

Gender as a property of words and as a property of structures

Paolo Acquaviva

University College Dublin
paolo.acquaviva@ucd.ie



Received: March 24, 2020

Accepted: June 26, 2020

Abstract

A reflection on the essence of gender as a grammatical concept leads to examining some phenomena which suggest that gender is fundamentally a property of nominal syntactic structures, rather than a word-property shared between morphemes in a structure. It is true that what drives the choice of gender value are mainly ‘controller’ lexical items; but the facts discussed suggest that this should not justify a theoretical understanding of gender as primarily a property of nouns, which happens to be shared in syntactic agreement. There are empirical grounds, it is claimed, for interpreting gender as a morphological expression of agreement relations. The crucial assumption that lexical items are also grammatically represented as formal structures makes it possible to encompass under this analysis cases where gender is and is not associated to the mental representation of a lexical word.

Keywords: gender; interpretable features; nouns; nominality

Resum. *El gènere com una propietat dels mots i com una propietat de les estructures*

Una reflexió sobre l'essència del gènere com a concepte gramatical ens porta a analitzar alguns fenòmens que suggereixen que el gènere és fonamentalment una propietat de les estructures sintàctiques nominals més que no pas una propietat dels mots compartida entre diferents morfemes d'una estructura. És cert que el que determina la tria del gènere són principalment els elements lèxics ‘controladors’, però els fenòmens que estudiem suggereixen que això no hauria de justificar la comprensió teòrica del gènere com una propietat primària dels noms, que es comparteix mitjançant la concordança sintàctica. Argumentem que hi ha evidència empírica per interpretar el gènere com l'expressió morfològica de les relacions de concordança. La premissa bàsica que els elements lèxics també es representen gramaticalment com a estructures formals permet que la nostra anàlisi doni compte dels casos en què el gènere està i no està associat amb la representació mental d'un mot lèxic.

Paraules clau: gènere; trets interpretables; noms; nominalitat

Table of Contents

1. Introduction 2. Does gender have meaning? 3. The morphosyntax of gender in Italian	4. The broader picture: ‘High’ gender in DP and the Spanish neuter 5. Conclusion Acknowledgments References
--	---

1. Introduction

If we ask ourselves what gender is, as opposed to what it does or how it is expressed, it is hard to go beyond the simple idea that it is a grammatical feature; a piece of information playing a role in the grammatical system of those languages that have it. We generally understand gender as a feature on a word that may or may not involve a choice between morphologically contrasting exponents, and that determines for each choice (a gender value) a particular pattern of exponence on all elements that agree with that word in that value. Gender characterizes nouns and pronouns inherently, and determines the shape of adjectives, articles, and other modifiers insofar as these agree with their ‘controller’ nouns or pronouns (this classic conception, and the corresponding terminology, is based on Corbett 1991; Corbett & Fraser 2000, and Aronoff 1994). The choice of gender values partitions the set of nouns into classes, and provides pronouns with a paradigmatic contrast which typically expresses oppositions such as male-female or animate-inanimate (as well as others). Agreement complements this categorizing function with the morphologically overt expression of the grammatical relatedness between an agreement controller and its targets (see Senft 2000 and Aikhenvald 2003 for a characterization of gender systems as a type of noun class systems, and opposed to the non-agreeing nature of classifiers).

Many interesting complications immediately arise; for instance, the fact that agreement controllers can be partitioned into more values than the targets, as in Corbett’s (1991: 151-154) classic description of the way Romanian nouns fall into masculine, feminine, and neuter, while their agreeing modifiers only into masculine and feminine (see also Loporcaro 2018 for a recent in-depth discussion). Over and above all these important ramifications, what is not often made explicit is that this notion of gender is a rather abstract one: a gender specification does not directly determine a form or a meaning, but provides a grammatical characterization which drives (or rather contributes to determine) a type of exponence and / or a semantic interpretation. The characterization itself, however, is purely grammar-internal: typically, the values of gender do not correspond to a straightforward formal or semantic interpretation, but contribute to determine it via an abstract marking, which is not interpretable outside of the grammatical system.

It is of course perfectly possible that the exponents realizing agreement in gender value should happen to be identical, but such formal alignments are accidental and in no way follow from the notion of gender itself. Even the noun class

system of Swahili, most often referenced for examples of alliterative concord like (1a), introduces in many cases exponents on some agreement targets that differ from those on the controller or on other targets, like in (1b) (from Contini-Morava 1996: 257):

- (1) a. ki-ti ki-le ki-li-anguka (Swahili)
 7-chair 7-that 7-PST-fall
 b. m-toto yule a-li-anguka
 1-boy 1-that 1-PST-fall

Likewise, but more subtly, what is shared in gender agreement is not an inflectional class. Consider the following elementary examples:

- (2) a. vetus loc-us (Latin)
 old.M.SG.NOM place-M.SG.NOM
 b. vet-eris loci
 old-M.SG.GEN place.M.SG.GEN
- (3) a. nov-aja zeml-ja (Russian)
 new-F.SG.NOM land.F.SG.NOM
 b. nov-oj zeml-i
 new-F.SG.GEN land-F.SG.GEN
- (4) a. nov-aja peč' (Russian)
 new.F.SG.NOM stove.F.SG.NOM
 b. nov-oj peč-i
 new-F.SG.GEN stove-F.SG.GEN

The feminine noun and the agreeing adjective in (2a) would seem to have the same ending, but in fact they inflect according to different patterns, and the genitive brings this out in (2b) (*vetus* is unsegmented in (2a) because *-us* is in fact not an inflectional ending but part of the stem). The contrasts in (3)-(4) illustrate the complementary point that the form of an agreeing adjective stays the same whether the controller noun belongs to the same (as in (3)) or to a different (as in (4)) inflectional class. Gender thus introduces a morphosyntactic dimension in which the agreeing elements match each other, but what is the same on the level of gender value is not the same both for phonological exponence and also for the more abstract categories identified by what we call inflectional classes. These may well be predictors of gender value, but they are not what is shared by agreement in that value.

That a choice of gender value does not directly translate into a semantic interpretation is even better known, and needs no real argument – even though the next section will look deeper into the issue. While gender oppositions very often reflect

semantic ones, and the choice of a gender value may more or less systematically correlate with a range of interpretations in systems with many values (noun class systems), such regularities are rarely absolute; but above all this sort of semantic motivation is different from an interpretation. Gender oppositions may give grammatical expression to semantically relevant categories, but this does not mean that gender marking ‘means’ those semantic categories and adds them to the meaning of the morpheme it marks. This is because in many cases the meaning component is not separable from the content of the lexeme in question. If the Swahili noun *mafuta* ‘oil’ has the denotation of an unbounded mass domain and its class 6 prefix is typical for denotations with that part structure (mass denotations like *maji* ‘water’, and plural denotations like *ma-gari* or *ma-blanketi* from *gari* ‘car’ and *blanketi* ‘blanket’), it seems odd to claim that class 6 is what makes ‘oil’ denote over masses. Not all cases are so clear, and gender marking does seem to add a meaning component especially when the alternative values correspond to alternative interpretations (see next section). But this limitation, the rather vague nature of semantic common denominators typically associated with a gender value, and the fact that a systematic, consistent match between gender value and a given interpretation is so relatively rare, all suggest that while gender is relevant for meaning, the relation is indirect: what is meaningful is a certain linguistic object corresponding to a grammatical structure, and gender is part of this object. Gender cannot be definitionally anchored into meaning, any more than it can be anchored into form.

Given that gender is such an irreducibly grammar-internal concept, perhaps the correct move is to acknowledge that there is no theoretical question: gender systems just provide a grammar-internal set of oppositions to define paradigmatic classes. Loporcaro’s (2018: 10) informal definition of genders (that is, gender values) as ‘paradigmatic classes of nouns, established on syntagmatic evidence’ is fit for purpose, even though it does not mention pronouns and non-lexical controllers (like nominalized non-finite clauses when they trigger a certain choice of gender value on their determiner or an anaphoric pronoun). Even a revised formulation that covered such items would still amount to an extensional definition, to the effect that gender ‘is’ the sum total of phenomena of this or that kind, and gender values ‘are’ the sets into which the oppositions partition the lexicon.

The following sections take a different direction, and aim for a more precise understanding of gender as part of the grammatical system. This theoretically sharper understanding, it is claimed, reconciles the evidence for gender as a lexical property with various phenomena that escape this definition. Underlying the argument there is an assumption that an ‘understanding’ of a set of linguistic phenomena requires modelling them as part of a more general theory. This should be made explicit, because it is what motivates the search for a characterization of gender(-values) that is not only extensional. Given the question ‘what is gender a property of?’, the claim that it is a property of pieces of nominal morphosyntactic structure related by agreement (in very loose terms) provides an answer that is more precise and provides a better empirical fit than the traditional view.

2. Does gender have meaning?

2.1. Sex-differentiable individuals as a special case

Gender is a morphological marking that is shared across agreeing morphemes. Apart from the question of semantic motivation, where a certain interpretation of a noun can be a predictor of its gender value, does this morphological marking also contribute a meaning by itself?

The question seems too simple to have an unambiguous answer. In one sense, the answer must clearly be ‘no’: it is not the masculine and feminine values of the Russian *otec* ‘father’ and *mat* ‘mother’ that make them denote males and females. Odd (but real) cases like the Italian masculine *mammo* ‘male playing a motherly role’ may well impose some qualifications, but in general, even in the idealized situation where a gender value on a class of lexical nouns *aligns* perfectly with a semantic interpretation, it does not *provide* this meaning component to a noun that would otherwise lack it. There is another side to the question however, brought into focus by doublets where one noun stem differs minimally from another only for its denotation and for its gender choice. Insofar as the morphosyntactic properties, and also the other lexical semantic properties, all remain constant, pairs like those in (5) evidence a direct role of the choice of gender value in driving semantic interpretation:

- | | | |
|--------|---|-----------|
| (5) a. | buco - buca
hole.M - hole.F
‘hole, gap’ - ‘hole, cavity’ ¹ | (Italian) |
| b. | mālus - mālum
apple.F - apple.N
‘apple tree’ - ‘apple’ | (Latin) |

The role of gender in shaping a noun’s lexical semantic content is evidently a vast and nuanced question. With another class of phenomena, however, the role of gender value is much more straightforward.

When a value is called ‘feminine’ because it is the canonical one for nouns denoting females, especially in basic oppositions like those between male and female family members, that value systematically determines a reading as ‘female’ on personal pronouns and on gender-paired nouns denoting humans and sex-differentiated animals, like the Portuguese pairs *filho-filha* ‘son-daughter’ and *operário-operária* ‘male worker-female worker’.² Apparent counterexamples, like the French

1. These glosses attempt to convey a distinction that is clear enough in usage, if elusive-sounding. A black hole is always a *buco nero*, and the same masculine form is the one used for a gap in the finances or a hole in a two-dimensional object like a wall or a blanket. By contrast, the feminine is used for a post-box, for a pocket in a pool table, or for a foxhole dug in the ground. ‘Gap’ and ‘cavity’ point to this two- vs. three-dimensional distinction.
2. I will use the transparent if clumsy label ‘sex-differentiable nouns’ for the cases under discussion, because more technical labels can create confusion. ‘Epicene’ refers to nouns like *person*, rather

elle which routinely resumes grammatically feminine nouns like *statue*, or indeed the English *she* when it refers to ships or nations, are no longer problematic once we realize that the female-denoting interpretation is conditional on a suitable interpretation as ‘human’ and/or ‘sex-differentiable animate’. When this obtains, the unmarked gender value denotes either males or non-differentiated individuals in the denotation, while what we call ‘feminine’ denotes females, and does so deterministically. It is certainly noteworthy that gender value selection should have such a clear-cut and rigid effect on semantics, not just for a grammatical category like pronouns but for a subset of lexical nouns too.

2.2. Delineating the grammatical expression of sex-differentiable individuals

The systematic nature of this interpretation has motivated formal semantic analyses of the contribution of gender as a co-determinant of the interpretation of personal pronouns, along with number and person. The most widely accepted view holds that gender determines a presupposition in the content of pronouns: a masculine pronoun, say, denotes an individual under a certain assignment only if that individual is male, otherwise the semantic value of the pronoun is undefined (much as the truth of *the present king of France* is undefined if there is no king of France). *He left*, then, presupposes that the individual denoted by *he* under an assignment is male, and asserts that this male individual left. In Heim’s (2008) analysis, pronouns are identified by an index and each index corresponds to a variable. Each variable receives a semantic value by an assignment function, which may be determined by the speaker in a context (so-called free pronouns), or by the linguistic structure itself (as in ‘bound’ readings like *she* in *each girl thought that she would get the prize*). Information encoded in person, number, and gender restrict the possible interpretation, and gender does so by adding a presupposition to the effect that the possible denotations must range over the set of males/females, otherwise the value is undefined. Heim (2008: 36) spells out the denotation of the feminine value of gender as in (6):

$$(6) \quad [[\text{feminine}]] = \lambda x_e: x \text{ is female. } x$$

This is an identity function from any individual in the domain (a variable x typed e) to itself, with the crucial proviso that the individual must be a female.

As it is, a denotation like (6) cannot be complete, because it states the content of [feminine] generally, while Heim makes it clear that the analysis only concerns pronouns. In addition, the same obligatory interpretation extends to nouns like *operária* ‘female worker’, so we need a restriction but not one that singles out

than to doublets minimally differing in gender value. ‘Common nouns’ are called by Corbett (1991: 67, 181–182) those which enter distinct agreement patterns while retaining a unique form, again unlike doublets. The most precise label would be ‘motion nouns’ also used by Corbett (1991: 67) (see Thornton 2004 for in-depth analysis and description of the whole phenomenon in Italian); but the term is not immediately transparent.

pronouns alone. Obviously we cannot appeal to the fact that the interpretation in question concerns only sex-differentiable animate individuals, because nouns like the French *personne* are feminine and denote humans, but they are not restricted to females. The descriptive conclusion follows so straightforwardly that it is little more than a restatement: there is something in common between gender-marked pronouns and nouns, when this marking triggers a semantic entailment restricting denotation to a sex-based subset of animate beings. The proposal follows naturally that pronouns like *she* and nouns like *operária* derive their semantic properties from the combination of gender value marking with a grammatically represented [Animate], or indeed [Sex-differentiable] marking. This is not a purely semantic feature, but an element in the grammatical representation. More specifically, I suggest that the interpretation of both pronouns and sex-differentiable nouns involves a morphosyntactically represented element which is absent from the representation of other animate- or human-denoting nouns. In other words, the morphological and semantic alternation that singles out this contrast in nouns and pronouns is a function of their structural grammatical representation.

2.3. *Meaning as meaning of an interpreted grammatical structure*

My approach follows a line of analysis that has independently justified positing a discrete syntactic locus for the encoding of gender, corresponding to the minimal nominal level notated [n] and occurring either as a gender-bearing categorizer of a root (Lowenstamm 2008; Kihm 2008; Kramer 2009, 2015) or pronominally, as part of a larger DP structure or on its own as an ‘empty noun’ (Panagiotidis 2002, 2003; Déchaine & Wiltschko 2012).³ The hypothesis of a grammatically discrete locus for animacy which interacts with gender and results in the grammar-driven interpretation as ‘female’, in particular, comes directly from Panagiotidis (2019) and develops the analysis of Percus (2011). In turn, it presupposes that the distinction between ‘semantically relevant’ and ‘purely morphological’ gender should not be approached in terms of two types of marking which happen to share the same morphological set of oppositions (say, masculine vs. feminine, or Class 1 vs. Class 2), but as the same set of oppositions encoded on distinct loci in the syntax of DP (and in the inner syntax of the noun) and *thereby* having potentially distinct semantic relevance. Familiar possible mismatches like those in (7) between the gender value of the head noun and that of its agreement targets have a structural basis: they arise when the language admits a semantically motivated choice of

3. A more remote antecedent is Alexiadou (2004), who analyzed the difference between nouns with single gender value and nouns like *operária* in syntactic terms: the former are inherently marked for gender, while the latter are marked not for gender but as [+human/+animate], and receive gender morphology ‘via identification with a human referent’ (p. 33). By contrast, such nouns are indeed gender-marked in the present account, and what triggers the pronoun-like deterministic interpretation for them is that [Anim] is so marked. It is not the semantic content of the noun that drives this interpretation, but the grammatical (featural) encoding of this content. This distinguishes nouns like *operário* ‘worker’ (which is [Anim], except in cases like (12)) from *pessoa* ‘person’ (which is not).

gender value on the determiner and on structurally high regions of DP, distinct from that of its head noun:

- (7) a. *khorošiy / khorošaja* *vrač* (Russian)
 good.M.SG.NOM / good.F.SG.NOM doctor
 b. *ngwazi y-athu y-oyamba / ngwazi w-athu w-oyamba*
 hero 9-our 9-first / hero 1-our 1-first (Chichewa; Landau 2016: 983)
 ‘our first hero’

In (7a), *vrač* is regularly inflected as a masculine but it may optionally agree with feminine modifiers when referring to a woman doctor; in (7b), *ngwazi* would normally control agreement in class 9, but it optionally allows for modifiers of class 1 because that is the canonical class for human beings.⁴ As is well known (Steriopo & Wiltschko 2010; Pesetsky 2013; Corbett 1991), the semantically motivated choice of value does not simply recategorize the whole DP, but is sensitive to the syntactic structure: insofar as mixed-gender DPs are admissible, the semantically motivated value must appear outside (hierarchically higher than) the morphologically motivated one:

- (8) a. *ngwazi y-athu w-oyamba* (Chichewa; Corbett 1991: 239)
 hero 9-our 1-first
 b. * *ngwazi w-athu y-oyamba*
 hero 1-our 9-first
 (9) a. ? *interesnaja noviy vrač* (Russian; Pesetsky 2013: 38)
 interesting.F.SG.NOM new.M.SG.NOM doctor
 b. * *interesniy novaja vrač*
 interesting.M.SG.NOM new.F.SG.NOM doctor

These phenomena reveal that gender in these cases is marked higher up in the DP on semantic bases, rather than being inherited from the morphology of the head noun. The explanation extends to cases where all DP constituents agree with the head noun but the DP as a whole triggers semantically motivated agreement on a target in the clause, as in (10a). What refers to a female is the DP as a whole, and so this constituent can optionally trigger feminine agreement even if none of its constituents is so marked; but when this semantic-based choice of value is reflected within the DP, as in (10b), the DP is unambiguously feminine morphologically and it triggers feminine agreement.

4. The linear order of constituents in (7b) reflects the noun-initial nature of Chichewa, which with Landau (2016) I take to indicate that the leftmost ones are most peripheral.

- (10) a. [noviy vrač] prišla / prišyol
 new.M.SG.NOM doctor.M.SG.NOM arrived.F.SG.NOM / arrived.M.SG.NOM
- b. [novaja vrač] prišla / *prišyol
 new.F.SG.NOM doctor.M.SG.NOM arrived.F.SG.NOM
 ‘The new doctor arrived’.

The difference between ‘semantic’ and ‘grammatical’ agreement in gender, then, does not reflect a qualitative difference between ‘concord’ features (purely morphosyntactic) and ‘index’ features (reflecting grammaticalized meaning), but the different structural positions that can be attributed a gender value independently of morphosyntactic agreement requirements (see the in-depth discussion in Landau 2016).

Suppose, then, that if certain feminine pronouns and nouns mandate a reading ‘female’, this is because the gender value marks a grammatical structure enriched with a morpheme that expresses animacy (more precisely, the property of being sex-differentiable, as a linguistic reflex of a conceptualization). The claim is expressed informally in (11):

- (11) a. The presuppositional reading $\lambda x: x \text{ is male / female}$. x is triggered by grammar when the gender system of a language has a value F / M canonically associated with females / males, and when a morpheme [n Anim] is part of a nominal structure which is marked by F / M.
- b. Nouns denoting animate, sex-differentiable individuals do not automatically involve a [n Anim] morpheme as a function of their semantics; a suitable denotation is only a necessary condition for having such a complex morphosyntactic structure.

Note that there is no mention of ‘interpretable’ gender features. Discussions centred on when the feminine value entails ‘female’ often lose sight of the fact that this is not the only reading that a gender value can determine. Contrasts like those in (5) are semantic, beside morphological, but they don’t involve animacy; the same can be said about the widespread mass/kind categorization of Romance ‘neuter’ forms, like the Neapolitan *o kkafé* ‘the.n coffee’ (substance) vs. *o kafé* ‘the.m coffee’ (a serving, a cup) (Loporcaro 2018: 123, and *passim*). Animacy is instead a key component in the *systematic* gender doublets illustrated by *operário-operária*, as well as by pronouns.

In addition, even the classic sex-differentiable nouns do not invariably denote animate or human beings. In the underlined phrase of the following example, *sindaco* denotes a pure role, not an even generically designated set of people with that role:

- (12) In questa città, la funzione di sindaco è svolta da un programma informatico.
 ‘In this town, the function of mayor is fulfilled by a computer program’.

Contexts like (12) foreground the ‘role’ interpretation that many of these doublet nouns have (not all, if pairs like *lupo* - *lupa* ‘wolf - she-wolf’ belong here). A relation to animacy is still part of the lexical content of the predicate even in (12), insofar as the function can be paraphrased as ‘someone (human) performing such-and-such task’; however the noun does not denote animate individuals here, but a role performed by them. If we add that precisely these contexts appear to neutralize the masculine-feminine opposition, two conclusions must be drawn: first, the semantic restriction to sex-differentiable individuals ([n Anim]) is indeed crucial for alternative gender marking, and second, this restriction is not a definitional criterion for the lexical items involved (so, they must appear with [n Anim] in the gender-marked nominal structure, as per (7a) above, but they may go without it in other cases).

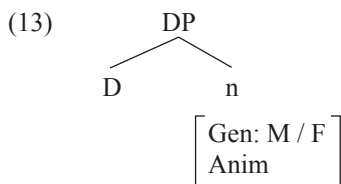
To sum up: it is not the property of being semantically relevant that singles out the function of gender in cases like *operário-operária*, but the systematic connections between a particular interpretation (animate, sex-differentiable) and a paradigmatic morphological choice. As a morphosyntactic category with a morphological reflex, gender expresses formal oppositions. Such formal oppositions often, and not unexpectedly, express semantic oppositions, as in *bucabuca* (cf. (5)) and in the count-mass contrasts alluded to after (11) above. But this is not mandated by the workings of the grammatical system in the present account. Instead, what is so mandated is that a nominal structure that involves both gender marking and [Anim] should determine interpretive oppositions aligned with the grammatical gender oppositions.

The analytic approach sketched out in (11) does not reference gender interpretability, because this is at best an irrelevant concept, and at worst a misguided one. The claim is not that gender is always [–interpretable], but that the [±interpretable] contrast does not apply – at least, not in the sense of distinguishing instances where gender marking is semantically relevant from others where it is not. This follows Hammerly’s (2019: 93-94) conclusion that ‘all features are *interpretable*, in the sense that they are visible to LF, but in any given derivation this does not necessitate that a feature be directly *interpreted*, in the sense that it compositionally contributes to the denotation of a sub-structure’ (emphasis in the original). Gender is always a formal, grammar-internal marker of paradigmatic oppositions, and as such it may or may not subserve semantic oppositions. The deterministic, systematic interpretations displayed by some pronouns and nouns is a distinct phenomenon, triggered by the presence of a gender-marked [n Anim] morpheme in a nominal structure. Nouns like *operário* involve this element quite regularly, but as (12) shows, not as a matter of definition. Gender does not have a meaning, by itself or with some lexical choices (including pronouns). It is a determinant of a grammatical complex object (a structure), and this object has a meaning. It will become increasingly clear that this constructionist approach to gender and noun categorization follows, in particular, Déchaine et al. (2014) and Fassi-Fehri (2018), who show how gender and noun class marking serves to express quite distinct content depending on where it is encoded along the DP spine.

3. The morphosyntax of gender in Italian

3.1. Pronouns and fixed-gender nouns

The singular subject personal pronouns *lui* (masculine, M) and *lei* (feminine, F) denote animate, usually human, individuals that are respectively male and female.⁵ I will assume the following morphosyntactic structure for them:⁶



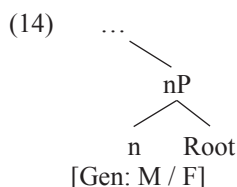
What matters in this case is not so much the location of the gender and animacy features, but their very presence. What is labelled ‘animate’ is not a purely semantic feature, but a grammatical formative which, like a classifier, categorizes nominal constituents and restricts their interpretation to sex-differentiable entities. It is conceived here as an implementation of the idea of empty nouns (Panagiotidis 2003) provided with a grammatical content which contributes to determine the morphosyntactic and semantic content of both pronouns and nouns. In particular, the gender-marked [n] in (13) directly adapts the ‘feminizer’ null morpheme of Percus (2011) and Pesetsky (2013), but differs because it encapsulates in grammatical form the information about sex-differentiable animate individual, rather than relying on the semantic entailments of roots or ‘lexical items’

The reason for doing so is that gender value assignment simply does not correlate with male or female denotation in nouns, and only does so with doublets for sex-differentiable individuals. Percus (2011) engaged particularly closely with this problem, and opted for a complex analysis that required assuming that some roots (but not those of doublets like *operaio* / *operaia*) entail sexedness, that gender features can be deleted subject to the condition that “we are not allowed to leave a feature uninterpreted when using an interpreted feature in its place would yield the same result”, and that a morpheme may be added at LF to make a non-female-entailing root compatible with a female-entailing environment when the F feature cannot be deleted. Even this complex network of assumptions incorrectly assigns a female reading to a feminine-only noun like *checca*, an insulting term for a male homosexual. After correctly noting this exception, Percus (p. 192) considers

5. This ignores the honorific use of *lei* as formal address form, which drives agreement in the third person singular feminine.
6. There is no real attempt in what follows to capture the internal complexity of gender and animacy, in relation to other pronominal features; Carvalho (2018) offers a detailed and reasoned proposal, motivated by an analysis of Brazilian Portuguese. In his account, however, [anim] and [hum] correspond to semantic entailments (so a DP that is not [anim] ‘conditions an inanimate reading in a given language’ (p. 153)).

that the problem is only apparent because the ‘female’ denotation must be understood loosely enough to include individuals with what is perceived as woman-like behaviour. But the fact that the referents of *checca* must be male is not an irrelevant accident: on the contrary, the lexical semantic content of this term consists precisely in the attribution of female nature to a referent that is presupposed to be a male. This is hardly an isolated oddity, because to express the same notion, the feminine gender contrasts with the assumed male denotation also in the French *pédale*, *tafirole*, *tapette*, *tarlouze*, the German *Schwuchtel*, the Spanish *loca*, and the Irish *píteog* (for the last one, derived with a feminine suffix, cf. Acquaviva 2018 and references therein).

It seems preferable to avoid postulating that gender value may ever mandate a type of reading *by itself*. Assuming that what drives a reading of ‘male’ or ‘female’ is a gender-marked [Anim] morpheme accords with the various analyses that view sex-differentiable doublets as morphosyntactically more complex,⁷ and makes clear the relation between such nouns and the pronouns that denote animate and human individuals. For these reasons, the representations of *terra* ‘earth’, *buco* ‘hole, gap’, *buca* ‘hole, cavity’, but also *persona* ‘person’ and *guardia* ‘guard’, as well as *checca* (all feminine except for *buco*) are all the same:⁸



Just as a possibly mono-morphemic pronoun is associated with a complex structure in (13), in the same way I am assuming that a noun realizes a complex structure which (unlike pronouns) involves a lexical root and a variable amount of extra grammatical information, modelled as a structure with distinct heads (this follows constructionist approaches in the wake of Borer 2005, in particular Déchaine et al. 2014 and Acquaviva 2019).

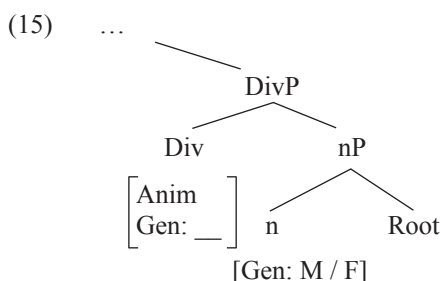
The representation in (14) is only partial, but it suffices to state that gender is marked on the abstract structural element which turns a root into a noun, and that the fact that some of these nouns denote sex-differentiable individuals is not grammatically represented. The specification of one or the other gender value for a given noun may well correspond to a selection of related but distinct semantic contents (and lexical-conceptual addresses), as in *buco* - *buca*; but this possibility simply arises from the paradigmatic opposition between two alternatively marked nouns, and is not due to any one value driving one reading or another.

7. They include Percus (2011), since he too resorts to a morphosyntactically discrete ‘feminizer’ morpheme.
8. Gender is fused with number in Italian inflection, which means that the gender value expressed on [n] must be copied on the Number head and/or that the two loci are fused in morphological realization.

Correspondingly, even the fact that the denotation of *checca* ranges only over males is not expressed grammatically.

3.2. Alternating-gender nouns

Doublets like *operaio* (M) - *operaia* (F) ‘worker - female worker’, but also *ospite* (M) ‘guest’ - *ospite* (F) ‘female guest’, or *turista* (M) - *turista* (F) ‘tourist - female tourist’ have a more complex structure:



Again, this substructure only indicates the elements that matter for the present analysis: how this abstract structure determines the morphological objects underlying the forms just listed is a distinct question, which is not addressed here.⁹ If [n] is the locus for the expression of nominality, what is noted ‘Div’ (after Borer 2005, on which the whole analysis is based) expresses the granularity of the denotation domain. It is not a single mass/count switch, but it is essential in the linguistic expression of countability properties (see Acquaviva 2019 for a detailed proposal). Given that such a classifier-like element has been independently motivated as part of the abstract representation of nouns, it seems very appropriate to locate here the [Animate] feature that, as we saw, is here construed also as a type of abstract classifier. After all, a sex-based restriction is a subtype of a restriction to humans and culturally salient animals, which is not only a prototypical value for classifiers, but also a prototypical way to identify individuals in a count domain. It should perhaps be made clear that the equation with a classifier should be understood in rather abstract terms. Classifiers in the narrow sense, as numeral, shape, or sort classifiers (see Aikhenvald 2003 and Svenonius 2008), are morphemes with

9. A reviewer points out that the relation between the various heads in the hypothesized syntactic structure and the morphemes making up *operaio* and the other forms is far from obvious. For concreteness I will hypothesize that in the examples the root is lexicalized by *operai-* and *ospit-*, and *tur-*, that *-ist-* realizes [Div], that the end vowels are portmanteau affixes spelling out the information on [n] and on a higher [Number] head (not included in (15)), and that the inflectional endings are specified for class diacritics which can restrict application to certain roots. So, *-o* and *-a* require no specification to spell out masculine and feminine singular, but the *-e* of *ospite* and the *-a* of *turista* are inserted conditionally on the presence of *ospit-* and *-ist-* alongside [masc]. This assumes that a derivational affix like *-ist-* is like a root in determining ‘lexical’ insertion contexts for affixes (the same reviewer aptly references Lowenstamm’s [2014]) analysis of some affixal heads as roots.

specific grammatical properties beyond their semantic role, and these properties define a morphosyntactic profile that is typologically quite distinct from that of number-based languages. The intuition behind [Div] (which goes back to Borer 2005) is that a denotation-dividing function akin to that of classifiers proper is also grammatically expressed in the absence of such elements.

The structural proposal in (15) envisages two distinct loci for gender value assignment to the lexical core of the noun ([Gen] under [n]) and to the [Anim] classifier. The value of the latter is driven by the choice inherent in the former; this makes nouns like *amico*, *figlio*, *operaio*, or *ospite* akin to adjectives in that they are one lexical base with two possible gender-valued realizations. More precisely, the structural hypothesis expressed in (15) tells apart the properties of having the content and syntactic distribution of a noun, which is a function of the whole structure (including [nP] and [DivP]), from the properties of lacking a fixed choice of gender value, which characterizes [n] when this occurs with a choice of lexical roots. Already Alexiadou (2004) hypothesized that such nouns lack a gender specification; (15) refines this idea, insofar as it spells it out structurally. This structural interpretation allows for a principled extension to cases where the nouns in question have different distribution and semantic content: namely, when they appear as adjective-like modifiers of another noun, as in (16a), and when they do not denote individuals but the role associated with it, as in (12), here repeated as (16b):

- (16) a. la classe operaia / la squadra ospite / una voce amica
 ‘the working class’ / ‘the guest team’ / ‘a friendly voice’
 b. In questa città, la funzione di sindaco è svolta da un programma informatico.
 ‘In this town, the function of mayor is fulfilled by a computer program’.

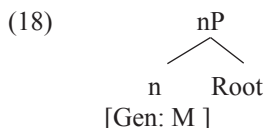
It is certainly significant that precisely terms like *operaio*, but not for instance *persona*, occur in the modifying function shown in (16a). The hypothesis suggests itself that this function differs from the nominal one in lacking the [n] and the [Anim] characterizations.

- (17)
- $$\begin{array}{c}
 \text{AdjP} \\
 \swarrow \quad \searrow \\
 \text{Adj} \quad \text{Root} \\
 [\text{Gen: } __]
 \end{array}$$

Of course this is not an analysis of adjectives, nor could it possibly be. It is simply meant to make explicit how the morphosyntactic structure of these nouns changes when they appear in adjectival use: the claim is that they lack [Div] and [n]. What is notated ‘Adj’ indicates that there is a locus for encoding the gender (and number) feature value determined by an agreement trigger.¹⁰

10. It might be suggested that adjectives too have something like [Div], as some seem to denote properties that necessarily apply to discrete identifiable elements, like *round* (or indeed *individual*). But

As for readings where these terms are clearly nouns but do not actually denote either male and female referents, as in (16b), all we need to assume is that this reading lacks the extra classifier-like [Div]. *Operaio* or *sindaco* in those readings do not differ from any other simplex noun morphosyntactically, and they are open to interpretations like ‘the function of mayor’.¹¹



In sum, nouns like *operaio* have one morphosyntactic representation with [Anim] as members of a gender doublet, and then masculine and feminine values deterministically denote males and females; and they have another representation, without [Anim], when they are not so interpreted. This does not mean that the general and the sex-specific reading should correspond to two distinct ‘lexical items’: it is precisely in cases like this that the notion of ‘lexical item’ shows its limitations as an analytic tool in morphosyntactic analysis. If we hypothesize that the same morphological and phonological representation *operaioi* can spell out either a structure like (15) or one like (18), we may define them as distinct lexical items for that. But that would be a precise choice about what it means to be a lexical item in the context of this theory, not a self-evident conclusion imposed by the data. More importantly, such a choice would not by itself license further theoretical conclusions or empirical predictions. This does not mean that the question about the paradigmatic relationship between (15) and (18) is irrelevant. It does make a difference whether the two grammatical objects correspond to one or two mental addresses, in a sense that can be made empirically precise. But the present account concerns the mental symbolic representation modelled as a grammatical structure, not the psycholinguistic analysis of how this abstract representation is cognitively implemented. In language-internal terms, the claim is that the two structures differ in the presence of [Anim]; this is what determines the morphological and semantic properties we have reviewed.¹²

the fact that the same adjective, with the same denotation, features in the count DP *round pieces of furniture* and in the non-count (syntactically mass) DP *round furniture* shows that modifying adjectives cannot determine the countability properties of a DP.

11. The structure in (18), without [Div], might also be involved in cases like (i), where the masculine value has a default choice and does not restrict the denotation to males:
 - (i) Questa ditta impiega 30 operai di cui 10 donne.
 ‘This firm employs 30 workers including 10 women’.

However this default reading requires plural number. Whether or not [Div] is involved (and the denotation here does necessarily range over individuals, unlike in (18)), the explanation for lack of ‘male only’ entailment cannot be the same as for (18) (namely that a strict reading is no longer mandated in the absence of [Div]). Thanks to a reviewer for noting this.

12. The analysis applies to terms like *amico* - *amica* ‘male - female friend’, but not to *fratello* - *sorella* ‘brother’ - ‘sister’, because the latter consist of distinct stems and do not minimally differ in gender

Note that personal pronouns cannot freely dispense with the representation of [Anim], since they are entirely grammatical objects: either they include this specification, behaving like the Italian *lui* (animate male denotation), or they do not, and then they behave like the French *il* (no such restriction, just agreement in masculine gender value).

3.3. Mixed gender?

A different class of phenomena impinges on the overall question of how gender is morphosyntactically represented, and I will extend the analysis to them in order, first, to show how the proposal can successfully account for different patterns, and second, to re-direct the discussion back to the broader issue of gender as a determinant of grammatical structures rather than of lexical items.

Italian does not generally allow for mixed-gender DPs like those illustrated in (8) and (9) above. However, three types of somewhat eccentric constructions must be addressed to ensure a full empirical coverage to the analysis in 3.2.

First, we should consider mismatches between the gender/number of N and those of D (and related modifiers) which most plausibly involve the ellipsis of an agreeing head noun (19a) or the reanalysis of an inflecting noun as a fixed-form name (19b):¹³

- (19) a. *la ricorsi* (= *la [commissione] ricorsi*)
 the.F.SG appeals.M.PL
 ‘the appeals [committee]’
- b. *la Topolino*
 the.F.SG Topolino.M.SG
 ‘the Topolino [car / car model]’

(19a) is not lexicalized, but it can straightforwardly be interpreted contextually as arising from the ellipsis of the feminine singular *commissione* ‘committee’. *Topolino* in (19b) is formally a masculine singular noun (the diminutive of *topo* ‘mouse’), but it acts as a (real) name of a car model and in that function its inflectional features remain invisible grammatically. It may be opined that the feminine on D formally agrees with a null feminine element, or that there is no such morphosyntactically represented controller head and the name *Topolino* has been assigned a gender on the basis of the feminine *macchina* ‘car’ (see Thornton 2009). What matters is that these are invariable nouns, which can be pluralized (*le ricorsi* ‘the appeals [committees]’, *le Topolino* ‘the Topolino’s’). Whether the

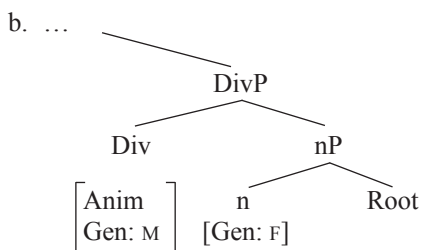
value. For the same reason I am not claiming that English terms like *horse* involve [Anim, M] when denoting males and lack it when they denote both males and females. That may or may not be true, but gender plays no role in those cases and my analysis does not apply to them.

13. On the relation of these cases with invariable nominal exponence in Italian, see also Acquaviva (2009), as well as Thornton (2004) for a full discussion of nouns arising from ‘conversion’ in Italian.

reanalysis of these terms involves or not a null element determining the gender value of the target D, it is this reanalysis that causes the DP to display a mixed gender value.

Cases like *il barba* or *il bomba* are also invariable nouns but they are structurally different, and I suggest that they are best viewed as abstract compounds:

- (20) a. *il barba*, *il bomba*
 the.M beard.F.SG, the.M bomb.F.SG
 ‘the one [male] related to a beard’, ‘the one [male] related to a bomb’



What I notated [DivP] corresponds to a nominal structure that includes information about part structure; [nP] is a less richly determined one, which only specifies the nominal nature and gender. The lexical root that we see in *il barba* has the form of a feminine singular noun, and there is no reason to suppose it does not indeed bear those markings. I am leaving number outside of the picture for the moment, but Italian nominal morphology requires a cumulative exponence for gender and number; we could equivalently assume that *barba* in *il barba* is numberless and appears with the default form, or that the representation in (20b) must be supplemented by a singular value for number for the Root. What matters is that these features do not drive agreement. The reason, it is claimed, is that *il barba* is really something like *the [beard man]*, that is a compound headed by a null element. Unlike the category illustrated in (19), this sort of mixed-gender DPs obligatorily denotes humans, and males when the value on D is masculine. The reason for this is then the same as for *operaio* in (15), except that the [Anim] element does not specify a lexical root, but more indirectly a structure of which that lexical root is a part (and not the projecting syntactic head).

I have called this hypothetical complex an ‘abstract compound’, because the notion of compound is morphological and a two-stem construction where one of the stems is not there is a contradiction. *Il barba* is only a compound in a structural morphosyntactic sense; however, nothing bars this construction from involving, instead of a single lexical root, a compound proper. This is what I think happens in *portalettere* ‘postman’, literally ‘carry-letters’, which interestingly contrasts with the superficially identical *portaerei* ‘aircraft carrier’. Both are exocentric V-N compounds, yet only the human-denoting *portalettere* can be masculine or feminine, denoting mandatorily males or females respectively, while *portaerei* denotes an object and has a fixed gender value. It would seem simpler to assume that if a noun

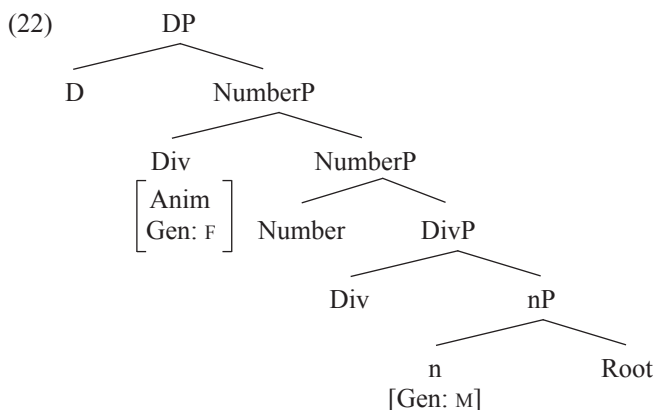
(simple or compound) denotes humans, leaving open the choice of gender value may be a natural way to let the term denote males or females. But this assumes that nouns denoting sex-differentiable individuals (like *portalettere*) just ‘can’ be assigned masculine or feminine. Yet this is not the case: some of them do, but many, like *persona* or *guardia*, cannot. I have hypothesized that nouns in the former class have a different and more complex structure; it is the lack of [Anim, Gen: __] as a grammatical element (as opposed to a semantic entailment) that explains why a male guard cannot be **un guardia*, even though (19) shows that such superficial mismatches are not by themselves problematic.

A final case of mixed-gender DP is quite exceptional in Italian, but it deserves a mention for its theoretical significance. It is illustrated in (21), which also shows its most significant property: it cannot be pluralized, whether in the gender value of D or of the head noun:

- (21) a. la sindaco
 the.F.SG mayor.M.SG
- b. * le sindaco
 the.F.PL mayor.M.SG
- c. * i sindaco
 the.M.SG mayor.M.SG

A feminine D can impose a female reading on the denotation of the head noun even if this remains in its masculine form, as in (21a); but unlike other mixed-gender patterns, this is incompatible with pluralization. In itself, this exceptional behaviour would not justify any general conclusions about the morphosyntax of gender in Italian. But in the context of our discussion, this pattern stands out as a cross between (19)–(20), where gender on D overrides that on the noun, and DPs like *la guardia* which, as just noted, also denote sex-differentiable individuals but do not allow such an override. In this perspective, (21) gives us further reason (beside **il guardia*) to doubt that the semantic reassignment to female or male referents may just involve a switch in the gender value of D. So it would be desirable to have an idea of how this case may be integrated into our account.

The analysis I will propose follows the more detailed account in Acquaviva (2019). Here too, the restriction to females is deterministically driven by grammar, and modelled by means of a classifier-like [Anim] element marked as feminine. However, this case is truly exceptional because the ‘feminizer’ occurs higher up in the DP spine, above Number:



As explained in Acquaviva (2019), which follows the semantic analysis of Number advanced by Harbour (2014), a plurally-valued Number normally selects just the non-atomic elements (the pluralities) in the denotation domain. Here however the classifier-like element which restricts the denotation to females appears above Number, and so its content is not affected by pluralization. A plurally-valued Number would thus coexist with a classifier imposing a restriction that the denotation should be a female, and the structure is uninterpretable because a plurality of females is not itself a female.

This conclusion is not truly explanatory, because I can find no independent reason for positing the crucial gender-marked [Anim] in this unusual structural position. It is however a way of modelling the facts in terms the overall analysis, which indeed proves flexible enough to accommodate this exceptional pattern. The real result is a negative conclusion, which this last discussion reinforces: at least for Italian, the attribution of a male- or female-denoting reading to nouns is not a simple matter of re-categorizing certain nouns as masculine or feminine. The data are instead consistent with an analysis where those readings arise when gender marks a certain nominal structure, specifically, one containing an [Anim] morpheme.

4. The broader picture: ‘High’ gender in DP and the Spanish neuter

The phenomena we have considered inscribe themselves in a broader context. Within and alongside the discussions of ‘semantic’ and ‘morphological’ gender briefly referenced in 2.3, a strand of morphosyntactic research focuses on DPs where D has a semantically motivated gender value different from that of the head noun (beside the Russian data analysed by Steriopolo & Wiltschko 2010 and Pesetsky 2013, Panagiotidis 2019 discusses Greek and Brazilian Portuguese examples). This falls under the rubric of gender determining different readings according to where it is marked in DP, not just in terms of ‘high’ and ‘low’ (Steriopolo & Wiltschko 2010; Steriopolo 2018) but also on intermediate loci along the DP spine. A prominent example is Fassi-Fehri’s (2018) analysis of feminine marking

in Arabic. Beside appearing on nouns interpreted as females, as well as on other nouns whose denotation is not sex-differentiable, feminine marking in this language is also involved in the expression of singulatives (one among many) and of unitized (many-as-one) nouns, and can have a diminutive or augmentative function.

Languages evidently differ in how freely they make use of gender marking to express grammaticalized content in DP which is independent of the choice of the head noun – or, in this framework, of the [nP] which turns a lexical root into a noun. This is a different dimension of variation from those usually considered for gender, namely its morphological regularity, its predictability from form or meaning, and the number of controller and target values. We can reconsider in this light the phenomenon of special gender values which do not appear on controller nouns but correlate with specific interpretations; the best-known example is probably the Spanish ‘neuter’ article *lo* distinct from masculine and feminine (with the same ending as the demonstratives *eso* ‘this’, *esto* ‘that’, *aquello* ‘that’), as in the nominalized property reading in (23a) and the ‘non-lexical controller’ configuration in (23b) (where a predicate is resumed by a neuter pronoun):¹⁴

- (23) a. *lo gracioso de esta propuesta* (Spanish)
 the.N.SG funny.M.SG of this.F.SG proposal
 ‘the funny aspect of this proposal’
- b. *Está enojado pero no lo muestra.*
 is.3.SG annoyed.M.SG but neg it.N.SG show.3.SG
 ‘He is annoyed but he does not show it.’

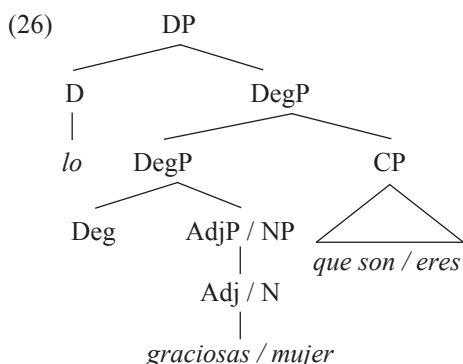
On the face of it, Spanish differs from, say, Italian because it allows for a special form of D in these contexts. However, the difference does not reduce to one language making a morphological opposition that another language neutralizes. While the Italian counterpart of (23) could feature masculine forms instead of the neuter *lo*, any such direct translation leads to unacceptability in cases like (24):

- (24) a. *lo graciosas que son* (Spanish)
 the.N.SG funny.F.PL that be.3.PL
 ‘how funny they [fem.] are’
- b. *lo mujer que eres*
 the.N.SG woman that be.2.SG
 ‘what a woman you are’

14. See Acquaviva (2019) for the view that Spanish and Italo-Romance ‘neuter’ DPs illustrate the general principle that ‘gender, when present, is part of the grammatical construction which globally determines how speakers define their ontology (entity types) and what they talk about (discourse referents)’ (p. 55). The distinctive interpretations of such cases is a function of the whole DP, it is claimed, not of a certain sense of the noun which may condition a special morphology on D.

- (25) a. * *il buffo che sono* (Italian)
 the.N.SG naive.F.PL that be.3.PL
 ‘how naive they [fem.] are’
- b. * *il donna che sei*
 the.N.SG woman that be.2.SG
 ‘what a woman you are’

The reason, I suggest, is that there is more to (24) than meets the eye. The neuter value in (23) arises as a function of the interpretation of the DP alone, as the nominalization of the property expressed by the adjective *gracioso* or as a pronoun resuming a proposition. By contrast, (24a) and (24b) denote degrees to which the adjective property holds: that is, the degree or the extent to which they are funny, and the degree or extent to which the property of being a woman holds of the addressee. It is certainly noteworthy that when a noun appears in this structure, as *mujer* in (24b), it is interpreted as a gradable property, here expressible as ‘how much of a woman you are’. I therefore propose the structure in (26) for both examples in (24), where NP is a shorthand for whatever layer in the nominal structure expresses a property:



Neuter, then, is the gender value assigned to D by Spanish morphology when the DP receives one of a range of ‘abstract’ readings, denoting nominalized properties, possibly facts, or degrees if that is the correct denotation. These may well constitute a semantic natural class, if recent analyses of degrees as nominalized properties denoting kinds are on the right track (Anderson & Morzycki 2015; Scontras 2017). In any case a neuter value for D reflects the interpretation(s) of a complex grammatical object. Italian does not allow a counterpart of (24) because its grammar does not allow for a construction like (26); the closest would be *quanto sono buffe* ‘how they are funny’, with *quanto* lexicalizing a measure/degree operator with Wh-morphology (notice that *quanto*, invariable and uninflected, parallels the Spanish *lo* in being independent of the inflection of the feminine plural adjective *buffe* / *gracioso*).

It goes without saying that much more should be said to support and sharpen this analysis, as well as to clarify to what extent it is applicable to the many similar instances of neuter in Romance varieties (see the very thorough account of Loporcaro 2018; the analysis must certainly be different, at least in part, for varieties where the neuter is directly dependent on the lexical choice of the head noun). However, what matters is that this structural analysis is not isolated. By linking the neuter form of D not just to certain interpretations, but also to some structures that underlie those interpretations, we can appreciate the similarities and the differences with other cases of non-straightforward gender value on D. Like other instances of ‘high’ gender, neuter is inherently marked on D and not copied there by agreement; and it differs from them in that it is outside of the inflectional gender values of the language, and does not categorize sex-differentiable denotations. As the marker of complex DPs with a nominalizing semantics, it is a property of certain grammatical structures with a certain semantic content.

5. Conclusion

That there is more to gender than the list of controller nouns / pronouns and their agreement targets is well known. The phenomenology of gender also includes so-called non-lexical controllers (like predicates or clauses) and instances where a semantics-related value is marked on DP heads, possibly extending to elements in the peripheral layers, as the label ‘high gender’ suggests. In addition, the very notion of ‘lexical’ controller must be nuanced enough to account for cases like exocentric compounds: see *portalettere* and *portaerei* mentioned in section 4, or indeed *fine settimana* ‘weekend’, which is masculine despite both *fine* ‘end’ and *settimana* ‘week’ being feminine.

Simply listing the controllers, the targets, and their combinations is not enough, because it does not explain why the phenomenon of ‘mixed’ or ‘high’ gender is constrained in the way it is. We have only explored a small sample of the relevant empirical landscape, but it has been enough to show that there can be no real understanding of the phenomena without a principled way to rule out cases like **il guardia* ‘the [m] guard [f]’ while ruling in doublets (*operaio* - *operaia*): DPs do not freely acquire a new gender on semantic bases. The account proposed here follows previous analyses in relying on a discrete piece of the morphosyntactic representation, as opposed to purely semantic entailments; it departs from them because it claims that these are not the only cases where a gender value is semantically interpreted, and because it aims at a unified account of semantically relevant gender choice on nouns, pronouns, and on Ds when they do not derive it from a nominal head.

The analyses of personal pronouns and of Italian doublets which forms the core of the proposal pursues the intuition that, while gender oppositions can variously signal semantic contrasts, a family of morphosyntactic structures (where [Anim] is gender-marked) determines an interpretation as a grammatical fact, based on syntagmatic structural properties. There is no need to posit two different abstract natures for what is one morphologically visible gender marking, interpretable and

non-interpretable; and there is no need to posit that gender, being one category, splits into two linguistic entities (again without direct evidence) according to its function as expression of morphosyntactic properties (concord) or of grammaticalized meaning (index). Keying the forms of agreeing elements to a paradigmatic choice defines gender as a unitary, if abstract, category; what gets interpreted, sometimes in a grammatically regimented way, is not the feature itself but the grammatical structures it marks. Gender is a property of structures, because (or rather insofar as) structural properties can predictively derive the constrained distribution of semantically motivated gender assignment.

This contention should not be taken simplistically to imply that gender is a matter of syntax rather than the lexicon, as noted. There are phenomena that are best understood by reference to lexical items as multi-dimensional clusters of listed knowledge, including (say) the gender value of a noun; and there are phenomena that are best understood by reference to the grammatical representations into which this lexical knowledge is encoded. The foregoing discussion has focused specifically on the latter, to argue that an understanding of gender as a formal property of grammatical structures provides a revealing way of making sense of both aspects of the gender phenomenology: it is a property of certain pieces of nominal structure, which sometimes are inherently part of a noun and sometimes are not.

The question ‘what is gender?’ goes back to the initial question ‘what is gender?’. However we characterize it, gender differs from number and person because it does not directly determine or constrain interpretation, any more than it directly determines form. Given its plausible function of highlighting syntactic cohesion between agreeing elements and the tracking of discourse referents, one might wonder why language should make use of such an abstract marking, which does not dialogue with either the conceptual nor the motor system. Assuming a realizational architecture where abstract symbols determine a realization, why cannot this morphosyntactic glue simply share across agreeing elements whatever abstract input determines a uniform realization? Why is gender not always alliterative concord?

The reason for posing such a wide-ranging question at the end of this study is that a plausible answer ties in with the conclusions just stated. On the hypothesis of Late Insertion for lexical roots (Haugen & Siddiqi 2013; Harley 2014), the (morpho-)syntactic representation which feeds morphological realization does not contain information about the form of lexical roots. But if gender necessarily involves morphosyntactic agreement, and so is part of the syntactic representation, it must characterize this purely abstract representation, not the morphologically interpreted one which derives from it. It stands to reason that gender marking, then, should not contain phonological or even purely morphological information like inflectional class: by hypothesis, this is simply not there at the level where features are shared across a morphosyntactic representation. Whether or not this broad claim is sustainable in detail, it sits quite well with a conception of gender as first and foremost a constituent of a morphosyntactic representation; only as an ingredient of ‘grammar’, in this sense, is it also an ingredient of those clusters of linguistic and non-linguistic knowledge we call lexical items.

Acknowledgments

I would like to thank Olga Borik and Boban Arsenijevic for inviting this contribution as well as for their work and support, and two anonymous reviewers for useful and constructive critical observations. The shortcomings are entirely my responsibility.

References

- Acquaviva, P. 2009. The structure of the Italian declension system. In F. Montermini, G. Boyé & Jesse Tseng (eds.), *Selected proceedings of the 6th Décembrettes: Morphology in Bordeaux*, 50-62. Somerville: Cascadilla Proceedings Project.
- Acquaviva, P. 2018. On the exponence of gender in the Irish DP. In J. Merchant, L. Mikkelsen, D. Rudin & K. Sasaki (eds.), *A Reasonable Way to Proceed: Essays in Honor of Jim McCloskey*, 7-30. Oakland: University of California Scholarship.
- Acquaviva, P. 2019. Categorization as noun construction: gender, number, and entity types. In É. Mathieu, M. Dali & G. Zareikar (eds.), *Gender and Noun Classification*, 41-63. Oxford: Oxford University Press.
- Aikhenvald, A. 2003. *Classifiers*. Cambridge: Cambridge University Press.
- Alexiadou, A. 2004. Inflectional class, gender, and DP internal structure. In L. Gunkel, G. Müller & G. Zifonum (eds.), *Explorations in Nominal Inflection*, 21-49. Berlin: Mouton de Gruyter.
- Anderson, C. & Morzycki, M. 2015. Degrees as kinds. *Natural Language and Linguistic Theory* 33: 791-828.
- Aronoff, M. 1994. *Morphology by Itself: Stems and Inflectional Classes*. Cambridge, Mass. MIT Press.
- Borer, H. 2005. *In Name Only*. Oxford: Oxford University Press.
- Carvalho, D. 2018. Gender and agreement in Brazilian Portuguese. In É. Mathieu, M. Dali & G. Zareikar (eds.), *Gender and Noun Classification*, 138-158. Oxford: Oxford University Press.
- Contini-Morava, E. 1996. "Things" in a Noun-Class Language. In E. Andrews & Y. Tobin (eds.), *Toward a calculus of meaning: Studies in Markedness, Distinctive Features and Deixis*, 251-290. Amsterdam: John Benjamins.
- Corbett, G. 1991. *Gender*. Cambridge: Cambridge University Press.
- Corbett, G. & Fraser, N. 2000. Gender assignment: A typology and a model. In G. Senft (ed.), *Systems of Nominal Classification*, 293-325. Cambridge: Cambridge University Press.
- Déchaïne, R.-M. & Wiltschko, M. 2002. Decomposing pronouns. *Linguistic Inquiry* 33: 409-442.
- Déchaïne, R.-M., Mudzingwa, C., Wiltschko, M. & Girard, R. 2014. The internal syntax of Shona Noun Class Prefixes. *Language sciences* 43: 18-46.
- Fassi-Fehri, A. 2018. *Constructing Feminine to Mean*. Lanham: Lexington Books.
- Hammerly, C. 2019. Limiting gender. In É. Mathieu, M. Dali & G. Zareikar (eds.), *Gender and Noun Classification*, 93-118. Oxford: Oxford University Press.
- Harbour, D. 2014. Paucity, abundance, and the theory of number. *Language* 90: 185-229.
- Harley, H. 2014. On the identity of roots. *Theoretical Linguistics* 40: 225-276.

- Haugen, J. & Siddiqi, D. 2013. Roots and the derivation. *Linguistic Inquiry* 44: 493-587.
- Heim, I. 2008. Features on bound pronouns. In D. Harbour, D. Adger & S. Béjar (eds.). *Phi-Theory. Phi-Features Across Modules and Interfaces*, 35-46. Oxford: Oxford University Press.
- Kihm, A. 2008. Noun class, gender, and the lexicon-syntax-morphology interfaces: A comparative study of Niger-Congo and Romance languages. In G. Cinque & R. S. Kayne (eds.). *The Oxford Handbook of Comparative Syntax*, 459-512. Oxford: Oxford University Press.
- Kramer, R. 2009. *Definite markers, phi-features and agreement: A morphosyntactic investigation of the Amharic DP*. PhD dissertation, University of California at Santa Cruz.
- Kramer, R. 2015. *The Morphosyntax of Gender*. Oxford: Oxford University Press.
- Landau, I. 2016. DP-internal semantic agreement: A configurational analysis. *Natural Language and Linguistic Theory* 34: 975-1020.
- Loporcaro, M. 2018. *Gender from Latin to Romance: History, Geography, Typology*. Oxford: Oxford University Press.
- Lowenstamm, J. 2008. On N, $\sqrt{\text{ }}$, and types of nouns. In J.M. Hartmann, V. Hegedüs & H. van Riemsdijk (eds.). *Sounds of Silence: Empty Elements in Syntax and Phonology*, 107-144. Amsterdam: Elsevier.
- Lowenstamm, J. 2014. Derivational affixes as roots. In A. Alexiadou, H. Borer & F. Schäfer (eds.). *The Syntax of Roots and the Roots of Syntax*, 230-259. Oxford: Oxford University Press.
- Nóbrega, V. A. & Panagiotidis, P. 2020. Headedness in exocentric compounding: a new approach to exocentricity. *Word Structure* 13(2): 211-249.
- Panagiotidis, P. 2002. *Pronouns, Clitics and Empty Nouns: "Pronominality" and Licensing in Syntax*. Amsterdam; Philadelphia: John Benjamins.
- Panagiotidis, P. 2003. Empty nouns. *Natural Language and Linguistic Theory* 21: 381-432.
- Panagiotidis, P. 2019. (Grammatical) gender troubles and the gender of pronouns. In É. Mathieu, M. Dali & G. Zareikar (eds.). *Gender and Noun Classification*, 186-199. Oxford: Oxford University Press.
- Percus, O. 2011. Gender features and interpretation: a case study. *Morphology* 21: 167-196.
- Pesetsky, D. 2013. *Russian Case Morphology and the Syntactic Categories*. Cambridge, MA: MIT Press.
- Scontras, G. 2017. A new kind of degree. *Linguistics and Philosophy* 40: 165-205.
- Senft, G. (ed.). 2000. *Systems of Nominal Classification*. Cambridge: Cambridge University Press.
- Steriopolo, O. & Wiltschko, M. 2010. Distributed gender hypothesis. In G. Zybatow, P. Dudchuk, S. Minor & E. Pshehotskaya (eds.). *Formal Studies in Slavic Linguistics. Proceedings of Formal Description of Slavic Languages 7*, 153-172. Frankfurt: Peter Lang.
- Steriopolo, O. 2018. Two syntactic positions of grammatical gender. *Yearbook of the Poznań Linguistic Meeting* 4: 125-151.
- Svenonius, P. 2008. The position of adjectives and other phrasal modifiers in the decomposition of DP. In L. McNally & C. Kennedy (eds.). *Adjectives and Adverbs: Syntax, Semantics, and Discourse*, 16-42. Oxford: Oxford University Press.

- Thornton, A. M. 2004. Conversione in sostantivi. In M. Grossmann & F. Rainer (eds.). *La formazione delle parole in italiano*, 218-227. Tübingen: Niemeyer.
- Thornton, A. M. 2009. Constraining gender assignment rules. *Language Sciences* 31: 14-32.