The definitional dilemma of gender in language

Ben Papadopoulos (*he or they, él o elle*)
University of California, Berkeley

Abstract: Current theories of linguistic gender define any language with a requisite system of nominal classification as having gender, in turn defining languages that manifest gendered distinctions in other ways as genderless. These definitions fail to capture the many ways that binary gender can be encoded in the grammar and lexicon crosslinguistically. Drawing from data on twelve typologically distinct languages, I argue that a reimagined theory of gender in language which centers the concept of social gender must be constructed in order to empirically situate the concept of gender-inclusive language. Data on the realization of gender-inclusivity in different languages is essential in this task, as it identifies the features of language that speakers determine to mark gender and signals the capacity of a gendered language to change.

Keywords: grammatical (morphological) gender, gender-inclusive language, nonbinary gender, language and gender, linguistic typology

El dilema definicional del género en el lenguaje

Resumen: Las teorías actuales del género lingüístico definen cualquier lengua con un sistema requerido de clasificación nominal como una lengua "con género" y las lenguas que manifiestan distinciones de género social de otras maneras como lenguas "sin género". Estas definiciones no logran describir las múltiples maneras en que el género binario puede codificarse en la gramática y el léxico de forma interlingüística. A partir de datos sobre doce lenguas tipológicamente distintas, argumento que se debe construir una nueva teoría del género lingüístico que se centre en el concepto del género social. Tal teoría es necesaria para situar el concepto del lenguaje no binario empíricamente. Los datos sobre la realización del lenguaje no binario en diferentes lenguas son esenciales en esta tarea, porque identifican los rasgos de la lengua que les hablantes determinan que marcan el género y señalan la capacidad de cambio de una lengua "con género".

Palabras clave: género gramatical (morfológico), lenguaje no binario, lenguaje inclusivo, género no binario, lengua y género, tipología lingüística

O dilema definicional do xénero na linguaxe

Resumo: As teorías actuais sobre o xénero lingüístico definen calquera lingua cun sistema obrigatorio de clasificación nominal como unha lingua "con xénero" e aquelas que manifestan distincións de xénero social doutras formas como linguas "sen xénero". Estas definicións non acadan describir as múltiples maneiras nas que o xénero binario pode codificarse na gramática e léxico de forma interlingüística. A partir dos datos sobre doce linguas tipoloxicamente distintas, argumento que se debe construír unha nova teoría do xénero lingüístico que se centre no concepto do xénero social. Tal teoría é necesaria para situar o concepto da linguaxe non binaria empiricamente. Os datos sobre a realización da linguaxe non binaria en diferentes linguas son esenciais nesta tarefa, xa que identifican trazos da lingua que as persoas falantes determinan que marca o xénero e sinalan a capacidade de campo dunha lingua "con xénero".

Palabras chave: xénero gramatical (morfolóxico), linguaxe non binaria, linguaxe inclusiva, xénero non binario, lingua e xénero, tipoloxía lingüística.

1. Introduction

Binary gender is grammaticalized to varying degrees in a majority of the world's languages. While in some, gender is marked only on personal pronouns, in oth-

ers, virtually every part of the grammar and lexicon must be marked as masculine or feminine, leaving behind no opportunities to express gender neutrality or any other gender identity besides *male* and *female*. The appearance of gender in language is well understood by nonbinary, trans, and other gender-nonconforming speakers, whose gender identities often cannot be expressed in normative grammar. Over the past decade, these speakers have popularized methods of expressing their identities in many languages with gendered distinctions. Gender-inclusive language, which can involve the innovation of nonbinary linguistic forms of gender self-expression where none exist prescriptively, identifies the features of language that speakers perceive to mark the gender of the person being referenced and introduces neutral and/or specifically nonbinary alternatives. This form of language change is a necessity for those who have little to no adequate forms of gender self-expression available to them in the languages they speak. It also reveals fundamental truths about the privileging of binary gender in society. The proliferation of gender-inclusive language crosslinguistically reveals binary gender to be a unifying feature of both language and culture.

In formal linguistics, the same understanding of gender in language is not necessarily shared. Linguistic gender is currently defined as a particular feature of morphosyntax. It is not defined in relation to the concept of social gender. At times, linguistic theory recognizes masculine and feminine linguistic gender in the same way as it is recognized among women and queer communities, whom its inequalities affect most, but all too often, it is blind to the many ways that masculine-feminine gendered distinctions can be encoded in language. In this paper, I argue that the emergence of gender-inclusive forms crosslinguistically demands a reanalysis of what linguistic gender is and isn't. Drawing from data on typologically distinct languages that have innovated nonbinary forms of expression, I argue that analyses of gender-inclusive language can improve linguistic theory in two principal ways: by identifying the relationship between social gender and linguistic gender, and by uniting the many ways that gendered distinctions can appear in language. First, I will describe how linguistic gender is currently defined in the literature and compare this definition to data on lan-

guages that form masculine-feminine gender distinctions in different ways. Then, I will argue the importance of elaborating a new theory of gender in language that takes as starting point speakers' identifications of binary gender categories in the grammar.

2. The definitional dilemma of gender in language

Language and gender are inherently related. Language is the one of the primary tools used to construct gender categories in society. It is expressed through socially-meaningful gender identities (e.g. *woman, transgender, agender*) as well as words for people that are socially marked as normatively masculine or feminine (e.g. *brother/sister*, *boyfriend/girlfriend*). It can also be expressed in grammatical categories. Although neutral alternatives exist to varying degrees, when gender is grammaticalized in language, it is most often that those genders are masculine and feminine genders to the exclusion of all others. Yet for a language to be recognized as "having gender" in the current theory, its grammar must be structured in a particular way, and if it is not, it is considered "genderless." In this section, I will outline the definitional dilemma of gender in language, beginning with an exploration of the concept in linguistic theory, and entering into an analysis of how different languages challenge these definitions.

In linguistic theory, the term *linguistic gender* is usually synonymous with *morphological gender*. Colloquially known as *grammatical gender*, morphological gender refers to systems of nominal classification in which all nouns in a given language are assigned to one of two or more classes called *genders* (Corbett, 1991). These classes may indeed be based on (social) gender, or they may be based on other semantic (meaning-based) or formal (form-based) criteria. Canonical definitions of morphological gender enumerate three criteria for its identification:

1. The invariable assignment of all nouns in a given language to one of two or more nominal classes;
2. Reflected in patterns of agreement on dependent elements of the noun;

3. Usually assigned based on formal criteria, such as morphological or phonological form, and/or semantic criteria, like animacy (animate, inanimate), humanness (human, nonhuman), or biological sex and/or social gender (masculine, feminine).

   (Corbett, 1991; Dixon, 1982; Kramer, 2015)

In this way, for a language to be identified as "having gender," *all* of its nouns must have a gender value, and these values must control patterns of morphosyntactic agreement. A language like Spanish conforms to all three of these criteria. It has a system of morphological gender in which there are only two noun classes, explicitly named *masculine* and *feminine*. Every noun in the language, whether animate or inanimate, is assigned to one of these two genders, which control agreement patterns on dependent elements, like articles and adjectives. Gender assignments are based on both formal criteria (i.e. certain nouns are assigned to one gender or the other based on their morphological form) and semantic criteria. This is exemplified by words for people in Spanish, the vast majority of which are assigned to the masculine gender if they refer to men or assigned to the feminine gender if they refer to women (Batchelor & San José, 2010, p. 64). This is usually the case in masculine-feminine morphological gender languages, which may be described as a subtype of nominal classification in which "biological sex" and/or social gender are central to the morphological gender assignment of words for humans. Yet Spanish also demonstrates its own peculiarities considered noncanonical in the same literature. For instance, there exists a subset of invariant nouns like *[el/la] estudiante* 'student' which may be variably assigned masculine or feminine gender dependent upon the gender of to whom they refer, defying the criterion that nouns may only ever have one gender value (Corbett & Fedden, 2018, p. 9). Yet because every noun is forcibly assigned a gender within the masculine-feminine gender system, Spanish is identified as a morphological gender language. English, on the other hand, violates the first criterion entirely, leading many to identify it as a "genderless" language. All nouns in English are not sorted into classes based on semantic or formal

criteria, though in the subset of nouns referring to people, many are so socially marked as normatively masculine or feminine that in normative contexts, they most often appear in complementary distribution according to the perceived or known gender of the person referenced (e.g. *John* is my *brother* v. *Jane* is my *sister*), much like masculine-feminine morphological gender languages. There are even a small number of words that may be identified as feminine-marked based on formal as well as semantic criteria (see Table 2.1).

Table 2.1: Feminine Gender Morphology in English

-ess	**-trix**	**-ette**	**-euse**
seamstress	aviatrix	bachelorette	chanteuse
priestess	dominatrix	brunette	masseuse

(Sources: Baron, 1986; Livia, 2001)

Yet all nouns do not belong to different noun classes, disqualifying English as a morphological gender language according to current definitions. English does, however, conform to the other two criteria. With reference to the agreement criterion, when a person is the subject of discourse, we see agreement patterns reflected on personal pronouns (*he, she, they,* etc.), possessive pronouns (*his* brother*, her* brother, *their* brother, etc.), reflexive pronouns (*himself, herself, themself,* etc.), and as mentioned before, in the choice of words that are socially marked as normatively masculine or feminine (*brother/sister, boyfriend/girlfriend,* etc.). There are other items, like honorifics (*Mr., Mrs./Miss/Ms., Mx.,* etc.), that are also chosen based on the perceived or known gender of the person referenced. These agreement targets may be found in the same sentence and are controlled by the same property of gender that originates from the person referenced. Finally, as to the third criterion, is it obvious that the perceived or known "biological sex" or gender of the person referenced plays a defining role in these phenomena. English is identified as a "pronominal gender language" in which "gender is marked solely on personal pronouns" (Corbett, 1991, p. 12; Audring, 2008, p. 95). In this way, while pronouns may be implicated in patterns of agreement in morphological gender languages, nouns themselves are imagined to have

no gender values in pronominal gender languages, though it is clear that nouns can be implicated in patterns of gender agreement in English.

An analysis of English alone reveals a number of limitations about current definitions of linguistic gender. First, in order for a language to be identified as "having gender," it must have a comprehensive system of nominal classification. Consequentially, all languages with a requisite system of nominal classification are identified as "having gender" whether or not their systems have any basis on "biological sex" and/or social gender, and languages that do not see every noun demonstrating classification and patterns of agreement on dependent elements are identified as "genderless." Yet proponents of gender-inclusive language continue to identify content in languages of different types (including those considered genderless) that is so socially marked as normatively masculine or feminine that many trans and nonbinary people do not identify with it. Second, with reference to the masculine or feminine gender assignment of nouns referring to people, the property of gender does not originate from within the language in a manner divorced from the gender identity of the referent. In masculine-feminine morphological gender languages, it is most often that the assignments of words for people are based on "biological sex" and/or social gender. This semantic basis upon which words are assigned to different morphological genders is the same in masculine-feminine gender languages as it is in languages without a system of morphological gender (as currently defined) that exhibit masculine-feminine gendered distinctions in other ways. Third, the observation that the property of linguistic gender can originate from people, not from language, leads us to examine how gender may be encoded in language in ways that escape definitions of morphological gender. As currently defined, morphological gender alone does not encompass all of the ways that gendered distinctions appear crosslinguistically. To this end, I will describe four crosslinguistic features implicated in gender-inclusive language transformation, describing their potential theoretical contributions with reference to traditional definitions of linguistic gender: morphological gender, pronominal gender, lexical gender, and radical gender. Drawing from typologically distinct languages, I will describe how data

on gender-inclusivity is central in challenging the concept of linguistic gender and how it is identified.

3. The realization of gender in language

3.1 Morphological gender

Only two grammatical features are predicted by linguistic theory to originate the property of gender: pronouns and nouns, though with restrictions. The term *pronominal gender system* is attributed to languages that have gendered distinctions in their personal pronominal systems to the exclusion of all nouns in the language. This divorces the concepts of morphological and pronominal gender completely even though they share fundamental commonalities and often coexist in a single language. Morphological gender languages, especially masculine-feminine morphological gender languages, identify nouns as "having gender," but only if every noun in the language has a gender value that is reflected in patterns of morphosyntactic agreement. This property of gender is imagined to be an invariable property of the noun, yet inflectional gender morphology often operates on the same root to make what is recognized as a single word available in multiple genders. Morphological gender as currently defined does not identify all gender morphology crosslinguistically, nor does it explain all socially gender-marked features of language more broadly. In this section, I place focus on the concept of gender morphology, analyzing languages considered both gendered and genderless for the ways that they do or do not conform to traditional definitions of morphological gender.

Table 3.1.1: Canonical Gender Morphemes in Romance Languages[1]

	MASCULINE	**FEMININE**	**INCLUSIVE**
Spanish	**-o**	**-a**	**-e, -x**
	hijo	hija	hije, hijx
	'son'	'daughter'	'child'
	[ˈi.xo]	[ˈi.xa]	[ˈi.xe], [ˈi.xeks]
Catalan	**(-ø)**	**-a**	**-i**
	advocat	advocada	advocadi
	'lawyer'	'lawyer'	'lawyer'
	[əd.buˈkat]	[əd.buˈka.də]	[əd.buˈka.di]
French	**(-ø)**	**-e**	**-·e**
	étudiant	étudiante	étudiant·e
	'student'	'student'	'student'
	[e.tyˈdjɑ̃]	[e.tyˈdjɑ̃t]	[e.tyˈdjɑ̃ t]
Portuguese	**-o**	**-a**	**-e/-u, -i**
	menino	menina	menine, menini
	'boy'	'girl'	'young person'
	[miˈni.nu]	[miˈni.nɐ]	[miˈni.ne]

(Sources: Papadopoulos, Duarte, Duran, & Fliege, 2022; Papadopoulos, 2022; Duarte, 2022; Valenzuela Sanz et al., 2021; Liberman, 2017; Clendenning-Jiménez et al., 2022; Gaigaia, 2014; Valente, 2020)

Table 3.1.1 displays data from four Romance languages identified as having masculine-feminine morphological gender: Spanish, Catalan, French, and Portuguese. These languages fulfill all of the requisite criteria: they have a system of nominal classification in which every noun is assigned a gender, the gender of nouns controls the realization of morphosyntactic agreement on dependent elements, and there is a semantic basis to the gender assignment of words for people. They also provide a number of challenges to the current theory. First, these languages feature many invariant noun forms that may be variably assigned masculine or feminine gender prescriptively (e.g. *[o/a] concorrente* 'contestant' in

[1]Note: This is not an exhaustive list. Phonological representations and data on inclusive gender in each table are based on extant attestations and are subject to widespread variation.

Portuguese), violating the criterion that all nouns in the morphological system must have invariable gender assignments. An extension of this logic is found in the idea that a single word referring to people may have forms in multiple genders, often only differing by suffixal gender morpheme (e.g. *alumno [M.], alumna [F.]* 'student' in Spanish). These forms are rarely identified as separate words, but rather one word that may be inflected for gender depending on the gender of the referent. The emergence of gender-inclusive forms in these languages both adheres to and complicates traditional definitions of morphological gender. In the same languages, additional gender categories designating gender-inclusivity (gender-neutrality and/or specifically nonbinary gender) are being created, replete with their own canonical personal pronouns and canonical gender morphology. While inclusive genders in Romance differ from masculine and feminine genders in that they do not contain nonhuman nouns (as gender-inclusivity is semantically controlled), they adhere to the three criteria nonetheless. Yet gender-inclusive language also complicates theories of morphological gender in that these additional inclusive genders are not created along a documented trajectory (Corbett, 1991, p. 312-315). Gender-inclusive language also exacerbates the question of where the property of gender originates in personal reference.

With this background, it is crucial to analyze other languages that feature gender morphology but are not defined as morphological gender languages. In these languages, subsets of the lexicon feature morphologically marked masculine-feminine gendered distinctions, yet not every noun belongs to system-wide classes identifiable through patterns of morphosyntactic agreement. Some of these languages, like English and Danish, used to have system-wide masculine-feminine (or masculine-feminine-neuter) morphological gender. English is largely recognized to have lost its system of morphological gender, while Danish's former masculine-feminine-neuter gender system is largely recognized to have collapsed into a common-neuter gender system wherein formerly masculine and feminine nouns are now common gender in most dialects. In both of these languages, there exist subsets of words with feminizing or matrimonial suffixes that may or may not survive from older varieties (see Table 3.1.2):

Table 3.1.2: Feminine Forms in English and Danish

English	**-ess**	**-ix**	**-ette**
	actress	aviatrix	suffragette
	waitress	dominatrix	rockette
Danish	**-esse**	**-inde**	**-ske**
	prinsesse 'princess'	veninde 'friend [F.]'	samleverske 'co-habitee [F.]'
	baronesse 'baroness'	værtinde 'hostess'	naboerske 'neighbor [F.]'

(Sources: Baron, 1986; Livia, 2001; Allan *et al*., 2000)

These languages provide an important complication to theories of morphological gender in that neither is imagined to "have gender" under current class-based analyses. Yet it is undeniable that both have traces of feminine gender morphology.

Table 3.1.3: Morphological Gender in Tagalog Loanwords

MASCULINE	**FEMININE**	**MASCULINE**	**FEMININE**
iho 'son'	iha 'daughter'	abarisyoso 'greedy'	abarisyosa 'greedy'
lolo 'grandfather'	lola 'grandmother'	direktór 'director'	direktóra 'director'

(Sources: Blazado, 2022; Oficina de Educación Iberoamericana, 1972)

There exist more blatant examples (see Table 3.1.3). Prior to Spanish colonialism, Tagalog did not have a system of morphological gender. It was only after Spanish colonialism that hundreds of Spanish loanwords were introduced into Tagalog, and many of these words that refer to people retained inflectional Spanish gender morphology (Oficina de Educación Iberoamericana, 1972). As a result, many Tagalog words have masculine-feminine gender in much the same way that Spanish does, though only in a subset of the lexicon. Some of these words are especially prominently in the language, including *Pilipino/Filipino* and *Pilipina/Filipina*, which designate people from the Philippines. Another resemblance to Spanish is found in the fact that speakers have begun to neutralize this gendered distinction by using the term *Pilipinx/Filipinx* (FIERCE, 2018). The *-x*

morpheme is characteristic of gender-inclusive Spanish and is now being used in Tagalog in much the same way that it is used in Spanish. The discovery of gender morphology in languages considered genderless leads us to posit that not even current definitions of morphological gender adequately account for all gender morphology, let alone other gendered features of language.

3.2 Pronominal gender

The issue of access to gender-inclusive language is often emblematized by the use of nonbinary personal pronouns. Pronoun culture, wherein speakers ask their interlocutors what pronouns they use instead of simply assuming, has been normalized in many English-language queer community settings, making personal pronouns perhaps the most visible feature of English gender-inclusive language activism. It is ironic, then, that English and other languages with pronominal gender are disqualified as "having gender" linguistically if they lack a system of morphological gender as currently defined. Personal pronouns hold a privileged place in discussions of what linguistic gender is and isn't: similarly to morphological gender, they represent an embedded grammatical category and are frequently referred to as a closed class crosslinguistically. They are also a widespread feature of gender in language. Many typologically distinct languages show masculine-feminine gendered distinctions in their personal pronominal systems, as shown in Table 3.2.

Table 3.2: Masculine, Feminine, and Inclusive Third-Person Singular Personal Pronouns

	MASCULINE	FEMININE	INCLUSIVE
Spanish	él	ella	elle, ellx
	[el]	[ˈe.ja]	[ˈe.je], [ˈe.jeks]
Catalan	ell	ella	elli, ellx
	[ˈeʎ]	[ˈe.ʎa]	[ˈe.ʎi], [ˈe.ʎɛks]
French	il	elle	iel
	[il]	[ɛl]	[jɛl]
Portuguese	ele	ela	elu, ile, ili
	[ˈe.li], [ˈe.le]	[ˈɛ.lɐ]	[ˈe.lu], [ˈi.li]
Hebrew	הוּא	הִיא	הֶא
	[hu]	[hi]	[hɛ̞]
Irish	sé	sí	siad
	[ʃe]	[ʃi]	[ʃiəd̪ˠ]
English	he	she	they, xe
	[hi]	[ʃi]	[ðeɪ], [zi]
Danish	han	hun	de, hen
	[hæn]	[hun]	[di], [hɛn]
Mandarin	他	她	TA, X也, 无也
	[ta˧]	[ta˧]	[ta˧]

(Sources: Papadopoulos, Bedin, Clendenning-Jiménez, & Miller, 2022; Duarte, 2022; Knisley, 2020; Clendenning-Jiménez *et al.*, 2022; Gross & Rivlin, 2021; Colleluori, 2022; Papadopoulos, Duarte, Duran, & Fliege, 2022; Hjorth-Nebel Miltersen *et al.*, 2022; Tang *et al.*, 2022)

In masculine-feminine gender languages like Spanish, Catalan, French, Portuguese, Hebrew, and Irish, masculine and feminine personal pronouns are often directly related to masculine and feminine morphological genders. For instance, the third-person singular feminine personal pronouns in many of these languages (e.g. *ella* in Spanish) share the canonical gender morpheme of the language's feminine gender (e.g. *-a*). Many of the nonbinary personal pronouns that have been popularized in the same languages follow a similar pattern: they overtly feature the suffixal gender-inclusive morphemes that have been popularized, and may similarly be identified as belonging to distinct nonbinary grammatical genders. An illustration of this is found in the innovative nonbinary Spanish personal pronouns *elle* and *ellx*, which are used together with *-e* and *-x* inclusive

genders in nonbinary reference. Nonbinary pronouns in masculine-feminine morphological gender languages come from a variety of sources: in Spanish, Catalan, French, Portuguese, and Hebrew, they are all innovative, meaning that they do not exist in the extant inventory of the language. The French pronoun *iel* is recognized to be a combination of the masculine and feminine personal pro-nouns *il* and *elle*. Spanish, Catalan, and Portuguese nonbinary pronouns are based on innovative gender morphemes (e.g. *ellx* (*-x*) in Spanish, *elli* (*-i*) in Catalan, *elu*, *ile* (*-u*/*-e*), in Portuguese). In Hebrew, the nikkud sign *segol* (IPA: [ɛ̟]) has been used to phonologically distance nonbinary personal pronouns (הֶא) from masculine (הוּא) and feminine (הִיא) personal pronouns. In Irish, nonbinary speakers have borrowed from the extant plural paradigm of personal pronouns (*siad*) to collapse a gendered distinction in the singular paradigm.

Other languages that lack a system of morphological gender, like English, Danish, and Mandarin Chinese have normatively masculine and feminine personal pronouns that represent one of the only grammatical categories in the language to have masculine-feminine gender distinctions. In the case of English, one gender-neutral/gender-inclusive pronoun comes from within the extant inventory of the language (*they*), and others are neopronouns (innovative, non-prescriptive pronouns) created specifically for the purpose of expressing nonbinary gender identities (e.g. *xe*). Other languages that have no such extant gender-neutral forms, like Danish, have borrowed from the plural paradigm (*de*) or innovated other gender-inclusive forms (e.g. *hen*). In Mandarin Chinese, as the masculine-feminine gender distinction exists only orthographically in gen-der-marked radicals, the pinyin (TA) is used, and the semantic radicals 无 ('none, without') and X, which come from within and outside of the extant radical in-ventory, respectively, replace gender-marked radicals in the nonbinary personal pronouns 无也 and X也.

The innovation of nonbinary pronouns in languages with normatively masculine and feminine gendered pronouns, regardless of their morphological gender status, demands that we accept a broader definition of linguistic gender

wherein gender is not exclusively a property of the noun. Pronominal gender may be unified with other masculine and feminine gender-marked features of language to account for how the gender identity of the speaker affects similar targets in languages of different morphological types. The emergence of nonbinary pronouns also serves to deconstruct the notion of a closed class. In languages that have exhibited linguistic innovation in their personal pronominal systems, it is evident that there is a practice of expanding the number of genders available, as in morphological gender languages.

3.3 Lexical gender

In languages considered both gendered and genderless, the formation of masculine-feminine gendered distinctions in noun pairs sometimes cannot be traced back to inflectional gender morphology, or what appears to be inflectional gender morphology (see Table 3.3). There exist many pairs of words like *uncle* and *aunt* in English that are so normatively marked as masculine or feminine socially that in normative contexts, they are used in complementary distribution with other masculine- and feminine-marked words, like personal pronouns. I refer to these words as lexical gender items to distinguish them from the concept of inflectional gender morphology. The lack of a prescriptive gender-neutral alternative for many of these pairs reveals that maximally binary masculine-feminine distinctions in a language can be constructed in ways distinct from inflectional gender morphology. Processes of linguistic innovation have provided gender-inclusive alternatives for some of these pairs in English (e.g. *niece/nephew/nibling*). In Vietnamese, a language also considered genderless, there exist forms like *ông* 'older male' and *bà* 'older female' that are featured in other kinship terms to pro-duce gendered distinctions (e.g. *ông nội* 'grandfather' and *bà nội* 'grandmother'). What these processes suggest is not only that masculine and feminine gender may be marked on roots and form binary gender distinctions by way of seman-tics only, but also that gender-marked roots themselves can gender-mark other words through processes of compounding.

Table 3.3: Lexical Gender in Spanish, Vietnamese, Irish, and Korean

	MASCULINE	FEMININE		MASCULINE	FEMININE
Spanish	yerno 'son-in-law'	nuera 'daughter-in-law'	**Vietnamese**	ông nội 'grandfather'	bà nội 'grandmother'
	varón 'male'	hembra 'female'		bố 'father'	mẹ 'mother'
Irish	fear 'male'	bean 'female'	**Korean**	아버지 'father'	어머니 'mother'
	mac 'son'	iníon 'daughter'		오빠 'older brother'	언니 'older sister'

(Sources: Papadopoulos, Duarte, Duran, & Fliege, 2022; Ha, 2022; Colleluori, 2022; Sim, 2022)

A similar phenomenon is found in masculine-feminine gender languages. In Spanish, there exist many pairs forming maximally binary masculine-feminine distinctions like *varón* 'male' and *hembra* 'female' that have completely different roots and are not entirely the result of inflectional gender morphology. In Irish, there exist pairs like *mac* 'son' and *iníon* 'daughter' that are not formed by processes of inflectional gender morphology even in part. The most challenging aspect of these nouns to capture empirically is the fact that the social construction of their meaning as male-specific or female-specific in normative situations is not related to any linguistic process like morphological gender inflection that would describe their language-internal formation. Lexical gender nouns force us to accept as fact that words referring to people are identified as normatively masculine or feminine based on their socially-constructed meaning. This is not dissimilar from the basis of masculine and feminine morphological genders on binary social gender, an extralinguistic property.

3.4. Radical gender

Least studied of all is what I term *radical gender*. In Mandarin Chinese, the majority of characters are composed of both semantic and phonetic radicals, the first of which contributes meaning and the second of which informs the character's phonological form (Li & Zhou, 2007, p. 622). There exist two distinct semantic radicals in the language frequently identified as the 'male' or 'human' radical (亻) and the 'female' or 'woman' radical (女) (Chin & Burridge, 1993). These radicals

form masculine-feminine distinctions in some pairs, and in others, the word in the pair featuring the 'woman' radical has some sort of derogatory meaning, as seen in Table 3.4.

Table 3.4. Radical Gender in Mandarin Chinese

In Kinship Terms		**As Pejorative**	
爸爸	妈妈	仅	奴
'father'	'mother'	'only'	'slave'
儿子	女儿	伎	妓
'son'	'daughter'	'skill'	'prostitute'

(Source: Tang *et al*., 2022)

These radicals are also featured in third-person personal pronouns, though there did not used to be a masculine-feminine gendered distinction in the personal pronoun paradigm. Prior to the May Fourth Movement of the late 1910s, the only third-person singular personal pronoun in the inventory of the language was 他. The semantic radical featured in this character, now identified as the 'male' radical (亻), did not used to be normatively gender-marked, as in the meaning 'human'. The feminine third-person singular personal pronoun 她 was introduced by feminists during this period as a way to increase the visibility of women in written Mandarin. These two personal pronouns share one phonological form, so gender-inclusive forms have focused on neutralizing the gendered distinction orthographically. This has been done by substituting the Chinese character for its pinyin form (TA) and by substituting the semantic radical in the extant characters (Mair, 2013). Similar to Spanish and Catalan, Mandarin Chinese speakers have proposed the use of an innovative X radical that does not come from within the extant radical inventory of the language to neutralize the gendered distinction between masculine (他) and feminine (她) personal pronouns in a new pronoun (X也; Lai, 2020). Another innovative pronoun replaces the gender-marked semantic radical with 无, the rad-ical meaning 'none' or 'without' (Zhu, 2021). The use of these radicals has been proposed to collapse other masculine-feminine gendered distinctions formed by the gendered radicals 亻 and 女.

That Mandarin Chinese is a language that largely lacks inflectional processes yet still manifests normatively masculine and feminine gendered distinctions in its characters is further evidence that these distinctions are sometimes formed in ways not unified by current theories of linguistic gender. It reminds us that a theory explaining the realization of masculine-feminine gender distinctions crosslinguistically must be sensitive to typological and featural differences between languages, and it must also describe how these features may be found in combination.

4. A new theory of gender in language

The four features I have described (morphological gender, pronominal gender, lexical gender, and radical gender) are found in typologically distinct languages, and they are not unified by the current theory of gender in language. The concept of linguistic gender as currently defined identifies systems of nominal classification that may or may not have any basis on (social) gender to the exclusion of almost all other gendered distinctions in language. The emergence of gender-inclusive language reveals a different understanding of linguistic gender: binary gender can be grammaticalized and lexicalized in ways that transcend current definitions of morphological gender. By attending to semantics, linguistic features marking subjects as masculine or feminine can be transformed to become inclusive of people of other genders. A new theory of linguistic gender would separate all other forms of nominal classification having no basis in "biological sex" and/or social gender from the features of language that distinguish socially-meaningful gender categories, whether or not those languages have a system of morphological gender as currently defined. That all languages with a system of nominal classification are aggregated and identified as "having gender," whether or not they have any basis on "biological sex" and/or social gender, is evidence used to deny the basis of linguistic gender in any real-world reality. This claim has been weaponized by institutions of prescriptive language to condemn feminist and gender-inclusive language practices that aim to visibilize women and people who are gender-nonconforming. In this way, theo-

ries of what linguistic gender is and isn't have material effects on human rights: namely, nonbinary, trans, and other gender-nonconforming speakers' access to gender-inclusive language.

Gender-inclusive language demands that we accept that gender identities may be encoded in language in many different ways, often in ways that preclude the expression of nonbinary gender identities. Maintaining the name *linguistic gender* is important to this argument. There is evidence to show that Roman grammarians understood the basis of masculine and feminine morphological genders on "biological sex" in their overlapping use of the Latin words *genus* 'kind' and *sexus* 'sex' to describe the linguistic system (Corbeill, 2015). A reanalysis of linguistic gender that maintains separate the concepts of gendered distinctions in language and nominal classification, describing both their interconnection and their differences, would serve to empirically legitimize the necessity of gender-inclusive forms in languages with masculine-feminine gendered distinctions. The study of gender-inclusive language is central in this task for two main reasons: first, it helps to reestablish the concept of linguistic gender as the encoding of gender identities or normatively gendered meanings in the grammar and lexicon of different languages, as understood from the perspective of gender-nonconforming speakers who identify and confront masculine-feminine gendered distinctions in language. Within the realm of nominal classification, this would isolate masculine-feminine morphological gender systems and other masculine and feminine gender morphology, proving these features' overall basis on "biological sex" and/or social gender in words for people with the identity-based innovation of additional gender categories in the grammar. Second, it would identify the gendered features of language distinct from morphological gender in languages currently considered both gendered and genderless, demonstrating that gendered distinctions can manifest themselves in other ways. To this end, we must take seriously this form of language change and allow it to inform linguistic theory. Reanalyzing the concept of linguistic gender from the perspective of social gender is necessary to empirically explain the rise of gender-inclusive language crosslinguistically. It is also necessary to assert the necessity of accepting

gender-inclusive language among institutions of prescriptive language, who control popular perceptions about "good" and "bad" forms of speech and attempt to control the shape of language itself. In this way, a new theory of gender in language is as vital to linguistic theory as it is to the material realities of gendered language speakers, for whom there is no definitional dilemma, only a literature that underdescribes their reality.

Acknowledgements

I acknowledge with gratitude the founding members of the Gender in Language Project (genderinlanguage.com): Gabí Agramont-Justiniano, Cooper Bedin, Carmela Blazado, Sol Cintrón, Sebastian Clendenning-Jiménez, Keira Colleluori, Jesus Duarte, Julie Duran, Chandler Fliege, Julie Ha, Zaphiel Kiriko Miller, Ina Nierotka, Kalinda Reynolds, Serah Sim, Chelsea Tang, Xuedi Yang, and Irene Yi. I would also like to thank Ártemis López, Ehm Hjorth-Nebel Miltersen, Justin Davidson, and the attendees of the 27th Annual Lavender Languages and Linguistics Conference who heard me present this paper and offered their feedback.

References

Allan, R., Holmes, P. & Lundskær-Nielsen, T. (2000). *Danish: An essential grammar.* Routledge.

Audring, J. (2008). Gender assignment and gender agreement: Evidence from pronominal gender languages. *Morphology,* 18, 93-116.

Baron, D. (1986). *Grammar and gender*. Yale University Press.

Batchelor, R. E. & San José, M. Á. (2010). *A reference grammar of Spanish*. Cambridge University Press.

Blazado, C. (2022). *Tagalog*. Gender in Language Project. genderinlanguage.com/tagalog.

Chin, N. B. & Burridge, K. (1993). The female radical: Portrayal of women in the Chinese script. *Australian Review of Applied Linguistics,* 10(1), 54-85.

Clendenning-Jiménez, S., Miller, Z. K. & Agramont-Justiniano, G. (2022). *Global Portuguese*. Gender in Language Project. genderinlanguage.com/portuguese

Colleluori, K. (2022). *Irish*. Gender in Language Project. genderinlanguage.com/irish

Corbeill, A. (2015). *Sexing the world: Grammatical gender and biological sex in Ancient Rome*. Princeton University Press.

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

Corbett, G. G. & Fedden, S. (2018). New approaches to the typology of gender. In S. Fedden, J. Audring & G. G. Corbett (Eds.), *Non-Canonical Gender Systems* (pp. 9-35). Oxford University Press.

Dixon, R. M. W. (1982). *Where have all the adjectives gone? And other essays in semantics and syntax*. Walter de Gruyter & Co.

Duarte, J. (2022). *Catalan*. Gender in Language Project. genderinlanguage.com/catalan

Fierce. (2018). *Why we say Filipinx*. Filipinxs in Education Reaffirming Community Empowerment. kp-fierce.weebly.com/blog/why-we-say-filipinx.

Gaigaia. (2014). *Uso do 'i' como paradigma neutro em português*. Google Docs. docs.google.com/document/d/1KMEm9qe6G4dV8BH4seJHPzyu1e8SM9xUPrFWomzw8iQ.

Gross, L. & Rivlin, E. (2021). *Grammar and systematics*. Nonbinary Hebrew Project. www.nonbinaryhebrew.com/grammar-systematics.

Ha, J. (2022). *Vietnamese*. Gender in Language Project. genderinlanguage.com/vietnamese

Hjorth-Nebel Miltersen, E. (2018). De, den, hen, and the rest: A pilot study of the use of gender-neutral and nonbinary/genderqueer pronouns in Danish. *Journal of Language Works,* 3(1), 31-42.

Knisley, K. A. (2020). Le français non-binaire: Linguistic forms used by nonbinary speakers of French," *Foreign Language Annals 2020*, 1-27.

Kramer, R. (2015). *The morphosyntax of gender*. Oxford University Press.

Lai, C. (2020). *X也'and 'ta': The gradual rise of gender-neutral pronouns in Chinese.* Ariana Life. arianalife.com/topics/gender-equality/x也and-ta-the-gradual-rise-of-gender-neutral-pronouns-in-chinese/

Li, J., & Zhou, J. (2007). Chinese character structure analysis based on complex networks. *Physica A: Statistical Mechanics and its Applications* (280), 629-638.

Liberman, M. (2017). *Écriture inclusive*. Language Log. languagelog.ldc.upenn.edu/nll/?p=34912.

Livia, A. (2001). *Pronoun envy: Literary uses of linguistic gender*. Oxford University Press.

Mair, V. (2013). *He/she/it/none of the above*. Language Log. languagelog.ldc.upenn.edu/nll/?p=4576.

Oficina de Educación Iberoamericana. (1972). *Hispanismos en el tagalo*. Imnasa.

Papadopoulos, B. (2022). A brief history of gender-inclusive Spanish/Una breve historia del español no binario. *Deportate, esuli, profughe,* 48(1), 31-48.

Papadopoulos, B., Bedin, C., Clendenning-Jiménez & Miller, Z. K. (2022). *Global English*. Gender in Language Project. genderinlanguage.com/English

Papadopoulos, B., Duarte, J., Duran, J. & Fliege, C. (2022). *Global Spanish*. Gender in Language Project. genderinlanguage.com/spanish

Sim, S. (2022). *Korean*. Gender in Language Project. genderinlanguage.com/korean

Tang, C., Yi, I. & Bedin, C. (2022). *Mandarin Chinese*. Gender in Language Project. genderinlanguage.com/chinese

Valenzuela Sanz, A., Alonso Vinués, C. & Ordóñez Castellnou, M. (2021). *Guia pràctica de lleguatge inclusiu*. Chrysallis: Asociación de familias de menores trans.

Valente, P. (2020). *O 'x' e o '@' não são a solução: Sistema elu | Linguagem neutra em género*. Medium. www.medium.com/@pedrosttv/sistema-elu-linguagem-neutra-em-g%C3%A9nero-pt-pt-9529ed3885cf

Zhu, Z. [@zzdesign]. (2021, June 29). *Say hello to the FIRST NON-BINARY They/Them Pronoun in Mandarin. ✦ I created this for the Nonbinary Chinese speaking people* [Photo]. Instagram. www.instagram.com/p/CQt6kFDr05d/