of (7), where the conceptual gender of the antecedent *cowgirl* is female but the coreference between the masculine pronoun and the (female) antecedent is licit. That is, unless the source is “man<sub><masc></sub> [dressed as a cowgirl<sub><fem></sub>]”, the source could easily also be “rancher<sub><0></sub>” or “party-goer<sub><0></sub>”, which cannot match as they do not have the <masc> feature, which would form an identity relationship with *his*<sub><masc></sub>. Neither does this less strict criterion fully explain (8), in which the antecedents might or might not be interchangeable with sources that have matching gender  $\phi$ -features (a: ✓ “The man<sub><masc></sub>”, ✗ “The customer<sub><0></sub>”; b: ✓ “The king<sub><masc></sub>”, ✗ “The monarch<sub><0></sub>”). This might be accounted for in two ways. First, there might be a way to override the feature checking criteria through modeling the parser as having earlier access to pragmatics and world knowledge (consistent with Sigurðsson 2018), or second, the feature checking process has a broader criterion of what can count as matching. The latter could be formulated as such:

**Broad matching criterion:** Matching gender requires at least one level of the mental representation of gender to be identical to the candidate antecedent in order to match. A conceptual property might include a probabilistic representation of the semantic set of possible referents, but also would be susceptible to environmental context, e.g., pragmatics, world knowledge, or discourse context (Cai et al. 2017; Arnold et al. 2018).

<sup>9</sup> Another interesting point this reviewer notes is that coreference between pronouns seems to require much stricter feature matching than between a pronoun and a referring expression, at least in English, as illustrated in (i).

- (i) a. One person<sub>i</sub> said she<sub>i</sub> lost her<sub>i</sub> sunglasses.  
b. One person<sub>i</sub> said they<sub>i</sub> lost their<sub>i</sub> sunglasses.  
c. \*One person<sub>i</sub> said they<sub>i</sub> lost her<sub>i</sub> sunglasses.  
d. \*One person<sub>i</sub> said she<sub>i</sub> lost their<sub>i</sub> sunglasses.

Since (i-a) and (i-b) are considered acceptable, we can infer that both *she* and *they* can corefer with *one person*. However, mixing *she* and *they* within one sentence and thus one set of coreferring elements causes a noticeable reduction in acceptability. This cannot be due to a mismatch in gender between each of the pronouns and *one person*, as (i-a) and (i-b) demonstrate these are individually acceptable. Therefore, it seems likely that it is coreference between the pronouns that is unacceptable. In this case, I propose that English (or at least the English that is informing these judgments) employs a **Strict matching strategy** as defined in (11) to evaluate coreference between pronouns, but not between a pronoun and a referring expression.

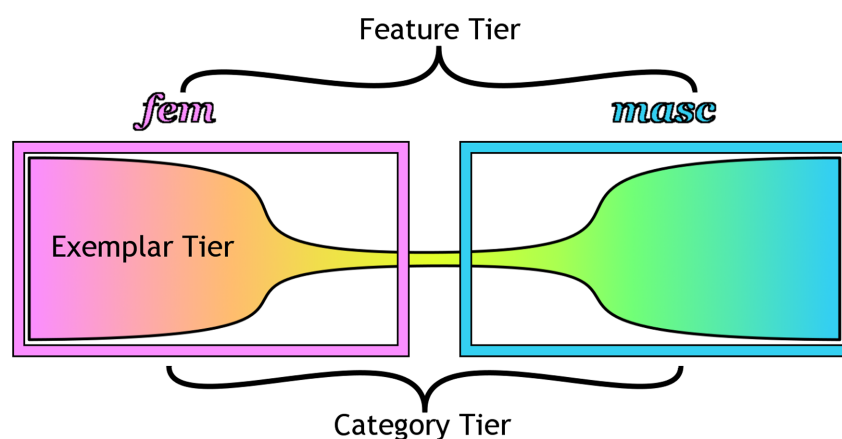
This final formulation can account for (7) as it directly references the conceptual gender of the referents. It can also account for some of the cross-linguistic variation observed in the literature (e.g., Comrie 2005). While one of the stricter formulations would be sufficient to account for data in some languages, the broad criterion allows for language-specific variation in strategy for checking gender matching, which can address language-internal variation and hypothetical change over time. This makes it both powerful and testable, as it still requires a parameter setting or a clearly defined discourse context and theory of gender categories. Languages with very strong or strict matching criteria would then find it difficult to have pragmatic context override the formal gender features of the pronoun (e.g., anaphor, cataphor) which triggered the coreference dependency.

However, all of this assumes that languages that have formal gender features on nominal elements also have formal features that can be checked on the candidate antecedents. What then, would happen if the candidate antecedent didn't have a  $\phi$ -feature in any instantiation? Would this cause a processing slowdown because the initial checking operation would automatically fail? If so, we should expect to see processing slowdowns for coreference dependencies which connect gender-unbiased or undefined antecedents and gender-specific pronouns as compared to coreference dependencies which connect gender-specific antecedents to gender-specific pronouns (cf. Foertsch & Gernsbacher 1997).

With this 'broad' matching criterion, I have shifted the formal problem of typological variation from the process of checking for gender congruency to the type of gender that is checked. This is addressed by the three-tier model illustrated in Figure 1, which provides a formal structure that languages and individuals can use to determine gender congruency all using the same standardized criterion.

#### 4.1 Checking for congruency

If formal morphosyntactic gender features are present in a language like English, but cannot be used to model how the parser checks for congruency in coreference dependency formation, what purpose do they serve? I will not argue for or against English having formal grammatical features for gender, but rather that such features are irrelevant during coreference dependency formation. Instead, English and languages with similar gender systems rely on conceptual gender for evaluating gender congruency in real time. In order to describe how such a system operates, a three-tiered scheme of linguistically and cognitively encoding gender is posited below.



**Figure 1:** A schema depicting the three proposed tiers, overlaid.

The three tiers comprise an exemplar tier, a category tier, and a feature tier (Figure 1). These tiers are not meant to represent actual processing mechanisms or structures in the mind. Rather, they are abstract categories of processes or representations that can be used to map behaviors and empirical observations to theoretical properties of grammars and other mental mechanisms and modules. Thus, each tier is designed to be as theory-agnostic as possible to provide the most utility across the various popular frameworks.

The first tier, the exemplar tier, is represented by a strongly bimodal continuum indicative of how biosocial gender and conceptual gender can vary within a population. Although only color and height vary in this diagram, one may imagine that this tier has many more dimensions that could align with variation in gender role, gender expression, and overt biosocial properties. The second tier, the category tier, comprises two discrete, non-overlapping spaces overlaid on the exemplar tier. These categorically distinct spaces represent the binary genders as might be conceptualized by someone from a society that reinforces a strictly binary gender schema. However, even so, one might not be able to categorize all individuals into one of these spaces, so the gap between the categories allows for ambiguous, nonconforming, and ‘other’ instances to exist outside the binary. If the exemplar tier is, indeed, multidimensional beyond what can be represented on paper, I request that the reader accept that these two categories are not as simple as the rectangles depicted, and their apparent shapes simply due to the limitations of the medium. For instance, if one dimension encodes hair length as a gendered property of appearance, then a man who otherwise fits all other stereotypically masculine traits but has long hair would align predominantly but not completely with the category tier’s binary (masculine) category. Finally, the third tier is the feature tier which, unlike the previous two, comprises labels associated with spaces rather than spaces themselves. In this illustration, the labels are the grammatical features <fem> and <masc>, corresponding to a language that has two noun classes. A language with more noun classes (or fewer) would have a different configuration for the labels.

#### 4.1.1 The exemplar tier

The exemplar tier consists of observations from individual’s exposure to the variety of observable gender expression. This may include tokens of phenotypic variation, non-conformity of gender expression, and variation of cultural norms. Crucially, most individuals will be primarily exposed to other individuals who have unambiguous binary gender expression and thus will have distinctly bimodal input represented in this tier (Fagot & Leinbach 1993; Glen & Hurrell 2012). Individuals who are members of or adjacent to non-conforming or nonbinary communities may have a different distribution of input, especially if exposure occurs during early acquisition of gender categories.

It cannot be that this tier includes the perceiver’s categorization of the gender of the person which they interact with, because that requires a secondary (categorical) behaviour that is crucially not a component of this tier. Instead, the tokens in this tier might be conceptualized as matrices of perceived properties that are used downstream to categorize the gender of the individual. For example, hair length and style, face shape, pitch range of voice, clothing style, sociolinguistically marked properties of speech, etc, could be dimensions of each token. These properties can be used to categorize an individual’s gender (Fagot & Leinbach 1993; Bussey & Bandura 1999; Armann & Bülthoff 2012; Fausto-Sterling 2012; Ansara & Hegarty 2013; Zimman 2017), but are not inherently properties of biosocial genders. Furthermore, few of these properties are purely linguistic, so the parser will not interact with the information stored in this tier. It therefore represents a way of organizing general perceptual input about individuals who a person encounters and interacts with throughout the lifespan.

#### 4.1.2 The category tier

The category tier consists of categories that are established through cognitive processes relying on bottom-up input from the exemplar tier and top-down information from semantics (e.g., gender schema; Bem 1981; Fagot & Leinbach 1993; Bussey & Bandura 1999; O'Brien et al. 2000; Zosuls et al. 2009). The categories of gender encoded in this tier may shift if the distribution of input to the exemplar tier changes. As an individual accumulates more exemplars over the lifespan, each new token will comprise a smaller proportion of the total input, thus will have less influence on the shape of the category tier.<sup>10</sup> The way someone sorts individuals into gender categories should take into account a subset of the dimensions catalogued in the exemplar tier. Whichever way an individual categorizes people into genders and whatever information is used to make those determinations, the category tier holds coarse-grained information about the parameters of each gender category. The structure and robustness of this tier relies on the assumption that gender is most frequently perceived categorically (Fagot & Leinbach 1993; Armann & Bülhoff 2012).

For example, this could manifest as recognition of variance in feminine gender expression and what it means to “self-identify” as having a particular gender (e.g., Zimman 2017). However, humans are still readily able to categorize people based on indices canonically associated with binary gender expressions into categories (leaving aside the accuracy or relevance of these categories) (Bussey & Bandura 1999; Waxman 2010). This suggests that the categorical perception of gender is complex and culturally specific. The details of this perceptual categorization process are beyond the scope of this paper. What remains relevant is that the boundaries of these categories may slightly differ between individuals within a culture or society. Thus the boundaries may differ more between individuals belonging to different cultures or societies.

These categories are not strictly linguistic, but contribute to assessments of whether linguistic meanings are consistent or felicitous when concerning the gender of referents. For instance, when discussing a known person (who is, say, categorized by both interlocutors as *female*), it may be relevant for the comprehension mechanism to refer to the category when assessing the plausibility of statements (Kreiner, Sturt & Garrod 2008; Prasad, Morris & Feinstein 2018).

(10) Did that student<sub>i</sub> email you her<sub>i</sub> follow-up questions yet?

Imagine that the person who uttered (10) was a guest lecturer and doesn't know the referenced student personally. The guest lecturer told the student to email the regular lecturer with any questions and those questions would be forwarded on. Then, when the guest lecturer approaches the regular lecturer to ask about the status of the awaited email, the gender of the student is assumed based on visual and perhaps auditory cues. In this type of situation, the gender of the student may also be important for communicative efficiency if it potentially disambiguates the referent (Newman 1992; Foraker & McElree 2007). However, in English, specifying the student's perceived (i.e., conceptual) gender is always optional, and the choice to include or omit it can be influenced by various social and pragmatic reasons.

The interaction of the exemplar tier and the category tier may generate and assign probabilities of genderedness to gender-biased (or equi-biased) lexical items, including names. In being exposed to instances of *surgeons* or *Michaels*, the tokens that have *surgeon* or the

<sup>10</sup> A reviewer points out that it not be the total cumulative number of tokens that shapes the category tier, but rather more marked, recent or salient tokens might be more heavily weighted in terms of their influence. This seems quite plausible and could potentially be investigated through experimental means, but I will leave this to future works.

name *Michael* as a property fall predominantly into the male category. If this is the mechanism for generating gender stereotyping, then the stereotype would be accessed in one of several ways (that all have the same consequence): An aggregate of all *surgeon/Michael* tokens is assessed as a probability; an individual token of *surgeon* or *Michael* is evaluated for gender category (thus drawn at random from all tokens of *surgeon/Michael*); or the evaluation of gender is assessed at an earlier time and is a property that is rarely updated in the lexicon, independent of the content and structure of the exemplar and category tiers. Crucially, whatever the process for determining gender bias associated with a lexical item, its meaning, or gender plausibility, this information is stored separately from the grammatical information stored in the feature tier.

Speculatively, if an individual were to have a substantial proportion of their lifetime experiences involving nonbinary people, we could assume that the distribution of their personal exemplar tier would not be so bimodal as depicted in Figure 1. This might make the shapes of the category tier more complex, or possibly create discontinuous categories, categories with fuzzy boundaries and other categories besides those designating the masculine and feminine modes of the exemplar tier.

#### 4.1.3 The feature tier

The feature tier consists of discrete  $\phi$ -features or labels which may include *<feminine>* and *<masculine>*, among others. These labels can be mapped one-to-one onto the conceptual categories in the category tier, but need not be. During coreference resolution, whether or not the feature tier is used to determine gender congruency is graded from languages that rigidly rely on the feature tier for coreference evaluation to languages without grammatical gender that do not map separate (grammatical) labels onto human gender categories (see examples in Corbett 2015).

This tier differs from the category tier in that the  $\phi$ -features are strictly linguistic and are formally encoded in the grammar of a language. That is, where the category tier concerns categorization of people and animate gendered referents based on social/cultural norms, the feature tier does not categorize anything: it consists of linguistic labels that are used in purely grammatical operations like agreement. These labels do not need to correspond to human gender (e.g., Bantu noun class systems, etc.), and can apply to inanimate lexical items. Furthermore, they do not apply to the *referents* of the relevant lexical items, but to the lexical items (antecedents) themselves. For instance, languages that have strict gender agreement will ignore the conceptual gender of the referents (category tier) in using grammatical gender to satisfy agreement relations (feature tier). This is elaborated on in Section 4.2.

These tiers are three levels at which the parser could assess gender congruency during coreference resolution. Once a pronoun is linked to a candidate antecedent, the parser may access one of the tiers to check gender congruency (Sturt 2003). If the feature tier does not supply relevant formal features for both lexical items (e.g., if it supplies gender  $\phi$ -features for pronouns but not unisex names or gender-stereotyped nouns in English, cf. Bjorkman 2017), it cannot compare like to like and an identity relationship will not be established. In this case, using the category tier as a holistic congruency assessment would be preferable because, presumably, any referring expression will be located in a category that can provide a property to be assessed against. Speculatively, if the exemplar tier were to have a third mode (e.g., a nonbinary human gender), this might affect the structure of the other tiers and provide organic support for the genesis of novel personal pronouns (e.g., Centauri 2013). That is, the space depicted between the two categories is present to suggest that ambiguous or distinctly nonbinary tokens in the exemplar tier can be accommodated by this model. The presence or increased prominence of these sorts of tokens may



lead the structure of the categories to adapt and develop a new categorical space, which may then provide a distinct space for a novel (grammatical gender) label to designate.

#### 4.2 *Typological evidence*

Together, these tiers describe three levels of encoding of gender, broadly construed, that a language (or an individual) may draw upon in order to determine the gender congruency of a pronoun and candidate antecedent during real-time coreference resolution. In (11), I describe three possible configurations of languages based on the broad criterion and the three tiers of mental representation of gender. These three configurations are points within a hierarchy of how rigidly a language (or individual) adheres to matching the gender encoded on the feature tier to assess gender congruency. While I list languages as examples of these points in the hierarchy, I also suggest that individuals may vary within what an individual language permits. That is, a speaker of French who finds any mixed agreement to be unacceptable would be applying the description of a Strict feature rather than where I have categorized French on the whole (Mixed feature). Similarly, English speakers who find singular *they* difficult to learn or use may be employing more of the ‘mixed’ matching strategy than ‘absent’. Furthermore, English speakers who do use the ‘absent’ strategy may also vary in the shape and adaptability of their category tiers, thus introducing intra-language variation in acceptability.

- (11) **Strict matching strategy:** Languages with no exception to grammatical gender agreement which access only the feature tier during coreference resolution. (e.g., Tsez, possibly German)  
**Mixed matching strategy:** Languages with grammatical gender (to any extent) will start with the feature tier, but draw on the category tier in certain specific contexts, such as when the feature tier is incongruent with referent’s conceptual gender. (e.g., Russian, possibly French)  
**Absent matching strategy:** Languages without grammatical gender do not have labels in the feature tier to be checked, so they must make use of the category tier where gender plausibility and discourse context is concerned. (e.g., Turkish, possibly English)

Tsez exemplifies a strict matching strategy, with (2) demonstrating a rigid grammatical gender system for anthropomorphic animals and other noun phrases (Comrie 2005; Corbett 2015). Thus, no matter what the conceptual genders of the characters in the story are, the agreement is consistent with the morphosyntactic features of the lexical items. This may also be the case for some speakers of French for whom *maire* is necessarily <masc> and *mairresse* (mayor<sub><fem></sub>) is a viable alternative. While I am being careful to avoid a neo-Whorfian claim that language shapes or limits our thought, I think it is reasonable to posit that the categories present in one language and not in another could draw attention to different non-linguistic properties of the members of those categories, thus creating subtle distinctions in the boundaries and shape of the categories. In this way, we might explain how grammatical gender can limit conceptual gender in practical translation without necessarily claiming that German speakers think toads are necessarily feminine (Konishi 1993; Irmen & Kurovskaja 2010). Languages with intermediate strategies like French and Russian would then show some mixed properties wherein formal features are checked during coreference resolution, but may be overridden given contextually appropriate information (Asarina 2009; 2011). Moreover, languages without grammatical gender would then rely entirely on the conceptual categorization of the antecedent to evaluate coreference feasibility.

- (12) Agreement patterns in Russian where *vrač* (*m*) refers to a woman (% = marked in certain registers) (Asarina 2009)
- a. Umnaja vrach prishla  
smart.F doctor(I) came.F
  - b. %Umnyj vrach prishel  
smart.M doctor(I) came.M
  - c. %Umnyj vrach prishla  
smart.M doctor(I) came.F
  - d. \*Umnaja vrach prishel  
smart.F doctor(I) came.M  
'The smart (female) doctor has come.'

Where does English fit into this hierarchy? As it has been claimed that English no longer has grammatical gender (except, possibly on pronouns) (Baron 1971), it might be an absent feature language. However, Bjorkman (2017) suggests that English does have limited use of grammatical gender agreement, particularly when referring to named individuals. If so, we might expect such cases to elicit psycholinguistic/cognitive behaviors that are similar to those observed in languages that make use of the feature tier. However, testing this is made difficult by the limited circumstances in which English could have grammatical gender. The potential environments for detecting grammatical gender in English overlap with environments where conceptual gender (as determined by the category tier) could be an alternative source for checking during coreference resolution. That is, words that could have formal gender features (as Bjorkman suggests, personal names) should also typically receive a gender property from the cognitive gender of the referent, encoded in the category tier.

## 5 Future directions and conclusions

There are myriad ways to test the hypotheses described in this paper. It is my hope that readers will be inspired to use this as a starting point for investigating this relatively new line of research into the links between cognition of gender (as a gradient, nonbinary property) and how gender is encoded linguistically. If definitionally gendered nouns or personal names have formal grammatical gender in English, then there should be a failure in coreference resolution for the link between *cowgirl*<sub><fem></sub> and *his*<sub><masc></sub> in (1)/(7), or *Johnathan*<sub><masc></sub> and *their*<sub><0></sub> in (3). At this stage of processing, the parser may need to draw upon the category tier (rather than feature tier, as it may have originally attempted). This could presumably cause a processing slowdown or electrophysiological effect comparable to one that might be observed for a plausibility mismatch.

Since the anaphor in (1)/(7) is also definitionally masculine/male, in conjunction with the pragmatic context (a Halloween party, in which costumes allow people some flexibility in identity performance), the parser may reassign the gender of the lexical item *cowgirl* in a process similar to that of impostor anaphora (Collins & Postal 2012). This should be detectable in behavioral and psychophysiological measures (e.g., Kuperberg et al. 2003; Nieuwland & Van Berkum 2006; Canal, Garnham & Oakhill 2015). However, the tiered schema I propose predicts that individuals who have extensive exposure to third genders or gender nonconforming communities will have differently shaped exemplar distributions, thus also differently shaped category tiers. If the category tier is shaped in such a way that the boundaries between gender categories are overlapping or 'fuzzy', this may ease the processing cost of reanalysis.

The three types of gender distinguished in this proposal comprise a model for exposure to variance in gender expression, cognition, and linguistic encoding. The model is designed to

be broadly applicable and testable across interfaces of linguistic, cognitive, psychological and sociological work. I describe some applications of the model to psycholinguistic topics and suggest future directions for development. Since forays into research on nonbinary gender are few and recent, the three-tiered model is intended to lead to better informed hypotheses about individual variation related to gender, language processing, and experience. Moreover, nonbinary people often suffer social stigma for their gender identities (McLemore 2015; K. Johnson et al. 2019). This puts empirical studies touching on nonbinary issues in a position to set the standard for ethical and compassionate research on and in conjunction with nonbinary people. This paper provides a set of terminology and the beginnings of a framework from which formal, empirical, and experimental linguistic research on nonbinary issues can grow, while incorporating the varied experiences of the people directly affected by it.

## Abbreviations

<<sub>MASC</sub>> = masculine  $\phi$ -feature, <<sub>FEM</sub>> = feminine  $\phi$ -feature, <<sub>0</sub>> = no gender  $\phi$ -feature, I = declension class I, III = noun class III, APUD = location near, COND = conditional, CVB = converb, ERG = ergative, F = feminine noun class, IMP = imperative, INF = infinitive, LAT = lative, M = masculine noun class, N = neuter noun class, NEG = negative, PST = past, PRS = present, PREP = prepositional, QUOT = quotative, UNW = unwitnessed

## Acknowledgements

I wish to thank Joel Wallenberg, Anders Holmberg, and Bronwyn Bjorkman for their helpful and insightful comments on earlier drafts of this paper and Kirby Conrod for their helpful discussion throughout. I am deeply grateful to all those who attended the conference *They, Hirsself, Em, And You: Nonbinary Pronouns in Research and Practice* (THEY2019) for thoughtful discussions and feedback. I also wish to thank the four helpful reviewers who contributed greatly to the improvement of this manuscript, particularly Daniel Currie Hall and Kirby Conrod, whose thoughtful and constructive comments were indispensable. This project is supported in part by a Wellcome Trust Capital Award to Newcastle (grant ref 092504).

## Competing Interests

The author has no competing interests to declare.

## References

- Ackerman, Lauren. 2018. Being themselves: Processing and resolution of singular (im)personal they. In *The 31st CUNY Conference on Human Sentence Processing*. osf.io/qba7d.
- Ansara, Y. Gavriel & Peter Hegarty. 2013. Misgendering in English language contexts: Applying non-cisgenderist methods to feminist research. *International Journal of Multiple Research Approaches* 7(2). 160–177. DOI: <https://doi.org/10.5172/mra.2013.7.2.160>
- Armann, Regine & Isabelle Bülhoff. 2012. Male and female faces are only perceived categorically when linked to familiar identities—and when in doubt, he is a male. *Vision Research* 63. 69–80. DOI: <https://doi.org/10.1016/j.visres.2012.05.005>
- Arnold, Jennifer E., Iris M. Strangmann, Heeju Hwang, Sandra Zerkle & Rebecca Nappa. 2018. Linguistic experience affects pronoun interpretation. *Journal of Memory and Language* 102. 41–54. DOI: <https://doi.org/10.1016/j.jml.2018.05.002>
- Asarina, Alevtina. 2009. Gender and adjective agreement in Russian. In *The 4th Annual Meeting of the Slavic Linguistics Society*. Zadar, Croatia.
- Asarina, Alevtina. 2011. *Case in Uyghur and beyond*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

- Baron, Naomi S. 1971. A reanalysis of English grammatical gender. *Lingua* 27. 113–140. DOI: [https://doi.org/10.1016/0024-3841\(71\)90082-9](https://doi.org/10.1016/0024-3841(71)90082-9)
- Barry, Herbert & Aylene S. Harper. 1982. Evolution of unisex names. *Names* 30(1). 15–22. DOI: <https://doi.org/10.1179/nam.1982.30.1.15>
- Barry, Herbert & Aylene S. Harper. 1993. Feminization of unisex names from 1960 to 1990. *Names* 41(4). 228–238. DOI: <https://doi.org/10.1179/nam.1993.41.4.228>
- Barry, Herbert & Aylene S. Harper. 2014. Unisex names for babies born in Pennsylvania 1990–2010. *Names* 62(1). 13–22. DOI: <https://doi.org/10.1179/0027773813Z.00000000060>
- Bem, Sandra Lipsitz. 1981. Gender schema theory: A cognitive account of sex typing. *Psychological Review* 88(4). 354–364. DOI: <https://doi.org/10.1037/0033-295X.88.4.354>
- Bjorkman, Bronwyn M. 2017. Singular they and the syntactic representation of gender in English. *Glossa: A Journal of General Linguistics* 2(1). 80: 1–13. DOI: <https://doi.org/10.5334/gjgl.374>
- Bradley, Evan D., Julia Salkind, Ally Moore & Sofi Teitsort. 2019. Singular ‘they’ and novel pronouns: Gender-neutral, nonbinary, or both? *Proceedings of the Linguistic Society of America* 4(1). 36: 1–7. DOI: <https://doi.org/10.3765/plsa.v4i1.4542>
- Brutt-Griffler, Janina & Sumi Kim. 2018. In their own voices: Development of English as a gender-neutral language. *English Today* 34(1). 12–19. DOI: <https://doi.org/10.1017/S0266078417000372>
- Bussey, Kay & Albert Bandura. 1999. Social cognitive theory of gender development and differentiation. *Psychological Review* 106(4). 676–713. DOI: <https://doi.org/10.1037//0033-295X.106.4.676>
- Cai, Zhenguang G., Rebecca A. Gilbert, Matthew H. Davis, M. Gareth Gaskell, Lauren Farrar, Sarah Adler & Jennifer M. Rodd. 2017. Accent modulates access to word meaning: Evidence for a speaker-model account of spoken word recognition. *Cognitive Psychology* 98. 73–101. DOI: <https://doi.org/10.1016/j.cogpsych.2017.08.003>
- Canal, Paolo, Alan Garnham & Jane Oakhill. 2015. Beyond gender stereotypes in language comprehension: Self sex-role descriptions affect the brain’s potentials associated with agreement processing. *Frontiers in Psychology* 6(1953). DOI: <https://doi.org/10.3389/fpsyg.2015.01953>
- Carnie, Andrew. 2007. *Syntax: A generative introduction*. Hoboken, NJ: Wiley-Blackwell.
- Centauri, Widow. 2013. *Gender variant neologisms*. San Diego, CA: San Diego State University dissertation.
- Cheshire, Jenny. 2002. Sex and gender in variationist research. In J. K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), *The handbook of language variation and change*, 423–443. Hoboken, NJ: Blackwell Publishing Ltd. DOI: <https://doi.org/10.1002/9780470756591.ch17>
- Collins, Chris & Paul Martin Postal. 2012. *Imposters: A study of pronominal agreement*. Cambridge, MA: MIT Press. DOI: <https://doi.org/10.7551/mitpress/9780262016889.001.0001>
- Comrie, Bernard. 1999. Grammatical gender systems: A linguist’s assessment. *Journal of Psycholinguistic Research* 28(5). 457–466. DOI: <https://doi.org/10.1023/A:1023212225540>
- Comrie, Bernard. 2005. Grammatical gender and personification. In *Perspectives on language and language development*, 105–114. New York: Springer. DOI: [https://doi.org/10.1007/1-4020-7911-7\\_9](https://doi.org/10.1007/1-4020-7911-7_9)
- Conrod, Kirby. 2018. Pronouns in motion. In *Lavender Linguistics 2018 (LavLang)*. Providence, RI.



- Corbett, Greville. 1991. *Gender (Cambridge Textbooks in Linguistics)*. Cambridge, UK: Cambridge University Press.
- Corbett, Greville. 2015. *The expression of gender*. Vol. 6. Berlin: Walter de Gruyter.
- Dillon, Asia Kate. 2017. *As it happens: CBC Radio*. Interview with Carol Off. DOI: <https://doi.org/10.22233/20412495.1017.28>
- Duffy, Susan A. & Jessica A. Keir. 2004. Violating stereotypes: Eye movements and comprehension processes when text conflicts with world knowledge. *Memory & Cognition* 32(4). 551–559. DOI: <https://doi.org/10.3758/BF03195846>
- Eckert, Penelope. 2014. The problem with binaries: Coding for gender and sexuality. *Language and Linguistics Compass* 8(11). 529–535. DOI: <https://doi.org/10.1111/lnc3.12113>
- Epple, Carolyn. 1998. Coming to terms with Navajo Nádleeh: A critique of “berdache,” “gay,” “alternate gender,” and “two-spirit”. *American Ethnologist* 25(2). 267–290. <http://www.jstor.org/stable/646695>. DOI: <https://doi.org/10.1525/ae.1998.25.2.267>
- Fagot, Beverly I. & Mary D. Leinbach. 1993. Gender-role development in young children: From discrimination to labeling. *Developmental Review* 13(2). 205–224. DOI: <https://doi.org/10.1006/drev.1993.1009>
- Fausto-Sterling, Anne. 2012. The dynamic development of gender variability. *Journal of Homosexuality* 59(3). 398–421. DOI: <https://doi.org/10.1080/00918369.2012.653310>
- Fausto-Sterling, Anne. 2019. Gender/sex, sexual orientation, and identity are in the body: How did they get there? *The Journal of Sex Research* 56(4–5). 529–555. DOI: <https://doi.org/10.1080/00224499.2019.1581883>
- Foertsch, Julie & Morton Ann Gernsbacher. 1997. In search of gender neutrality: Is singular they a cognitively efficient substitute for generic he? *Psychological Science* 8(2). 106–111. DOI: <https://doi.org/10.1111/j.1467-9280.1997.tb00691.x>
- Foraker, Stephani & Brian McElree. 2007. The role of prominence in pronoun resolution: Active versus passive representations. *Journal of Memory and Language* 56. 357–383. DOI: <https://doi.org/10.1016/j.jml.2006.07.004>
- Frazier, Michael, Lauren Ackerman, Peter Baumann, David Potter & Masaya Yoshida. 2015. Wh-filler-gap dependency formation guides reflexive antecedent search. *Frontiers in Psychology* 6(1504). DOI: <https://doi.org/10.3389/fpsyg.2015.01504>
- Gabriel, Ute, Pascal Gygax, Oriane Sarrasin, Alan Garnham & Jane Oakhill. 2008. Au pairs are rarely male: Norms on the gender perception of role names across English, French, and German. *Behavior Research Methods* 40(1). 206–212. DOI: <https://doi.org/10.3758/BRM.40.1.206>
- Garnham, Alan & Jane Oakhill. 1990. Mental models as contexts for interpreting texts: Implications from studies of anaphora. *Journal of Semantics* 7(4). 379–393. DOI: <https://doi.org/10.1093/jos/7.4.379>
- Garnham, Alan, Jane Oakhill & David J. Reynolds. 2002. Are inferences from stereotyped role names to characters’ gender made elaboratively? *Memory & Cognition* 30(3). 439–446. DOI: <https://doi.org/10.3758/BF03194944>
- Garrod, Simon & Melody Terras. 2000. The contribution of lexical and situational knowledge to resolving discourse roles: Bonding and resolution. *Journal of Memory and Language* 42(4). 526–544. DOI: <https://doi.org/10.1006/jmla.1999.2694>
- Glen, Fiona & Karen Hurrell. 2012. *Technical note: Measuring gender identity*. Tech. rep. Manchester: Equality & Human Rights Commission.
- Graham, Sharyn. 2004. It’s like one of those puzzles: Conceptualising gender among Bugis. *Journal of Gender Studies* 13(2). 107–116. DOI: <https://doi.org/10.1080/0958923042000217800>



- Gygax, Pascal, Ute Gabriel, Oriane Sarrasin, Jane Oakhill & Alan Garnham. 2008. Generically intended, but specifically interpreted: When beauticians, musicians, and mechanics are all men. *Language and Cognitive Processes* 23(3). 464–485. DOI: <https://doi.org/10.1080/01690960701702035>
- Hahn, Matthew W. & R. Alexander Bentley. 2003. Drift as a mechanism for cultural change: An example from baby names. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270(S1). S120–S123. DOI: <https://doi.org/10.1098/rsbl.2003.0045>
- Hall, Kira & Veronica O'Donovan. 1996. Shifting gender positions among Hindispeaking hijras. In Victoria Lee Bergvall (ed.), *Rethinking language and gender research*. London, UK: Longman. DOI: <https://doi.org/10.13140/2.1.3369.3760>
- Harley, Heidi & Elizabeth Ritter. 2002. Person and number in pronouns: A featuregeometric analysis. *Language* 78(3). 482–526. DOI: <https://doi.org/10.1353/lan.2002.0158>
- Hess, David J., Donald J. Foss & Patrick Carroll. 1995. Effects of global and local context on lexical processing during language comprehension. *Journal of Experimental Psychology: General* 124(1). 62–82. DOI: <https://doi.org/10.1037/0096-3445.124.1.62>
- Hord, Levi C. R. 2016. Bucking the linguistic binary: Gender neutral language in English, Swedish, French, and German. *Western Papers in Linguistics/Cahiers linguistiques de Western* 3(1). 4: 1–29.
- Hughes, Ieuan A., John D. Davies, Trevor I. Bunch, Vickie Pasterski, Kiki Mastroyannopoulou & Jane MacDougall. 2012. Androgen insensitivity syndrome. *The Lancet* 380(9851). 1419–1428. DOI: [https://doi.org/10.1016/S0140-6736\(12\)60071-3](https://doi.org/10.1016/S0140-6736(12)60071-3)
- Irmen, Lisa & Julia Kurovskaja. 2010. On the semantic content of grammatical gender and its impact on the representation of human referents. *Experimental Psychology*, 367–375. DOI: <https://doi.org/10.1027/1618-3169/a000044>
- Jackendoff, Ray. 1992. Mme. Tussaud meets the binding theory. *Natural Language & Linguistic Theory* 10(1). 1–31. DOI: <https://doi.org/10.1007/BF00135357>
- Johnson, Joy L. & Robin Repta. 2012. Sex and gender. In John L. Oliffe & Lorraine Greaves (eds.), *Designing and conducting gender, sex, and health research*, 17–37. Thousand Oaks, CA: SAGE.
- Johnson, Kelly, Colette Auerswald, Allen J. LeBlanc & Walter O. Bockting. 2019. Invalidation experiences and protective factors among non-binary adolescents. *Journal of Adolescent Health* 64(2). S4. DOI: <https://doi.org/10.1016/j.jadohealth.2018.10.021>
- Joseph, Brian D. 1979. On the agreement of reflexive forms in English. *Linguistics* 17. 519–523.
- Kasof, Joseph. 1993. Sex bias in the naming of stimulus persons. *Psychological Bulletin* 113(1). 140–163. DOI: <https://doi.org/10.1037/0033-2909.113.1.140>
- Kennison, Shelia M. & Jessie L. Trofe. 2003. Comprehending pronouns: A role for word-specific gender stereotype information. *Journal of Psycholinguistic Research* 32(3). 355–378. DOI: <https://doi.org/10.1023/A:1023599719948>
- King, Katherine E. 2015. *Mixed gender agreement in Russian DPs*. Seattle, WA: University of Washington dissertation.
- Konishi, Toshi. 1993. The semantics of grammatical gender: A cross-cultural study. *Journal of Psycholinguistic Research* 22(5). 519–534. DOI: <https://doi.org/10.1007/BF01068252>
- Konnolly, Lex & Elizabeth Cowper. 2017. The future is they: The morphosyntax of an English epicene pronoun. Ms., *University of Toronto*. <https://ling.auf.net/lingbuzz/003859>.
- Kratzer, Angelika. 2009. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40(2). 187–237. DOI: <https://doi.org/10.1162/ling.2009.40.2.187>

- Kreiner, Hamutal, Patrick Sturt & Simon Garrod. 2008. Processing definitional and stereotypical gender in reference resolution: Evidence from eye-movements. *Journal of Memory and Language* 58(2). 239–261. DOI: <https://doi.org/10.1016/j.jml.2007.09.003>
- Kuperberg, Gina R., Tatiana Sitnikova, David Caplan & Phillip J. Holcomb. 2003. Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research* 17(1). 117–129. DOI: [https://doi.org/10.1016/S0926-6410\(03\)00086-7](https://doi.org/10.1016/S0926-6410(03)00086-7)
- Lasnik, Howard & Juan Uriagereka. 1988. *A course in GB syntax: Lectures on binding and empty categories*. Cambridge, MA: MIT Press.
- Lieberson, Stanley, Susan Dumais & Shyon Baumann. 2000. The instability of androgynous names: The symbolic maintenance of gender boundaries. *American Journal of Sociology* 105(5). 1249–1287.
- Liu, Jinfeng, Dan Wang, Xiaoting Li & Ningyu Wang. 2017. Association between sex and speech auditory brainstem responses in adults, and relationship to sex hormone levels. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research* 23. 2275–2283. DOI: <https://doi.org/10.12659/MSM.904651>
- Lorber, Judith. 1996. Beyond the binaries: Depolarizing the categories of sex, sexuality, and gender. *Sociological Inquiry* 66(2). 143–160. DOI: <https://doi.org/10.1111/j.1475-682X.1996.tb00214.x>
- McConnell-Ginet, Sally. 2015. Gender and its relation to sex: The myth of ‘natural’ gender. In Greville G. Corbett (ed.), *The expression of gender*, 3–38. Berlin: De Gruyter Mouton.
- McLemore, Kevin A. 2015. Experiences with misgendering: Identity misclassification of transgender spectrum individuals. *Self and Identity* 14(1). 51–74. DOI: <https://doi.org/10.1080/15298868.2014.950691>
- Merritt, Rebecca Davis & Cynthia J. Kok. 1995. Attribution of gender to a genderunspecified individual: An evaluation of the people = male hypothesis. *Sex Roles* 33(3–4). 145–157. DOI: <https://doi.org/10.1007/BF01544608>
- Murray, David A. B. 2003. Who is Takatapui? Maori language, sexuality and identity in Aotearoa/New Zealand. *Anthropologica* 42(2). 233–244. DOI: <https://doi.org/10.2307/25606143>
- Newman, Michael. 1992. Pronominal disagreements: The stubborn problem of singular epicene antecedents. *Language in Society* 21. 447–475. DOI: <https://doi.org/10.1017/S0047404500015529>
- Nieuwland, Mante S. & Jos J. A. Van Berkum. 2006. When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience* 18(7). 1098–1111. DOI: <https://doi.org/10.1162/jocn.2006.18.7.1098>
- O’Brien, Marion, Vicki Peyton, Rashmita Mistry, Ludmila Hrudá, Anne Jacobs, Yvonne Caldera, Aletha Huston & Carolyn Roy. 2000. Gender-role cognition in three-year-old boys and girls. *Sex Roles* 42(11–12). 1007–1025. DOI: <https://doi.org/10.1023/A:1007036600980>
- Office québécois de la langue française. 2017. *Le grand dictionnaire terminologique*. [http://www.granddictionnaire.com/ficheOqlf.aspx?Id\\_Fiche=2077627](http://www.granddictionnaire.com/ficheOqlf.aspx?Id_Fiche=2077627).
- Page, Ann I. 2013. Shaping the English language: Gender-neutral pronouns in EIL. *Thammasat Review* 16(3). 164–175.
- Payne, John & Rodney D. Huddleston. 2002. Nouns and noun phrases. In Rodney D. Huddleston & Geoffrey Pullum (eds.), *The Cambridge grammar of the English language*. Cambridge, UK: Cambridge University Press.
- Prasad, Grusha, Joanna Morris & Mark Feinstein. 2018. The P600 for singular ‘they’: How the brain reacts when John decides to treat themselves to sushi. In *The 31st CUNY Conference on Human Sentence Processing*. <https://osf.io/2vjyp/>.

- Pyykkönen, Pirita, Jukka Hyönä & Roger P. G. van Gompel. 2010. Activating gender stereotypes during online spoken language processing. *Experimental Psychology* 57(2). 126–133. DOI: <https://doi.org/10.1027/1618-3169/a000016>
- Reynolds, David J., Alan Garnham & Jane Oakhill. 2006. Evidence of immediate activation of gender information from a social role name. *The Quarterly Journal of Experimental Psychology* 59(5). 886–903. DOI: <https://doi.org/10.1080/02724980543000088>
- Ritter, Elizabeth. 1993. Where's gender? *Linguistic Inquiry* 24(4). 795–803.
- Rubin, Donald L. & Kathryn L. Greene. 1991. Effects of biological and psychological gender, age cohort, and interviewer gender on attitudes toward genderinclusive/exclusive language. *Sex Roles* 24(7/8). 391–412. DOI: <https://doi.org/10.1007/BF00289330>
- Sanford, Anthony J. 1985. *Cognition and cognitive psychology*. New York: Basic Books, Inc.
- Schriefers, Herbert & Jörg D. Jescheniak. 1999. Representation and processing of grammatical gender in language production: A review. *Journal of Psycholinguistic Research* 28(6). 575–600. DOI: <https://doi.org/10.1023/A:1023264810403>
- Shroy, Alyx J. 2016. Innovations in gender-neutral French: Language practices of nonbinary French speakers on Twitter. Ms., *University of California, Davis*.
- Sigurðsson, Halldór Ármann. 2018. Gender at the edge. *Linguistic Inquiry* 50(4). 723–750. DOI: [https://doi.org/10.1162/ling\\_a\\_00329](https://doi.org/10.1162/ling_a_00329)
- Spivak, Michael. 1990. *The Joy of TEX: a gourmet guide to typesetting with the AMS-TEX macro package*. Providence, RI: American Mathematical Society.
- Sturt, Patrick. 2003. The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language* 48(3). 542–562. DOI: [https://doi.org/10.1016/S0749-596X\(02\)00536-3](https://doi.org/10.1016/S0749-596X(02)00536-3)
- Taylor, Marylee C. & Judith A. Hall. 1982. Psychological androgyny: Theories, methods, and conclusions. *Psychological Bulletin* 92(2). 347–366. DOI: <https://doi.org/10.1037/0033-2909.92.2.347>
- Van Fleet, David D. & Leanne Atwater. 1997. Gender neutral names: Don't be so sure! *Sex roles* 37(1–2). 111–123. DOI: <https://doi.org/10.1023/A:1025696905342>
- Waxman, Sandra R. 2010. Names will never hurt me? Naming and the development of racial and gender categories in preschool-aged children. *European Journal of Social Psychology* 40(4). 593–610. DOI: <https://doi.org/10.1002/ejsp.732>
- Welch-Ross, Melissa K. & Constance R. Schmidt. 1996. Gender-schema development and children's constructive story memory: Evidence for a developmental model. *Child Development* 67(3). 820–835. DOI: <https://doi.org/10.1111/j.1467-8624.1996.tb01766.x>
- Wilson, Cecil Leigh. 2018. Can you be nonbinary in Russian? *Slavic and East European Journal Blog*. <http://u.osu.edu/seej/2018/10/25/can-you-be-nonbinaryin-russian/>.
- Zimman, Lal. 2017. Trans people's linguistic self-determination and the dialogic nature of identity. In Evan Hazenberg & Miriam Meyerhoff (eds.), *Representing trans: Linguistic, legal and everyday perspectives*, 226–248. Wellington, New Zealand: Victoria University Press.
- Zosuls, Kristina M., Diane N. Ruble, Catherine S. Tamis-LeMonda, Patrick E. Shrout, Marc H. Bornstein & Faith K. Greulich. 2009. The acquisition of gender labels in infancy: Implications for gender-typed play. *Developmental Psychology* 45(3). 688–701. DOI: <https://doi.org/10.1080/00918369.2012.653310>

**How to cite this article:** Ackerman, Lauren. 2019. Syntactic and cognitive issues in investigating gendered coreference. *Glossa: a journal of general linguistics* 4(1): 117.1–27. DOI: <https://doi.org/10.5334/gjgl.721>

**Submitted:** 12 June 2018    **Accepted:** 31 July 2019    **Published:** 25 October 2019

**Copyright:** © 2019 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by Ubiquity Press.

