

Gender markedness: the anatomy of a counter-example

Jonathan David Bobaljik · Cynthia Levart Zocca

Received: 7 August 2008 / Accepted: 10 June 2009
© Springer Science+Business Media B.V. 2010

Abstract The morphological expression of gender on nouns displays a puzzling behaviour under ellipsis of nominal predicates. In some instances, it appears that gender can be ignored in the calculation of the identity/parallelism requirement. With other nouns, gender seems relevant and mismatch engenders parallelism violations. With yet a third group of nouns, there is an asymmetry—an overt masculine noun licenses ellipsis of the corresponding feminine, but not vice versa. The difference between the last two groups is exemplified by the English contrast in: *John is a {waiter/#prince} and Mary is too* (compare *#Mary is a waitress/princess, and John is too*). We examine six languages, and show that nouns for nobility/titles and kinship nouns form a systematic exception to an otherwise stable marked: unmarked opposition, and that when this class of nouns is factored out, the remaining two classes reflect the inflection/derivation distinction in the morphological realization of gender.

Keywords Gender · Ellipsis · Identity/parallelism requirement · Morphological realization · Distributed Morphology · Markedness

1 Introduction: a puzzle

The morphological expression of gender on nouns displays a puzzling behaviour under ellipsis of nominal predicates: in some cases mismatches are allowed, in others there is an asymmetry between masculine and feminine, and in a third group no mismatches are allowed at all. For example, in some instances, it appears that

J. D. Bobaljik (✉) · C. L. Zocca
Department of Linguistics, University of Connecticut, Unit 1145,
337 Mansfield Road, Storrs, CT 06269-1145, USA
e-mail: jonathan.bobaljik@uconn.edu

C. L. Zocca
e-mail: cynth@cynthax.net

gender can be ignored in the calculation of the identity/parallelism requirement: (1) from Brazilian Portuguese (BP) is well-formed, even though the corresponding overt nominal predicate would have a different final vowel than the antecedent noun. Put differently, (1a) does not assert: #*Marta é médico*, but rather: *Marta é médica*. We indicate the intended construal of the elided noun in square brackets after each example.

- (1) a. O Pedro é médic-o e a Marta também é. [médic-a]
 the Pedro is doctor-MASC and the Marta also is doctor-FEM
 b. A Marta é médic-a e o Pedro também é. [médic-o]
 the Marta is doctor-FEM and the Pedro also is doctor-MASC
 ‘Pedro/Marta is a doctor, and Marta/Pedro is too.’

With a different choice of nouns, a sharp asymmetry emerges. An overt masculine antecedent can license ellipsis of the corresponding feminine noun, but the reverse is impossible. The sentence in (2b) asserts (or presupposes) that Paulo is female, and is thus infelicitous.¹

- (2) a. ? O Paulo é ator e a Fernanda também é. [atr-iz]
 the Paulo is actor and the Fernanda also is actr-ess
 b. # A Fernanda é atr-iz e o Paulo também é. [ator]
 the Fernanda is actr-ess and the Paulo also is actor
 ‘Paulo/Fernanda is an actor/actress and Fernanda/Paulo is too.’

With still a different choice of nouns, the asymmetry disappears, but in this case neither form of the noun licenses ellipsis of the opposite gender.

- (3) a. # O Zé vai ser ti-o e a Lu também vai
 the Zé will be uncle-MASC and the Lu also will
 ser. [ti-a]
 be aunt-FEM
 b. # A Lu vai ser ti-a e o Zé também vai ser.
 the Lu will be aunt-FEM and the Zé also will be
 [ti-o]
 uncle-MASC
 ‘Zé/Lu will become an uncle/aunt and Lu/Zé will too.’

The contrast between the patterns in (2) and (3) is evident in other languages as well, including English, as the difference in (4) and (5) shows.²

¹ There is a weak contrast between the examples in (1), but the contrastive judgments between the different classes are quite sharp. Thus, (1b) is acceptable, perhaps slightly odd, while (2b) is strongly infelicitous. Additional remarks on speaker variation are noted below.

² We put aside here the issues surrounding gender-neutral usage for professional designations (such as the trend to avoid forms such as *waitress*, *actress* altogether). We believe that the contrastive judgments we report reflect the intuitions of those speakers who control a register in which, for example, #*Mary is a waiter* is infelicitous. Although there is some variation by speaker and by lexical item, the infelicity of the corresponding sentences outside of ellipsis contexts holds in the other languages considered as well, if anything, more strongly than in English (see below).

- (4) a. John is a waiter, and Mary is ... too. [waitress]
b. # Mary is a waitress, and John is ... too. [waiter]
- (5) a. # Andrew is a prince, and Anne is ... too. [princess]
b. # Anne is a princess, and Andrew is ... too. [prince]

Masculine–feminine contrasts of this sort have been a staple of the literature on semantic markedness since the pioneering work of Jakobson (see below). We suggest, though, that it is only under ellipsis that a three-way contrast emerges, and thus ellipsis provides a new perspective on this venerable topic. If we did not seek to describe all of the behaviours at once, theoretical tools are readily available for understanding any one of the pairs in (1)–(3). For example, the pair in (1) suggests that gender is irrelevant for the concerns of identity/parallelism in ellipsis; this would be consistent with a treatment of gender as an inflectional feature, since inflection is systematically ignored in the resolution of ellipsis identity (see, among others, Lasnik 1995; Nunes and Zocca 2005; Sauerland 2008; and references therein). On the other hand, the pair in (2) suggests that gender is relevant to the parallelism constraint in ellipsis resolution, but in an asymmetric manner: only overt feminine marking needs to be matched in the elided conjunct. This is readily expressible in terms of markedness and underspecification: if feminine is the marked gender on predicates, it must be copied into the elided conjunct, and yields a gender clash with a masculine subject. Masculine, being unmarked, yields no gender clash (in unification terms, it unifies with a subject of either gender). Finally, the pair in (3) suggests, in distinction to (1)–(2), that both genders matter for ellipsis, and that parallelism does not ignore either gender. Each solution accounts for exactly one pair, and yields the wrong predictions for the other pairs. The challenge lies in explaining all three patterns within an internally consistent set of assumptions, and, to the extent possible, predicting the behaviour of a given noun in a given language from independent characteristics.

It is to this puzzle that we turn our attention in this paper. We argue in effect that all three analyses are correct, but apply to different classes of nouns, and we offer a partial basis for the classification of the nouns that will predict their behaviour in ellipsis. We suggest that the difference between (1) and (2) lies in the distinction between inflectional and derivational manifestations of gender; nouns that fall into the *médica* class share gender morphology with adjectives (where gender is clearly inflectional), while nouns for which the gender morphology is not shared with adjectives do not fall into this class (see Sect. 4 for details). This clearly cannot be the whole story, though, in light of nouns such as those in (3) and (5). These nouns have the same morphological characteristics as nouns from the other groups, yet fail to pattern with either of them. We report below on a preliminary survey of six languages, and note that there appears to be a semantic regularity to the pattern of exceptions. Nouns denoting titles, such as ranks of nobility, high status professions, and (some) kinship terms are correlated in all six languages with the behaviour in

(3), regardless of their morphology. We thus suggest a semantic account for these classes of nouns. Specifically, we conclude:

- (i) that the morphology does consist (for these languages) of a simple two-way opposition, with marked feminine opposed to unmarked masculine,³
- (ii) that in the case of derivational affixation, the morphologically unmarked forms are indeed semantically unmarked, but
- (iii) that certain noun stems may nevertheless carry MALE/MASCULINE as a part of their lexical semantics.

Taken together, we thus recognize a possible three-way contrast in the semantics (male vs. female vs. unspecified) even where the morphology and morphosyntax draw only a two-way distinction (feminine vs. unspecified). This, combined with the difference between derivational and inflectional expression of gender, provides an account of the varying patterns illustrated in (1)–(3) and of (much of) their distribution in the languages surveyed here.

We begin the paper with a discussion of the examples that show an apparent markedness asymmetry, as in (2), situating our discussion in the general approach to markedness famously advocated by Jakobson (1932/1984). In Sect. 3 we turn to the nobility and kinship noun pattern in (3). We return to the *médico/a* type alternation in (1) in Sect. 4. This type is, within our survey, only attested in Brazilian Portuguese. In addition to providing an account of this behaviour, we offer some rather tentative suggestions in that section as to why it is absent from the other languages considered.

2 Ellipsis and markedness: the *actress* class

The pairs in (2) and (4) showed an asymmetry regarding ellipsis. A masculine predicate noun may serve as the antecedent for an elided form with a feminine subject, but the reverse is not possible. Further examples from Russian and German are given in (6)–(7). We will refer to noun pairs with this behaviour as the *actress* class.

- (6) a. Ivan moskvič i Marina tože. [moskvička]
 Ivan Muscovite.MASC and Marina too Muscovite-FEM
- b. # Marina moskvič-ka i Ivan tože. [moskvič]
 Marina Muscovite-FEM and Ivan too Muscovite.MASC
 ‘Ivan/Marina is a Muscovite and Marina/Ivan is too.’
- (7) a. Mein Onkel ist (ein) Österreicher, und meine
 my uncle is (an) Austrian.MASC and my
 Tante ... auch. [Österreicher-in]
 aunt ... too Austrian-FEM

³ For ease of exposition we ignore the neuter gender in the languages that have it. Consideration of neuter adds nothing to the points we wish to make for the languages we are discussing; however, in some languages, neuter is used as the unmarked gender, for example, in resolution rules for mixed gender systems (Corbett 1991, p. 298 cites Icelandic as one such language).

- b. # Meine Tante ist (eine) Österreicher-in, und Mein
 My aunt is (an) Austrian-FEM, and my
 Onkel ... auch. [Österreicher]
 uncle ... too Austrian.MASC
 'My uncle/aunt is (an) Austrian and my aunt/uncle is too.'

These examples are typical of the behaviour of nouns in the languages we surveyed that stand in a morphological unmarked (masculine) versus marked (feminine) opposition, i.e., where the feminine is derived from the masculine by the addition of a suffix. Such nouns include names for professions, general descriptive terms for humans (e.g., German: *Held/Held-in* 'hero/hero-ine', *Idiot/Idiot-in* 'idiot', *Pessimist/Pessimist-in* 'pessimist', etc.), demonyms (nouns indicating place of origin or residence) in many languages (though not English), as well as many animal names (see the Appendix). We suggest that this pattern is easily understood as the product of two assumptions: (i) an identity requirement on ellipsis, and (ii) a markedness asymmetry in gender. In order to be able to discuss the contrasting behaviour with other noun classes below, we will spend some time here spelling out some of our assumptions and the analysis explicitly.

As noted, we assume that ellipsis requires identity between the elided material (noun) and the antecedent. It is immaterial to the present account whether ellipsis is copying or deletion, so long as identity is met. (We will refine this below, but not in ways that affect the points made in this discussion). The (b.) examples in (2), (4), (6), (7) are thus all excluded as gender clashes. Identity forces the elided predicate nouns to be construed as feminine, which clashes with their subjects.

Why then are the corresponding (a.) examples acceptable? To understand this requires a particular view of underspecification of gender, and we turn now to a brief aside to review the classic presentation of this perspective.

2.1 Jakobson's *donkey* sentences (a review)

In a now-famous discussion, Roman Jakobson (1932/1984, pp. 2–3) observes that a morphological markedness asymmetry in masculine-feminine pairs is paralleled by a semantic asymmetry. Specifically, where the feminine form is morphologically marked (with respect to the masculine), the use of the feminine form indicates female sex, but the use of the masculine form is, at least in certain instances, neutral as to sex. Thus, in regarding an animal of unknown sex, a speaker of Russian may ask (8a) with the masculine form, and receive an affirmative answer (as in (8b)), along with the further specification that the animal is in fact female, without contradiction.

- (8) a. Èto osel?
 it donkey.MASC
 'Is that a donkey?'
 b. Da, no voobščè-to èto osl-ica.
 Yes but in.general-PRT it donkey-FEM
 'Yes, actually it is a jenny (female donkey).'

On the other hand, if the question is posed with the feminine form (9a), then an affirmative answer with the opposite gender is contradictory (as in (9b)); only a negative answer is felicitous if the animal is in fact male (as in (9c)).

- (9) a. Èto osl-ica ?
 it donkey-FEM
 ‘Is that a (female) donkey?’
 b. # Da, no voobšče-to èto osel.
 Yes but in.general-PRT it donkey.MASC
 ‘Yes, actually it is a (male) donkey.’
 c. Net, (èto) osel.
 No it donkey.MASC
 ‘No, it is a (male) donkey.’

Examples are readily replicated for various other (Indo-European) languages, as Jakobson noted, including English, as shown in (10), and German in (11).⁴

- (10) a. Is that a lion? Yes, (more precisely) it’s a lioness.
 b. Is that a lioness? # Yes, (more precisely) it’s a lion. / No, it’s a lion.
- (11) a. Ist das ein Löwe? Ja, das ist ein-e Löw-in.
 is that a-MASC lion.MASC Yes,that is a-FEM lion-FEM
 b. Ist das ein-e Löw-in? # Ja, das ist ein Löwe. / Nein, das
 is that a-FEM lion-FEM Yes,that is a.MASC lion.MASC No, that
 ist ein Löwe.
 is a.MASC lion.MASC

Jakobson concluded from these examples that the morphologically unmarked form is in fact also unmarked semantically. In other words, while the form with a feminine suffix marks an assertion of female sex, the form with no morphological mark (the masculine, grammatically) makes no assertion about sex. In particular, it does not assert “not female” (i.e., male). Jakobson extends this view to markedness in general:

if Category I announces the existence of A, then Category II does not announce the existence of A, i.e. it does not state whether A is present or not.

The general meaning of the unmarked Category II, as compared to the marked Category I, is restricted to the lack of ‘A-signalization’ (p. 1).

The morpho-semantic markedness asymmetry that Jakobson discussed is revealed by other diagnostics as well. A well-known example (at least in the languages discussed here) is that feminine plural forms can only be used to refer to a group of

⁴ The pattern is clear where the morphology consists of a marked (feminine) versus unmarked (masculine) opposition. The pattern breaks down when the morphology deviates from this form, as in suppletive pairings, especially where there is a specialized vocabulary that speakers are aware of, but possibly unsure about, as in the case of domesticated animals (*horse/stallion/mare*/etc.). We focus for now on the cases with transparent morphological markedness asymmetries, as our main point is to show that there are exceptions to the morpho-semantic parallels even where the morphology is straightforward. We return to this point briefly below.

females, whereas the masculine plural can be used for mixed groups, as well as all-male groups (cf. Greenberg 1966, p. 30; Corbett 1991, pp. 290–299; see especially the latter for discussion of languages which deviate from this pattern).

In our questionnaire pilot study, we established that the nouns that show an asymmetry in ellipsis contexts quite generally also pass the other diagnostics for feminine > masculine markedness in that the singular masculine (but not feminine) may be used to refer to an entity of unknown gender, and that the plural masculine (but not feminine) is used to refer to groups of mixed gender.

2.2 Ellipsis and markedness

We are now in a position to return to the ellipsis asymmetry. For concreteness, we assume that gender features on a noun introduce presuppositions (cf. Cooper 1983; and more recently Heim 2008; Sauerland 2008; Percus this vol., and references therein). Thus, all else being equal, the feminine form introduces a presupposition that the referent is female, while the masculine form crucially introduces no presupposition about gender/sex.⁵

Consider first (4b), analyzed as in (12). The antecedent has the marked, feminine form (12a). There are two choices for resolving the ellipsis at LF (we enclose material that is elided, but interpreted at LF, within angled brackets). Resolving the ellipsis with the masculine (unmarked) form as in (12b) is semantically appropriate, but violates parallelism, as it is not identical to the antecedent. On the other hand, (12b') satisfies parallelism, but is infelicitous, as it carries the presupposition that John is female.

- (12) a. Mary is a waitr-ess and John is ... too.
 PRESUPP: [FEM]
 LF: b. * and John is <a waiter>
 PRESUPP: [Ø]
 LF: b'. # and John is <a waitr-ess>
 PRESUPP: [FEM]

Now consider the reverse example (4a), with the ellipsis resolved as in (13). In this case, it is the semantically most appropriate form, (13b'), that is excluded as a parallelism violation. Our focus, then, is on (13b), the form that respects parallelism.

- (13) LF: John is a waiter and Mary is ... too
 PRESUPP: [Ø]
 LF: b. and Mary is <a waiter>
 PRESUPP: [Ø]
 LF: b'. * and Mary is <a waitr-ess>
 PRESUPP: [FEM]

⁵ One might hesitate about the presuppositional treatment of phi-features on the grounds that examples we treat as presupposition failures have the 'feel' of something stronger, e.g., of contradictions. We do not see that our main points would be affected if gender was taken to mark an assertion rather than a presupposition, and so see this as an issue we may put aside.

Under the markedness hypothesis, *waiter* does not introduce a presupposition of “not female” but rather introduces no presupposition (about gender). Parallelism is respected by positing the unmarked form in the elided conjunct as well, but in contrast to (4b)/(12), there is no gender clash. Mary is female, but the unmarked form introduces no presupposition and hence is compatible in principle with a subject of either sex. The sentence is thus felicitous and the markedness asymmetry accounted for directly.

The account is incomplete, though, since we must now consider the relevant forms outside of ellipsis contexts. Consider the pair in (14).

- (14) a. Mary is a waitr-ess.
b. # Mary is a waiter.

The judgment here is somewhat of an idealization, as some (many?) speakers of English accept forms such as (14b). (Although even speakers we consulted who accept (14b) find a contrast between that and the ellipsis context.) For other languages in our study, we generally found a contrast analogous to (14) for *actress*-class nouns, and typically, the contrast was sharper than in English. Compare Russian (15) to the ellipsis context in (6), and Brazilian Portuguese (16) to (1) and (2).⁶

- (15) a. Moja sestra moskvič-ka.
my sister Muscovite-FEM
'My sister is a Muscovite-FEM.'
b. # Moja sestra moskvič. (Rothstein 1973:463)
my sister Muscovite
'My sister is a Muscovite.'
- (16) a. A Maria é atr-iz / médic-a / american-a
the Maria is actr-ess doctor-FEM American-FEM
'Mary is an actress-FEM/doctor-FEM/American-FEM.'
b. # A Maria é ator / médic-o / american-o
the Maria is actor doctor-MASC / American-MASC
'Mary is an actor/doctor/American.'

Why should this be the case? Indeed, under a parallelism view of ellipsis, the unacceptable (b.) sentences in (14)–(16) are properly contained in the corresponding acceptable examples of mismatched ellipsis. We suggest that a straightforward account is available by assuming some version of a Gricean principle that favours the strongest (most explicit) form compatible with the context. For concreteness, we adopt Heim's (1991) *Maximize Presupposition* (cf. Sauerland 2008), where the

⁶ Rothstein notes significant variation in Russian, among speakers, and among lexical items, in the felicity of predicating a masculine noun of a feminine subject. We direct our attention here to the explanation of forms where structures like (14b) are infelicitous. Our account accommodates the variation, though, in the sense that it also readily characterizes those speakers and nouns where examples like (14b) are acceptable—in those situations, parallelism is satisfied and there is no puzzle to explain.

relevant notion of “strength” is computed over presuppositions: since the feminine introduces a presupposition (female referent) but the masculine does not, the feminine form must be used wherever possible.⁷ In this way, the (b.) examples in (14)–(16) are excluded as violations of *Maximize Presupposition*, and are not treated as true gender mismatch errors.⁸

What makes the ellipsis context in (13) crucially different is the parallelism requirement. In (14)–(16), there are two forms grammatically compatible with the context, and the stronger form (the feminine) is chosen in each case. But in (13), having the feminine form in the LF would violate parallelism, in failing to be identical to the overt antecedent (see (13b')). In the special context of a mismatched ellipsis, the strongest compatible form is therefore the unmarked form (13b). This form thus (trivially) satisfies *Maximize Presupposition*, despite having no gender presupposition. In sum, this ultimately Gricean account, coupled with the underlying markedness asymmetry and the parallelism condition on ellipsis, straightforwardly covers the general pattern seen in English, German, Russian, Brazilian Portuguese, etc. Yet in all of these languages, the pattern expected on the account just given fails to materialize with a different class of nouns. It is to this that we turn next.

3 Noble exceptions: the *princess* class

Although we now appear to understand the workings of the ellipsis diagnostic, the test fails to reveal an expected markedness asymmetry with nouns denoting titles/ranks. The examples in (17)–(20) constitute near-minimal pairs with those considered above. In particular, they show the same morphological structure/alternations as the *actress*-type nouns, yet contrast in their semantic behaviour, as revealed in ellipsis. We will refer to these as *princess*-type nouns. (18)–(20) are German, Russian and B.P., respectively.

- (17) a. # Humperdinck is a **prince** and Buttercup is too. [princ-ess]
 b. # Buttercup is a **princ-ess** and Humperdinck is too. [prince]
- (18) a. # Otto war (ein) **König** und Edith ... auch. [König-in]
 Otto was a king and Edith ... also queen-FEM
 ‘Otto was a king, and Edith was too.’
 b. # Edith war (eine) **König-in** und Otto ... auch. [König]
 Edith was a queen -FEM and Otto ... also king
 ‘Edith was a queen, and Otto was too.’

⁷ Alternatively, this could be a morphological principle, such as Andrews’s (1990) *Morphological Blocking Principle*. In the text, we treat the principle as pragmatic, though nothing of substance would change in the account if it is morphological. See also note 5.

⁸ On our reading, Jakobson tacitly held such a competition-based view, inasmuch as he took the unmarked form to assert [NOT F] just when it is juxtaposed with the marked member of the opposition. Thus, the example “*net—osel*” in (9c) is meaningful, and specifically means that the animal is masculine, in effect by asserting that the unmarked form is the strongest form that may be truthfully used. Given two natural genders, the denial of feminine invites the inference of masculine sex.

- (19) a. # Dolgorukij **knjaz'** i Volkonskaja tože. [knjag-inja]
 Dolgorukij prince and Volkonskaja too princ-ess
 'Dolgorukij was a prince, and Volkonskaja was too.'
- b. # Volkonskaja **knjag-inja** i Dolgorukij tože.[knjaz']
 Volkonskaja princ-ess and Dolgorukij too prince
 'Volkonskaja was a princess, and Dholgorukij was too.'
- (20) a. # Aquele senhor é **barão** e aquela senhora também é. [baron-esa]
 that mister is baron and that mrs. also is baron-ess
 'That man is a baron, and that woman is too.'
- b. # Aquele senhora é **baron-esa** e aquele senhor também é. [barão]
 that mrs. is baron-ess and that mister also is baron
 'That woman is a baroness, and that man is too.'

We tentatively include kinship terms in the *princess* class, as the following examples from Brazilian Portuguese and German illustrate, although we note that there was variation here and some kinship nouns for some speakers patterned instead with the *actress* class.⁹

- (21) a. # O Cláudio é um **ti-o** meu e a Denise também é. [ti-a]
 the Cláudio is an uncle my and the Denise also is aunt
 'Cláudio is an uncle of mine, and Denise is too.'
- b. # A Denise é uma **ti-a** minha e o Cláudio também é. [ti-o]
 the Denise is an aunt my and the Cláudio also is uncle
 'Denise is an aunt of mine, and Cláudio is too.'
- (22) a. # Der Richard ist ein **Schwäger** von mir, und die
 the Richard is a broth.-in-law of me and the
 Christine auch. [Schwäger-in]
 Christine also sister-in-law
 'Richard is a brother-in-law of mine, and Christine is too.'
- b. # Die Christine ist eine **Schwäger-in** von mir, und der
 the Christine is a sist.-in-law of me and the
 Richard auch. [Schwager]
 Richard also bro-in-law
 'Christine is a sister-in-law of mine, and Richard is too.'

⁹ *Enkel-Enkelin* 'grandchild' showed *actress*-class behaviour for one of two German speakers initially consulted, while *Vorfahre-Vorfahrin* 'forbearer' did for both speakers. Kinship terms presented an additional complication in that quite a few show suppletive gender pairs across languages: *father-mother* (**father-ess*); *uncle-aunt* etc. We have found no suppletive pairs that pattern with the *actress* or *médica* classes. This could be because suppletive pairs are predominant in kinship and nobility semantic fields; alternatively, it could be that the nouns in such gender pairings are not grammatically related to each other at all, and thus do not involve (true) suppletion at all (contra Osthoff 1899; see Corbett 2007 and references therein for discussion).

The striking property of nouns of the *princess* class is that both the masculine and feminine forms pattern as marked, under the ellipsis test. That is, they both fail to license ellipsis of the contrasting gender.¹⁰ These same pairs also fail to show a markedness asymmetry under other diagnostics mentioned above. For example, in contrast to *actress* nouns, neither the masculine nor feminine singular can be used for a referent of unknown gender; compare (23a) and (24a).

- (23) a. Is there a waiter in that picture? ?Yes – Mary.
 b. Is there a waitress in that picture? #Yes – John.
- (24) a. Is there a prince in that picture? #Yes – Princess Anne.
 b. Is there a princess in that picture? #Yes – Prince William.

Likewise in plurals there is a contrast between *actress* nouns and *princess* nouns:

- (25) a. *waiters* all males OR group of mixed gender
 b. *princes* all males only NOT group of mixed gender

This behaviour is fairly consistent for the *princess* class of nouns among the languages we surveyed.¹¹

The *princess* class is not a natural class in terms of morphological structure; indeed, the majority of the nouns considered show the same morphological pattern and range of feminine affixes as the *actress*-class nouns. However, the class is fairly well-defined semantically, and the membership of this class (at least for nobility nouns) is consistent across the languages surveyed. We therefore conclude that the account of this class must lie in the lexical semantics of these nouns. Specifically, we conclude that surface morphological parallels such as: *Löwe* : *Low-in* :: *König* : *König-in*, mask an underlying difference. Noun stems such as *Löwe* ‘lion’ are morphologically unmarked and unspecified for gender semantically, but *princess*-class nouns like *König* are semantically specified for masculine

¹⁰ Spathas (2008) reports similar behaviour for kinship nouns in ellipsis in Greek, a language we have not investigated further. For Spanish, the symmetrical behaviour of this noun class is also observed under nominal ellipsis in non-predicate position, as (i) illustrates.

- (i) #Juan visitó a su tí-o y Pedro prometió visitar a la ~~tía~~ de él.
 Juan visited to his uncle and Pedro promised visit to the aunt of he
 #‘Juan visited his uncle, and Pedro promised to visit his [i.e. aunt]’. (Kornfeld and Saab 2004)

¹¹ There was some discrepancy in the judgment for plurals. One Spanish consultant reports it to be acceptable to use *reyes* (‘kings’) to refer to a group containing kings and queens, as in (i):

- (i) Había much-os rey-es en la boda, entre ellos:
 have many-M.PL king-M.PL in the wedding among them:

Juan II de Dinamarca, Felipe VI de España y Cristina III de Suecia
 John II of Denmark, Philip VI of Spain and Christina III of Sweden

‘There were a lot of kings in the wedding. Among them: Juan II of Denmark, Philip VI of Spain and Christina III of Sweden.’

We might hazard to speculate further that the reason nobility nouns have the structure they do lies in the social/historical convention (at least in Western Europe) whereby titles of rank accrue only secondarily to women. As a matter of culture, the relationship of prince:princess, king:queen etc. is inherently asymmetrical in a way in which profession nouns, demonyms, animal terms etc. are not. In addition to the obvious fact of a male-prioritizing line of ascension (a queen is the reigning monarch only when there is no male member available) there is also an asymmetry in the bestowing of titles via marriage—when a king or prince marries, his wife becomes a queen, resp. princess, but when a queen or princess marries, her spouse does not receive the counterpart title of king, resp. prince. It may be this cultural fact that finds itself linguistically manifest, via the meaning of nobility noun roots, in the special behaviour of nobility nouns in ellipsis. This speculation would, if substantiated, require us to qualify the basis of our semantic classification. Specifically, we are led to predict that profession nouns would pattern with nobility nouns in the *princess* class wherever these nouns are used in a title-like fashion (for men), with marked feminine forms extended to wives by courtesy.

In fact, there is some evidence, at least from Brazilian Portuguese, that this might be correct. The noun *embaixador* ‘ambassador’ has two corresponding feminine forms: *embaixadora* and *embaixatriz*. The former refers to a female ambassador, the latter to the wife of an ambassador.¹⁵ Now, consider the behaviour of *embaixador* under ellipsis. The acceptability of the masculine form in the first conjunct, when the subject of the second is feminine, depends crucially on which of the two feminine meanings is to be construed, as shown in (27). The overt material is the same in both examples, but the ellipsis is licit only in case Maria is an ambassador in her own right (the construal of *embaixador* as a profession), and not if Maria is the wife of an ambassador (the construal of *embaixador* as a title).

- (27) a. O João é embaixador e a Maria também é.
 the João is ambassador and the Maria also is
 [embaixador-a]
 ambassador-FEM
- b. # O João é embaixador e a Maria também é.
 the João is ambassador and the Maria also is
 [embaixatr-iz]
 ambassador-FEM

This pattern conforms to our predictions. Under the professional construal, the recovered form in the ellipsis does not specify gender, and is true (she is an ambassador, and female). Under the titular construal, however, the form that satisfies parallelism in ellipsis yields a false assertion, when predicated of *Maria*, as she does not bear the title in her own right. Another example comes from pre-revolutionary Russian, where the feminine suffix *-ša* was added to high-status professional titles for

¹⁵ Note that this is not a matter of the morphemes themselves. The *-or/-riz* alternation does not exclusively refer to the ‘wife of’ counterpart relation, cf. *ator/atriz* ‘actor/actress’ discussed above. Note that the *embaixador/embaixadora* alternation is actually in the *médica* class discussed below, though this is not relevant to the point being made here.

the wife of the profession holder (*doktor/doktor-ša*); we would of course expect these noun pairs to pattern with (or more accurately ‘to have patterned with’) the titular use of *embaixador*.¹⁶ The fluidity of the notion of “high status” may shed light on some speaker variation we have encountered. A reviewer suggests that in the Dutch pair *boer/boerin* ‘farmer (m/f)’, the female *boerin* may mean either a female farmer or the wife of a farmer, a view confirmed by the entry in Van Dale’s *Groot Woordenboek der Nederlandse Taal* which gives: ‘wife of the farmer’ and also ‘woman belonging to the farmer-class’. Another Dutch speaker consulted felt by contrast that *boerin* could not be used (except with humorous intent) to denote the wife of a farmer, if she herself was not involved in the farming. A similar disagreement characterized the intuitions of two German speakers consulted regarding the cognate *Bäuerin* (corresponding to masculine *Bauer*). Tellingly, perhaps, the disagreement about whether the German *Bäuerin* may denote the wife of a farmer appeared to correlate with differences in the understanding of the term *Bauer*. The speaker who permitted the ‘wife of’ reading noted that *Bauer/Bäuerin* requires more than simply working on a farm, but implies also (at least) ownership of the land, and thus arguably relative status; but for the speaker who took *Bäuerin* only to mean a female *Bauer*, the connotations of the noun were broader, encompassing those who work or live on farms, and perhaps rural inhabitants generally.¹⁷ We may thus perhaps understand some of the speaker variation encountered with this term as reflecting variation in the status attributed to its denotation, which plays out in a predictable fashion in the range of meaning available to the feminine form.

In sum, we have offered here a (loose) semantic characterization that we take to underlie the classification of nouns into the *princess* class or the *actress* class. Before moving on to the third class of nouns (the *médica* class), we wish to briefly address an alternative account of similar facts advanced in Haspelmath (2006), embedded in a broad critique of markedness in general. Haspelmath notes that *prince/princess* and *king/queen* do not show the standard Jakobson/Greenberg markedness asymmetry, and in particular, that the masculine member of each pair is restricted to meaning males only. Rather than pursuing a semantic account, as we have done (with a language-external grounding in sociological facts), Haspelmath argues that the key factor in explaining the variable behaviour of unmarked (typically, masculine) nouns with respect to semantic markedness criteria is the relative frequency of use of the marked (feminine) noun as compared to the unmarked. His contention is that, to the extent that the feminine term in a masculine:feminine pairing is rare, the masculine will have a correspondingly wider distribution, encompassing reference to females as well as to males. Thus, he states: “in the pair *dog/bitch*, *bitch* has a much lower proportional frequency than *queen* has in the pair *king/queen*, so it is not surprising that it behaves more like a hyponym of *dog*” (p. 53). Haspelmath does not provide numerical evidence to support his claims, however. So, as a first approximation to quantifying this, we conducted a small

¹⁶ For discussion of this suffix in Russian, and the complications it has engendered in twentieth-century usage of gender designators with profession nouns, see Comrie and Stone (1978, Chap. 6) and Rothstein (1973).

¹⁷ Thus *Bauer/Bäuerin* is translated both as ‘farmer’ and as ‘peasant/peasant-woman’ (among other terms) at <http://dict.leo.org> [consulted May 2009].

sampling of the British National Corpus for eight male:female pairings in English. While there is indeed a sizeable difference between *dog/bitch* and *king/queen* (Haspelmath's primary examples), this difference does not appear to generalize in the manner suggested by Haspelmath. Our results are given in (28).¹⁸

(28) Proportional frequency (male/female) for English gender pairings

Masculine	Occurrences	Feminine	Occurrences	Proportion M/F
lion	2153	lioness	70	30.76
priest	3274	priestess	109	30.04
dog	12406	bitch	1006	12.33
duke	3552	duchess	849	4.18
actor	3540	actress	1051	3.37
baron	795	baroness	239	3.33
waiter	1005	waitress	353	2.85
king	11045	queen	4399	2.51
prince	3774	princess	1743	2.17

Under Haspelmath's proposal, the greater the proportion of masculine to feminine forms in the corpus, the greater the likelihood that the masculine form will behave in the manner we (following Jakobson and others) have termed unmarked. Conversely, the smaller the proportion, the more likely it is that the masculine form will be limited to male reference. Thus *dog* outnumbers *bitch* at $12\frac{1}{3}:1$, and *dog* is not restricted to males, while *king* outnumbers *queen* only $2\frac{1}{2}:1$, and *king* is restricted to males. However, the proposed correlation breaks down quickly in the small sample we examined. The actress-class nouns (*actor*, *waiter*) are interleaved among the princess-class nouns (*duke*, *baron*, *king*, *prince*), and all show roughly comparable proportions of masculine to feminine forms, but strikingly different behaviour under markedness diagnostics. At the opposite end of the spectrum, while the gender-neutral *lion* outnumbers *lioness* at about 30:1 in the corpus, the (arguably) high-status denoting *priest*, which for many speakers permits male reference only and is thus in the princess class, outnumbers the female *priestess* by a similarly wide margin. In sum, for the results we have been able to obtain, over an admittedly

¹⁸ The figures are taken from Adam Kilgarriff's frequency lists [<http://www.kilgarriff.co.uk/bnc-readme.html>, consulted May 2009]. For number of occurrences, the lemmatised list was used whenever both members occurred on that list, otherwise, figures are from the raw lists. Various qualifications detract from the usefulness of this data. For example, the frequency count for *bitch* is limited to occurrences of that word as a noun. However, the list does not distinguish between the sense of 'female dog', and the figurative (derogatory) sense. We included this item because it is Haspelmath's prime exemplar. Note though that correcting for this (by subtracting the figurative uses of *bitch*) will only increase the proportion of occurrences of *dog:bitch* and will thus not affect Haspelmath's point or our response.

small sample, Haspelmath's proposed correlation between proportional frequency and semantic markedness effects is strikingly not supported.¹⁹

4 Inflection versus derivation: the *médica* class

In the preceding section, we argued that nobility nouns (and some kinship terms) constitute a systematic exception to an otherwise regular markedness asymmetry in morphology and semantics. We understand this exception in terms of a distinction in the semantics of gender and noun roots, ultimately, we speculate, with a cultural (extralinguistic) basis. We now return to another class of nouns that fails to show the expected markedness asymmetry in ellipsis (or does so only very weakly), but for which no semantic characterization seems forthcoming. This class of nouns is represented by the BP pair *médico/médica*. This class is essentially restricted to BP within our sample, but within that language, it is a significant class.

Nouns of the *médica* class differ from those of both previously discussed classes, in that, for *médica*-class nouns, either form licenses the ellipsis of the other, with at worst a mild oddness to (29b), which stands in contrast to the strongly infelicitous examples considered above, such as (2b).

- (28) a. O Pedro é médic-o e a Marta também é. [médic-a]
 the Pedro is doctor-MASC and the Marta also is doctor-FEM
 b. A Marta é médic-a e o Pedro também é. [médic-o]
 the Marta is doctor-FEM and the Pedro also is doctor-MASC

Nouns of the *médica* type show this particular behaviour in the ellipsis test, but on the other markedness diagnostics, they pattern together with the regular nouns displaying a gender asymmetry (the *actress* class). Thus:

- (29) a. as médicas a group of female doctors only
 b. os médicos a group of male doctors, or a mixed group
- (30) a. Tem um médic-o na figura? Tem, a Maria.
 have a doctor-MASC in-the picture have the Maria
 'Is there a doctor in the picture? Yes, there is Maria.'
 b. Tem uma médica na figura? #Tem, o João.
 have a doctor-FEM in-the picture have the João
 'Is there a doctor-FEM in the picture? #Yes, there is João.'

¹⁹ Note also that Haspelmath's proposal is about the relative frequency of use of the two members of an opposition, and thus he explicitly extends it to so-called markedness reversals where the feminine term is the unrestricted/unmarked form (as in *cow/bull*). Under Haspelmath's account, since *cow* is not restricted to a single gender, as opposed to male-only *bull*, the ratio of *cow:bull* should pattern with *dog:bitch* and *lion:lioness*, as against *king:queen*. Instead, in the same corpus, the proportion of *cow/bull* is 1.81, a figure not only closer to the figure for *king/queen*, but in fact lower than for any item in (28).

For the *médica* class, then, what needs to be explained is not only these nouns' special behaviour regarding the ellipsis diagnostic (why both genders are permitted in mismatched contexts), but also why the unexpected behaviour is only evidenced in the ellipsis test and not with the other markedness diagnostics. A complete account should also shed light on why this class of nouns occurs in BP, but did not turn up in the other languages we investigated.

In contrast to the preceding section, we suggest here a morphological account of the behaviour of these nouns. Note first that nouns of the *médica* type differ from those considered so far, in that the morphological asymmetry is not in the presence versus absence of a (feminine) affix, but rather in an alternation in theme vowel (-*o/-a*). We will argue that this is part of the key to understanding the behaviour of these nouns. It cannot be the whole picture, since kinship terms that show the same -*o/-a* alternation pattern with (17)–(20), as evidenced in (21), repeated below as (31).

- (31) a. # O Cláudio é um **ti-o** meu e a Denise também
 the Cláudio is an uncle my and the Denise also
 é. [ti-a]
 is aunt
 'Claudio is an uncle of mine, and Denise is too.'
- b. # A Denise é uma **ti-a** minha e o Cláudio também
 the Denise is an aunt my and the Cláudio also
 é. [ti-o]
 is uncle
 'Denise is an aunt of mine, and Claudio is too.'

However, we have argued above that (many) kinship nouns pattern with nobility nouns and have a special semantics. Thus, the root *ti-* is not 'sibling of parent, or spouse thereof', but rather 'brother (in-law) of parent'. The feminine form here must mean 'female counterpart to X' and not 'female X'. The behaviour of the kinship nouns is thus understood as in Sect. 3. Indeed, it appears that with the exception of kinship and nobility nouns (a semantically defined class), all nouns in BP that form the feminine with *-a* added to the masculine stem are in this class.²⁰

The -*o/-a* alternation is significant, since it is the same morphology that is used for gender on adjectives in BP. Moreover, unlike the nominal gender discussed in Sect. 2, adjectival gender agreement is systematically ignored in calculating ellipsis identity, as shown in (32).

- (32) a. O Fernando é bonit-o e a Carol também
 the Fernando is beautiful-MASC and the Carol also
 é. [bonit-a]
 is
 'Fernando is beautiful, and Carol is too.'

²⁰ Morphologically, these are of two types (see the Appendix). On the one hand are nouns, like *médica*, which show an -*o/-a* alternation. On the other, are nouns that end in *-or* in the masculine, and in *-or-a* in the feminine (*professor(a)*, *cantor(a)* etc.).

- b. A Carol é bonit-a e o Fernando também
 the Carol is beautiful-FEM and the Fernando also
 é. [bonit-o]
 is
 ‘Carol is beautiful, and Fernando is too.’

Indeed, this appears to be part of a larger generalization, namely, that inflectional morphology is quite systematically ignored in ellipsis resolution (see, among others, Lasnik 1995; Stjepanović 1997; Nunes and Zocca 2005; Sauerland 2008). For BP, we illustrate as well with verbal agreement in person:

- (33) a. Nós sempre comprá-va-mos aqui, mas eles não. [compra-va-m]
 we always buy-PAST-1P.PL. here, but they not buy-PAST-3P.PL.
 ‘We used to shop here, but they didn’t.’
 b. Eles sempre compra-va-m aqui, mas nós não. [comprá-va-mos]
 they always buy-PAST-3P.PL. here, but we not buy-PAST-1P.PL.
 ‘They used to shop here, but we didn’t.’

There are various proposals in the literature to make agreement features invisible for the purposes of the identity condition on ellipsis. In their most general form, there are two families of proposal, either of which will suffice for present concerns. On the one hand, agreement may be treated as a syntactic operation (feature matching or sharing), with features on the agreement target deleted at LF. On the other, the syntax may operate without specifying the values of agreement features, and these features may be filled in by a transmission or copying rule operating only in the morphology. Since the effect is the same—features on agreement targets are not valued at LF—we need not decide between these here.²¹ What is important for present concerns is that examples like (32) have no gender features on the adjective at the point where identity is calculated. Thus, there is no mismatch, and both examples are acceptable.

- (34) a. LF: O Fernando é bonit- e a Carol também é <bonit->
 Morph: bonit-o
 b. LF: A Carol é bonit- e o Fernando também é <bonit->
 Morph: bonit-a

We therefore suggest that the *médico/médica* alternation has a morphosyntax more like that of adjectives than of the noun pairs considered above. In particular, we suggest that root *médic-* is not inherently specified for gender, and that moreover,

²¹ The agreement-plus-deletion view is superficially, at least, consistent with Chomsky’s view of agreement in the Minimalist Program writings. The agreement-as-morphology view is the preferred view among semanticists looking at related issues in bound variable pronouns, see Kratzer (1998, 2009), Schlenker (1999), Heim (2008) among others. See also Bobaljik (2008) for a syntactic proposal from the perspective of agreement-as-morphology. We do depart from feature-sharing versions of agreement, such as Pollard and Sag (1994) inasmuch as it is important for us that the phi-features on the targets of agreement do not introduce presuppositions.

the choice of theme vowel (*o/a*) arises via agreement.²² Since we know independently that agreement features are ignored in calculating identity for the purposes of ellipsis (see (32)–(33)), it follows that either form can license ellipsis of the other. The level of representation at which ellipsis identity is calculated thus looks as in (35).

- (35) a. LF: O Pedro é médic- e a Marta também é <médic->
 Morph: médic-o
 b. LF: A Marta é médic- e o Pedro también é <médic->
 Morph: médic-a

In neither case is there a gender mismatch at the relevant level (LF), and so the identity relation is satisfied. Other than the kinship terms, the membership of a given noun in BP in either the *actress* class or the *médica* class is straightforwardly determined by its morphology. If the noun shows a theme vowel alternation for gender (including *-Ø/-a* as a special case for nouns in *-or*), then the gender marking is inflectional, and the noun is in the *médica* class. If, by contrast, the feminine is derived by means of a derivational suffix, then the noun is in the *actress* class, with attendant behaviour in ellipsis.

At this point, two further questions arise. First, why do *médica*-class nouns differ from *actress*-class nouns only in ellipsis, but not on other markedness diagnostics? And second, why is the *médica* class only evident in BP, from among our sample of languages?

The answer to the first question is straightforward. We have assumed throughout that there is indeed a markedness asymmetry in the morphology, and for BP—indeed, for Romance generally—we assume the following (minimum) rules for the realization of the theme vowels/inflectional endings:

- (36) -a ⇔ FEM
 -o ⇔ <elsewhere>

Treating the *-o* as the unmarked exponent captures the traditional markedness diagnostics, requiring that the *-o* form be used in all non-feminine contexts. This includes masculines, and also, e.g., mixed groups (29), or where gender is uncertain (30), in the standard manner. The distinction between the *actress* and *médica* classes is therefore not observable in such contexts. It is only in ellipsis, where (36) is irrelevant (because the second noun is simply unpronounced/unrealized), that the difference between the two classes is manifest. On tests other than ellipsis, the

²² A reviewer asks about the analysis of such nouns in non-predicative contexts, for example, in argument position. We suspect that whatever needs to be said for uses of adjectives (without overt nouns) in argument position, as in “*The rich also cry*” (Spanish: *Los ricos también lloran*), may carry over here as well. See also Spencer (2002) on substantivized adjectives in Russian. The inflection vs. derivation contrast is of course a notoriously thorny division to make precise. Our use of the term here is intended to suggest, as we do in the main text, that the nouns with adjective-like morphology pattern with adjectives under ellipsis, at least in terms of the lack of presuppositions associated with either gender, in contrast to the other noun classes considered in this paper.

familiar masculine/feminine asymmetry will emerge with this class of nouns as well, but this time due to the morphological asymmetry expressed in (36).

We may now turn as well to the second question, though here we may offer only a partial answer. Nouns of the *médica* type are absent from the other languages in our survey. The discussion above yields a simple answer as to why this class is generally absent from English, German and Russian, but we must leave as a mystery why the distinction does not surface in Spanish and Italian.²³

English lacks a *médica* class as it lacks inflectional gender altogether. In German, although there is inflectional gender, there is no (inflectional) gender agreement on predicates, which are obligatorily bare. Assuming this to be a property of German morphosyntax (and not specifically of adjectives), the *médica* class is excluded from this language.

- (37) a. ein-e schlau-e Student-in / ein schlau-er Student
 a-FEM smart-FEM student-FEM a smart-MASC student
 b. der Student / die Student-in ist schlau(*-e/-er)
 the.MASC student / the.FEM student-FEM is smart(*-FEM/-MASC)

We turn now to Russian, in which, as in BP, predicate adjectives do show gender agreement with the subject. In this language, we might expect a class of nouns like the *médico/médica* class. However, Russian lacks nouns which express a gender alternation merely by changing declension class (in Russian, this would come out as $-\emptyset/a$). Instead, for true nominals, the feminine is apparently always derived by means of a derivational affix:

- (38) student / student-ka * student-a

While we have no synchronic explanation of this gap, we may reduce the absence of any nouns showing *médica*-class behaviour in ellipsis to this prior lexical gap. In fact, we may refine this statement somewhat. Russian does have a class of substantivized adjectives, which show adjectival inflectional morphology, but are syntactically nouns in their distribution. Spencer (2002) discusses these nominals in detail, and argues that for (only) this class of nominal elements in Russian, gender is inflectional. Among this class are some nouns referring to professions, which do show an inflectional masculine/feminine alternation. We therefore predict that exactly these nouns will show the *médica*-type behaviour in ellipsis. Preliminary inquiries suggest that this is correct.²⁴

²³ Spanish and Italian, like BP, display the same $-o/a$ alternation on some nouns as on adjectives. Gender is ignored in these languages in computing parallelism (as in (32)), yet so far as we can tell, nouns with the $-o/a$ alternation pattern with the *actress* class and not the *médica* class. This is a problem for our account that we leave open.

²⁴ As a substantivized noun, this was the designation for the monitor on each floor in Soviet-era hotels, but the judgment here may be influenced by a parallel use as a true adjective (meaning 'on duty'). The form *zavedujuščij* 'manager' should work this way as well, though the contrast as reported to us by one speaker was less sharp. We make the same prediction for the few profession nouns in German that show adjectival gender inflection, such as *Beamter-Beamtin* 'clerk'.

- (39) a. Vanja dežurn-yj, i Maša ... tože.
 Vanja on.duty-MASC, and Masha ... also
 ‘Vanja is the duty person and Masha is too.’
- b. Maša dežurn-aja, i Vanja ... tože.
 Masha on.duty-FEM and Vanja ... also
 ‘Masha is the duty person and Vanja is too.’

Our hypothesis is that the morphological characteristics of nominal gender morphology will predict the behaviour of a given nominal pair under ellipsis. Nouns that show inflectional morphology—i.e., the same morphology as adjectives—should pattern with the *médica* class, while nouns that show derivational gender morphology should fall into the *actress* class, once the semantically-based exceptions are filtered out. This hypothesis seems to work towards explaining why BP has a large *médica* class, but the Germanic and Slavic languages do not, but the hypothesis fails to make the right prediction for Spanish and Italian.

5 Conclusions/future research

We have presented in this paper an array of data where, contrary to Jakobson’s famous examples, morphological and semantic patterns of markedness in gender do not align. However, we have argued that the primary counter-examples to Jakobson’s thesis constitute a relatively homogenous class (nobility, high-status professions, and kinship) across the languages we have surveyed, suggesting a systematicity to the exceptions. We have provided an account consistent with Jakobson’s basic proposal, but embellished with a marginally richer semantics for these classes of nouns, where the masculine forms do indeed carry a presupposition/assertion of male-ness. In addition, we offered a tentative speculation about the extra-linguistic basis for this special property of nobility nouns. We have also noted a class of nouns (*médico/a*) where the gender asymmetry in ellipsis is weaker than for the cases discussed by Jakobson, if not absent altogether, even though other markedness diagnostics do indicate an asymmetry for these nouns. We argued that the account of this class of exceptions lies in the morphology, rather than the semantics. Specifically, we invoked the conservative view of a distinction between gender as derivation and as inflection, arguing that by and large, gender morphology on nouns is inflectional when it uses the same morphology as adjectival gender morphology, and is derivational when it uses dedicated nominal feminine suffixes. This account works for many of the languages considered here, and explains why the *médica* class is absent from English, German and Russian (except for substantivized adjectives in the latter); however, we are unable to offer an account of why Spanish and Italian pattern with English, German etc. and not with Brazilian Portuguese.

Our primary contribution, as we see it, has thus been to offer a refinement of the views of markedness and underspecification in the realm of gender morphology as inherited from Jakobson. As in any project of this sort, one may debate whether the classes of nouns that we have identified represent discrete categories, or points on a cline. The sample we have used is too small to shed light on this debate. In addition,

it remains to be seen how the paradigms discussed here will (or will not) scale up to consideration of languages beyond the genetically limited sample we have begun with, and especially, to those with a more diverse gender system. In the worst case, we suggest that we have shed some light on the inner workings of gender in the relationship of morphology and semantics in Western European languages; it remains open whether, or in what way, this reflects anything more fundamental.

Acknowledgements The work reported here forms a part of a joint project with Uli Sauerland, supported by the Alexander von Humboldt Foundation and the National Science Foundation, grant #BCS-0616339 (PI: Bobaljik). For discussion of the material presented here, we wish to thank: Benjamin Bruening, Andrea Calabrese, Jairo Nunes, Cilene Rodrigues, Uli Sauerland, Susi Wurmbrand, and the audiences at the MUMSA Workshop, the 34th Berkeley Linguistic Society Meeting, and the 2008 Penn Linguistics Colloquium, as well as two anonymous reviewers. For assistance with the data, and discussion of individual languages, we thank: Carlos Buesa García, Natasha Fitzgibbons, Simona Herdan, Harry van der Hulst, Nina Radkevich, Andrés Saab, Uli Sauerland, Oksana Tarasenkova, and Susi Wurmbrand. The usual disclaimers apply.

6 Appendix

Below is a sample of predicative nouns from the data we collected, grouped according to their behaviour in gender-mismatched ellipsis. As a reminder, here is a summary of each class.

(40) Classes of predicative nouns

Class	-MASC antecedent	-FEM antecedent
	-FEM ellipsis	-MASC ellipsis
<i>actress</i> nouns	✓	*
<i>princess</i> nouns	*	*
<i>médica</i> nouns	✓	✓/?

6.1 *Actress* nouns

English

lion/lioness
tiger/tigress
actor/actress
waiter/waitress
comedian/comedienne

Brazilian Portuguese

tigre/tigresa ‘tiger/tigress’
leão/leoa ‘lion/lioness’
ator/atriz ‘actor’
poeta/poetisa ‘poet’
garçom/garçonete ‘waiter’

Spanish

<i>león/leona</i>	'lion/lioness'
<i>tigre/tigresa</i>	'tiger/tigress'
<i>gato/gata</i>	'cat'
<i>oso/osa</i>	'bear'
<i>professor/professora</i>	'teacher'
<i>celador/celadora</i>	'janitor'
<i>poeta/poetisa</i>	'poet'
<i>actor/actriz</i>	'actor/actress'
<i>director/directora</i>	'director'

German

<i>Elefant/Elefantin</i>	'elephant'
<i>Hund/Hündin</i>	'dog'
<i>Ungar/Ungarin</i>	'Hungarian'
<i>Ausländer/Ausländerin</i>	'foreigner'
<i>Österreicher/Österreicherin</i>	'Austrian'
<i>Pilot/Pilotin</i>	'pilot'
<i>Student/Studentin</i>	'student'
<i>Stadtrat/Stadträtin</i>	'city councillor'
<i>Rektor/Rektorin</i>	'headmaster/headmistress'
<i>Leiter/Leiterin</i>	'principal'
<i>Kunde/Kundin</i>	'customer'

Russian

<i>xomjak/xomjačka</i>	'hamster'
<i>medved'/medvedica</i>	'bear'
<i>rabočij/ rabočaja</i>	'worker'
<i>krest'janin/ krest'janka</i>	'peasant'
<i>učitel'/ učitel'nica</i>	'teacher'
<i>pevec/ pevica</i>	'singer'
<i>tancor/tancovščica</i>	'dancer'
<i>dvornik/dvorničixa</i>	'yard keeper'
<i>prodavec/prodavščica</i>	'salesperson'
<i>laborant/laborantka</i>	'lab assistant'
<i>oficiant/oficiantka</i>	'waiter/waitress'
<i>portnoj/portnixa</i>	'tailor'

Romanian

<i>caine/catea</i>	'dog'
<i>urs/ursaica</i>	'bear'
<i>tigru/tigroaica</i>	'tiger/tigress'
<i>student/studenta</i>	'student'
<i>chelner/chelnerita</i>	'waiter/waitress'
<i>doctor/doctorita</i>	'doctor'
<i>secretar/secretara</i>	'secretary'
<i>dansator/dansatoare</i>	'dancer'

6.2 Princess nouns

English

prince/princess
king/queen
count/countess
baron/baroness
uncle/aunt
brother/sister
husband/wife
brother-in-law/sister-in-law

Brazilian Portuguese

<i>príncipe/princesa</i>	‘prince/princess’
<i>rei/rainha</i>	‘king/queen’
<i>conde/condessa</i>	‘count/countess’
<i>barão/baronesa</i>	‘baron/baroness’
<i>tio/tia</i>	‘uncle/aunt’
<i>irmão/irmã</i>	‘brother/sister’
<i>marido/mulher</i>	‘husband/wife’
<i>cunhado/cunhada</i>	‘brother-in-law/sister-in-law’
<i>genro/nora</i>	‘son-in-law/daughter-in-law’

Spanish

<i>príncipe/princesa</i>	‘prince/princess’
<i>rey/reina</i>	‘king/queen’
<i>conde/condesa</i>	‘count/countess’
<i>barón/baronesa</i>	‘baron/baroness’
<i>tío/tía</i>	‘uncle/aunt’
<i>hermano/hermana</i>	‘brother/sister’
<i>verno/nuera</i>	‘son-in-law/daughter-in-law’
<i>cuñado/cuñada</i>	‘brother-in-law/sister-in-law’

German

<i>Prinz/Prinzessin</i>	‘prince/princess’
<i>König/ Königin</i>	‘king/queen’
<i>Graf/Gräfin</i>	‘count/countess’
<i>Baron/Baronin</i>	‘baron/baroness’
<i>Enkel/Enkelin</i>	‘uncle/aunt’
<i>Bruder/Schwester</i>	‘brother/sister’
<i>Schwäger/Schwägerin</i>	‘brother-in-law/sister-in-law’
<i>Cousin/Kusine</i>	‘cousin’

Russian

<i>princ/princessa</i>	‘prince/princess’
<i>dvorjanin/dvorjanka</i>	‘nobleman, noblewoman’
<i>graf/grafinja</i>	‘count/countess’

<i>svjokor/svekrov'</i>	'father-in-law/mother-in-law'
<i>brat/sestra</i>	'brother/sister'
<i>svat/svat'ja</i>	'son-in-law's father/daughter-in-law's father'
<i>kum/kuma</i>	'godfather/godmother of one's child'
<i>deduška/babuška</i>	'grandfather/grandmother'

Romanian

<i>rege/regina</i>	'king/queen'
<i>baron/baronesa</i>	'baron/baroness'
<i>print/printesa</i>	'prince/princess'
<i>unchi/matusa</i>	'uncle/aunt'
<i>nepot/nepoata</i>	'nephew/niece'
<i>bunic/bunica</i>	'granfather/grandmother'
<i>cumnat/cumnata</i>	'brother/sister'
<i>socru/soacra</i>	'father-in-law/mother-in-law'

6.3 Médica nouns

Brazilian Portuguese

<i>médico/médica</i>	'doctor'
<i>engenheiro/engenheira</i>	'engineer'
<i>enfermeiro/enfermeira</i>	'nurse'
<i>secretário/secretária</i>	'secretary'
<i>diretor/diretora</i>	'director'
<i>professor/professora</i>	'teacher'
<i>cantor/cantora</i>	'singer'
<i>produtor/produzora</i>	'producer'

References

- Andrews, A. (1990). Unification and morphological blocking. *Natural Language and Linguistic Theory*, 8, 507–557.
- Bobaljik, J. D. (2008). Where's Phi? Agreement as a post-syntactic operation. In D. Adger, S. Béjar, & D. Harbour (Eds.) *Phi-theory: Phi features across interfaces and modules* (pp. 295–328). Oxford: Oxford University Press.
- Cooper, R. (1983). *Quantification and syntactic theory*. Dordrecht: Reidel.
- Comrie, B., & Stone, G. (1978). *The Russian language since the revolution*. Oxford University Press.
- Corbett, G. (1991). *Gender*. Cambridge: Cambridge University Press.
- Corbett, G. (2007). Canonical typology, suppletion and possible words. *Language*, 83, 8–42.
- Greenberg, J. (Ed.). (1963). *Universals of language*. Cambridge, MA: MIT Press.
- Greenberg, J. (1966). *Language universals, with special reference to feature hierarchies*. The Hague: Mouton & Co.
- Haspelmath, M. (2006). Against markedness (and what to replace it with). *Journal of Linguistics*, 42(1), 25–70.
- Heim, I. (1991). Artikel und Definitheit. In A. von Stechow & D. Wunderlich (Eds.), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung* (pp. 487–535). Berlin: de Gruyter.
- Heim, I. (2008). Features on bound pronouns. In D. Adger, D. Harbour, & S. Bejar (Eds.), *Phi theory: Phi-features across modules and interfaces* (pp. 35–52). Oxford: Oxford University Press.
- Jakobson, R. (1984). Structure of the Russian verb. In L. R. Waugh & M. Halle (Eds.), *Russian and Slavic Grammar—Studies 1931–1981*. Berlin: Mouton de Gruyter. (Translation of "Zur Struktur des

- russischen Verbums” in *Charisteria Gvilelmo Mathesio qvinqvagenario a discipulis et Circuli Lingvistici Pragensis sodalibus oblata*, Prague, 1932.)
- Kornfeld, L., & Saab, A. (2004). Nominal ellipsis and morphological structure in Spanish. In R. Bok-Bennema, B. Hollebrandse, B. Kampers-Manhe, & P. Sleeman (Eds.), *Romance languages and linguistic theory 2002* (p. 183). Amsterdam: John Benjamins.
- Kratzer, A. (1998). More structural analogies between pronouns and tenses. In D. Strolovitch, & A. Lawson (Eds.), *Proceedings of SALT VIII*. Ithaca: Cornell Linguistics Circle.
- Kratzer, A. (2009). Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry*, 40(2), 187–237.
- Lasnik, H. (1995). Verbal morphology: *Syntactic structures* meets the Minimalist Program. In H. Campos, & P. Kempchinsky (Eds.), *Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero* (pp. 251–275). Georgetown: Georgetown University Press.
- Lumsden, J. (1992). Underspecification in grammatical and natural gender. *Linguistic Inquiry*, 23, 469–486.
- Nunes, J. & Zocca, C. L. (2005). Morphological identity in ellipsis. In *Leiden Working Papers in Linguistics 2.2* (pp. 29–42). Leiden: Leiden University.
- Osthoff, H. (1899). *Vom Suppletivwesen der indogermanischen Sprachen*. Heidelberg: Kommissionsverlag von Alfred Wolff.
- Percus, O. (this volume). Gender features and interpretation: A case study.
- Pollard, C. & Sag, I. (1994). *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.
- Rothstein, R. (1973). Sex, gender, and the October Revolution. In S. Anderson, & P. Kiparsky (Eds.), *A Festschrift for Morris Halle* (pp. 460–466). New York: Holt, Rinehart and Winston.
- Sauerland, U. (2008). Implicated presuppositions. In A. Steube (Ed.), *Sentence and context, language, context and cognition*. Berlin: Mouton de Gruyter.
- Schlenker, P. (1999). Propositional attitudes and indexicality (A cross-categorial approach). PhD Dissertation, MIT.
- Spathas, G. (2008). On the interpretation of gender on nouns and pronouns. Ms., UiL-OTS, Utrecht.
- Spencer, A. (2002). Gender as an inflectional category. *Journal of Linguistics*, 38, 279–312.
- Stjepanović, S. (1997). VP ellipsis in a verb-raising language. In A. Jennifer & A. Lawson (Eds.), *Proceedings of ESCOL 1997*. Cornell University, Ithaca: CLC Publications.