The structure of hybrid nouns in Bosnian/Croatian/Serbian:
Evidence from mismatches under NP ellipsis

Andrew Murphy, Zorica Puškar & Matias Guzmán Naranjo *

Universität Leipzig

FDSL12, Berlin, 09.12.16

Based on their agreement patterns, a class of hybrid nouns in Bosnian/Croatian/Serbian (henceforth: BCS) show that they can simultaneously have both grammatical and natural gender features. Using agreement mismatches under NP ellipsis as a diagnostic for structure and feature specification of hybrid nouns, we show experimental evidence that these nouns disallow agreement mismatches under NP ellipsis. We argue that mismatches are the result of introducing natural masculine gender on top of their grammatical feminine, which violates the identity requirement between the antecedent and the ellipsis site (cf. Merchant's (2001) *e-GIVENness*).

1 Hybrid nouns in BCS

- Hybrid nouns have been a challenge for theories of agreement and NP structure (see Corbett 1991; Wechsler & Zlatić 2003; Pesetsky 2013; Kramer 2015; Landau to appear; Smith 2015, 2016), as it seems that they can simultaneously bear two types of gender specification:
  - natural gender (reflecting the gender of the referent)
  - grammatical gender (assigned arbitrarily, e.g. according to noun's inflection class in Slavic)
- We focus on a particular type of hybrid nouns in BCS, which have grammatical feminine gender, but the natural gender can alternate depending on the discourse referent.
- These nouns belong to declension class II in BCS, which includes nouns ending in *-a*,

(1) *Some hybrid nouns in BCS under examination*:

-a		*-ica*		*-(č)ina*	
mušterija	'customer'	*varalica*	'cheater'	*junačina*	'great hero'
budala	'fool'	*propalica*	'loser, failure'	*dobričina*	'very good person'
sudija	'judge'	*pijanica*	'drunkard'	*drugarčina*	'a great friend'
tužibaba	'telltale'	*spavalica*	'sleeper'	*lažovčina*	'a big liar'

- With male referents, these nouns mostly trigger feminine agreement on the adjectives and predicates (grammatical agreement, (2a)-(2b)).

*andrew.murphy@uni-leipzig.de; zorica.puskar@uni-leipzig.de; matias.guzman_naranjo@uni-leipzig.de

- Some speakers also allow masculine agreement (2c)-(2d).

(2) a. Milan nam je nov-a mušterija.
Milan us is new-F customer
'Milan is our new customer.'
b. Nov-a mušterija je kupila jaknu.
new-F customer is bought.F jacket.
'A new (male or female) customer bought a jacket.'
c. %Milan nam je nov-i mušterija.
Milan us is new-M customer
'Milan is our new customer.'
d. %Nov-i mušterija je kupio jaknu.
new-M customer is bought.M jacket.
'A new (male) customer bought a jacket.'

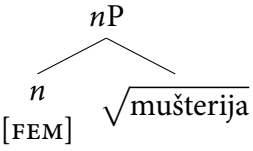
- With female referents, these nouns only trigger feminine agreement.

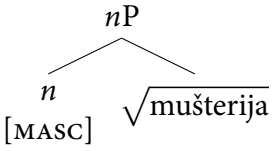
(3) a. Marija nam je nov-a mušterija.
Marija us is new-F customer
'Marija is our new customer.'

b. *Marija nam je nov-i mušterija.
Marija us is new-M customer
'Marija is our new customer.'

- One could treat these as so-called 'epicene' nouns of the type found e.g. in Brazilian Portuguese and Greek, which can be used with both masculine and feminine referents without change in form (Bobaljik & Zocca 2011; Merchant 2014; Kramer 2015).

- It has been proposed that such nouns are listed in the lexicon twice (e.g. Merchant 2014:19):

(4)


(5)


- However, there is evidence that these (and similar hybrid nouns in BCS) nouns can *simultaneously* bear natural and grammatical gender.

1. DP-internal agreement mismatches:

(6) ov-i privatn-e zanatlije
these-**M.PL** private-**F.PL** artisan.PL
'these private artisans' (Corbett 2006:206)

2. Mismatches between concord and verbal agreement:

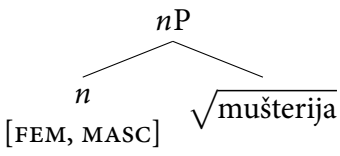
(7) Osm-a budala je bio-∅ mnogo kul tip ali su ga se drugi malo plašili.
eighth-**F.SG** fool is been-**M.SG** very cool guy but are him REFL others little feared
'The eighth fool was a very cool guy, but others were a bit afraid of him.'[1]

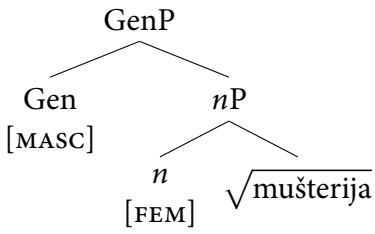
[1] http://magdajanjic.tumblr.com/post/85348961537/budala <accessed 26.11.2016>

3. Mismatches between concord and agreement of the relative pronoun:

(8) Lokaln-a pijanica, koj-i je završio sa igranjem fudbala...
local-**F.SG** drunkard, who-**M.SG** is finished.M.SG with playing football
'A local drunkard, who's finished playing football...'[2]

- Thus, there is an open question about the structural representation of such nouns – do they contain both feminine and masculine gender features?

(9)


(10)


- We argue that gender mismatches under NP ellipsis can shed light on the possible structure of hybrid nouns.

2 Gender mismatches under NP ellipsis

- It is well-known that, while ellipsis imposes an identity requirement on deleted material, some featural mismatches are tolerated under ellipsis, whereas others are not.

(11) *Tense/finiteness/agreement mismatches:*
a. John plans to write a novel, whereas Mary already has ⟨written a novel⟩.
b. Usually, the professor writes the introduction of the paper and his students ⟨write⟩ the rest.

(12) **Voice/auxiliary mismatches:*
a. *Someone murdered Joe, but we don't know who by ⟨Joe was murdered⟩.
b. *Emily was beautiful at the recital, and her sister will ⟨be beautiful⟩ too.

- Recent work has shown that gender mismatches are sometimes possible under NP ellipsis (e.g. Nunes & Zocca 2010; Bobaljik & Zocca 2011; Merchant 2014; Sudo & Spathas 2016; Barrie 2016).
- Broadly speaking, there seem to be three types of nouns regarding their behaviour under NPE:

❶ Two-way alternating nouns (e.g. *doctor*-class):

(13) a. O Pedro é médic-o e a Marta também é ⟨medic-a⟩.
the Pedro is doctor-MASC and the Marta also is doctor-FEM
'Pedro is a doctor and Marta is too.'
b. A Marta é médic-a e o Pedro também é ⟨medic-o⟩.
the Marta is doctor-MASC and the Pedro also is doctor-FEM
'Marta is a doctor and Pedro is too.'
(*Brazilian Portuguese*; Bobaljik & Zocca 2011:142)

[2] http://vukajlija.com/seoski-fudbalski-tim-iz-beton-lige <accessed 26.11.2016>

(14) a. O Petros ine kalos jatros, ala i Maria ine mia kakia ⟨jatros⟩.
the Petros is good.MASC doctor but the Maria is a.FEM bad.FEM doctor
'Petros is a good doctor, but Maria is a bad one.'
b. I Maria ine kali jatros, ala o Petros ine enas kakos ⟨jatros⟩.
the Maria is good.FEM doctor but the Petros is a.MASC bad.MASC doctor
'Maria is a good doctor, but Petros is a bad one.'
(*Greek*; Merchant 2014:15)

❷ Non-alternating nouns (*brother*-class):

(15) a. *O Zé vai ser tio e a Lu também vai ser ⟨tia⟩.
the Zé will be uncle and the Lu also will be aunt
'*Int.* Zé will become an uncle and Lu will become one (an aunt) too.'
b. *A Lu vai ser tia e o Zé também vai ser ⟨tio⟩.
the Lu will be aunt and the Zé also will be uncle
'*Int.* Lu will become an aunt and Lu will become one (an uncle) too.'
(*Brazilian Portuguese*; Bobaljik & Zocca 2011:142)

(16) a. *O Petros ine kalos adherfos, ala i Maria ine mia kakia ⟨adherfi⟩.
the Petros is good.MASC brother but the Maria is a.FEM bad.FEM sister
'*Int.* Petros is a good brother, but Maria is a bad one (sister).'
b. *I Maria ine kali adherfi, ala o Petros ine enes kakos ⟨adherfos⟩.
the Maria is good.FEM sister but the Petros is a.MASC bad.MASC brother
'*Int.* Petros is a good brother, but Maria is a bad one (sister).'
(*Greek*; Merchant 2014:12)

❸ One-way alternating nouns (*actor*-class):

(17) a. ?O Paulo é ator e a Fernanda também é ⟨atriz⟩.
the Paulo is actor and the Fernanda also is actress
'Paulo is an actor and Fernanda is too.'
b. *A Fernanda é atriz e o Paulo também é ⟨ator⟩.
the Fernanda is actress and the Paulo also is actor
'*Int.* Fernanda is an actress and Paulo is (an actor) too.'
(*Brazilian Portuguese*; Bobaljik & Zocca 2011:142)

(18) a. O Petros ine kalos dhasakalos, ala i Maria ine mia kakia
the Petros is good.MASC teacher.MASC but the Maria is a.FEM bad.FEM
⟨dhaskala⟩.
teacher.FEM
'Petros is a good teacher, but Maria is a bad one.'
b. *I Maria ine kali dhasakala, ala o Petros ine enas kakos
the Maria is good.FEM teacher.FEM but the Maria is a.MASC bad.MASC
⟨dhaskalos⟩.
teacher.MASC
'*Int.* Maria is a good teacher, but Petros is a bad one.' (*Greek*; Merchant 2014:16)

- We therefore find the following patterns:

(19) *Classes of predicative nouns* (cf. Bobaljik & Zocca 2011:162):

Class	MASC antecedent FEM ellipsis	FEM antecedent MASC ellipsis	Type
doctor-class	✓	✓	Two-way alternating
brother-class	*	*	Non-alternating
actor-class	✓	*	One-way alternating

- We would expect that hybrid nouns in BCS pattern with one of these types. The complication is that they allow variation between feminine and masculine agreement, as well as simultaneous presence of two gender features.
- Thus, what we want to find out is the extent to which NP ellipsis is possible given mismatched gender on referents and variable agreement on the adjectives:

(20) a. Milan mu je star-a mušterija, a Marija mu je nov-a ⟨mušterija⟩.
Milan him is old-F customer but Marija him is new-F
'Milan is his old customer and Marija a new one.'
b. ?Milan mu je star-i mušterija, a Marija mu je nov-a ⟨mušterija⟩.
Milan him is old-M customer but Marija him is new-F
c. ?Marija mu je star-a mušterija, a Milan mu je nov-i ⟨mušterija⟩.
Marija him is old-F customer but Milan him is new-M

3 The experiment

- **Aim:**
  - Test the grammaticality of sentences in which natural gender licenses the ellipsis of grammatical gender and vice versa.
  - Test whether there is a significant difference between the acceptability of the two combinations, which would in turn indicate the level of markedness of gender features.
- **Task:**
  - Grammaticality judgement; 7-point Likert scale (1=completely bad, 7=sounds excellent).
  - Each sentence contained two clauses (with subjects of different natural genders), in both of which the predicate noun was a *customer*-type hybrid noun, present in the first clause as the antecedent, and elided in the second clause.

3.1 Method

3.1.1 Factors and conditions

- **Factors**
  1. AGREEMENT WITH MASCULINE SUBJECT; two levels: grammatical (MF) and natural (MM)
  2. AGREEMENT WITH FEMININE SUBJECT; one level: FF
  3. CLAUSE; two levels: FIRST and SECOND

- **Conditions**
- The combinations of factors above yield four basic test conditions:

	Factor				Condition
	M.Subj.Agr		Clause	F.Subj.Agr	
1	MF	(Gram.Gen)	First	FF	MFFF
2	MM	(Nat.Gen)	First	FF	MMFF
3	MF	(Gram.Gen)	Second	FF	FFMF
4	MM	(Nat.Gen)	Second	FF	FFMM

Table 1: Test items

3.1.2 Stimuli

Test items

- Sentences containing two clauses coordinated by the conjunction *a* 'but' of the following type:

(21) Jovan je redovn-a mušterija, a Milica povremen-a _____.
Jovan is regular-F customer but Milica occasional-F
'Jovan is a regular customer and Milica an occasional one.' MFFF

(22) Jovan je redovn-i mušterija, a Milica povremen-a _____.
Jovan is regular-M customer but Milica occasional-F
'Jovan is a regular customer and Milica an occasional one.' MMFF

(23) Milica je povremen-a mušterija, a Jovan redovn-a _____.
Milica is occasional-F customer but Jovan regular-F
'Milica is an occasional customer and Jovan a regular one.' FFMF

(24) Milica je povremen-a mušterija, a Jovan redovn-i _____.
Milica is occasional-F customer but Jovan regular-M
'Milica is an occasional customer and Jovan a regular one.' FFMM

- 96 test items; 24 items per condition; varying the proper names (24 male, 24 female), adjectives (48) and hybrid nouns (6).
- *Latin square design*: The 96 test items were distributed in 4 lists such that each list contained different items for every condition.
- Each participant thus saw only items from one list, i.e. 24 test items, 6 items per condition.

Control items and fillers

- Three types of control items:

1. Only masculine subjects with masculine agreement (MMMM) (25); used to establish how much speakers like masculine agreement.

(25) Uroš je redovn-i mušterija, a Tomislav povremen-i _____.
Uroš is regular-M customer but Tomislav occasional-M
'Uroš is a regular customer and Tomislav an occasional one.' MMMM

2. Good baseline; feminine subjects with feminine agreement (26); expected to be fully grammatical:

   (26) Jelena je tešk-a dobričina, a Ljubica umerenij-a _____.
   Jelena is heavy-F good-person but Ljubica moderate-F
   'Jelena is a really good person and Ljubica less of a one.' FFFF

3. Bad baseline; feminine subjects with masculine agreement (27); expected to be completely ungrammatical:

   (27) Stanislava je velik-i škrtica, a Dušanka darežljivij-i _____.
   Stanislava is big-M scrooge but Dušanka more.generous-M
   'Stanislava is a big scrooge and Dušanka a more generous one.' FMFM

- A total of 18 controls (6 per combination), all of them the same in every list.
- Fillers:
  - Sentences containing other types of hybrid nouns of declension class II. Different form of sentences in (28), with different agreement patterns and different degrees of grammaticality:

    (28) Stare vladike su se posvađale na ulici.
    old.F.PL bishops are refl. argued.F.PL on street
    'Old bishops argued on the street.'

  - 20 items per list; 80 items total.

3.1.3 Procedure

- The experiment was coded using LimeSurvey (LimeSurvey Project Team 2012) and run online via the LimeService platform.
- Sentences were presented one by one in a random order.
- Each participant saw 62 sentences sentences (24 test items + 18 controls + 20 fillers).
- The participant was asked to give a grammaticality judgment on a 7-point Likert scale (1=completely bad, 7=sounds excellent) by dragging a slider from the middle of the scale towards the number signaling their response.

(29)

Figure 1: Example experimental item

3.1.4 Participants

- A total of 164 volunteers, 131 female and 33 male, aged 16–66.
- Different varieties of Bosnian (22 speakers), Croatian (5) and Serbian (136).
- None of the participants were paid or otherwise compensated for their participation.

4 Results

- Figure 2 shows the distribution of responses under each conditions from all participants.
- Strong grammaticality effects: FFFF (grammatical), FFMF (grammatical), MFFF (grammatical) and FMFM (ungrammatical).
- More gradient unacceptability for FFMM and MMFF. A u-shaped type of distribution for MMMM suggests that although most speakers disliked it, some speakers found it grammatical.

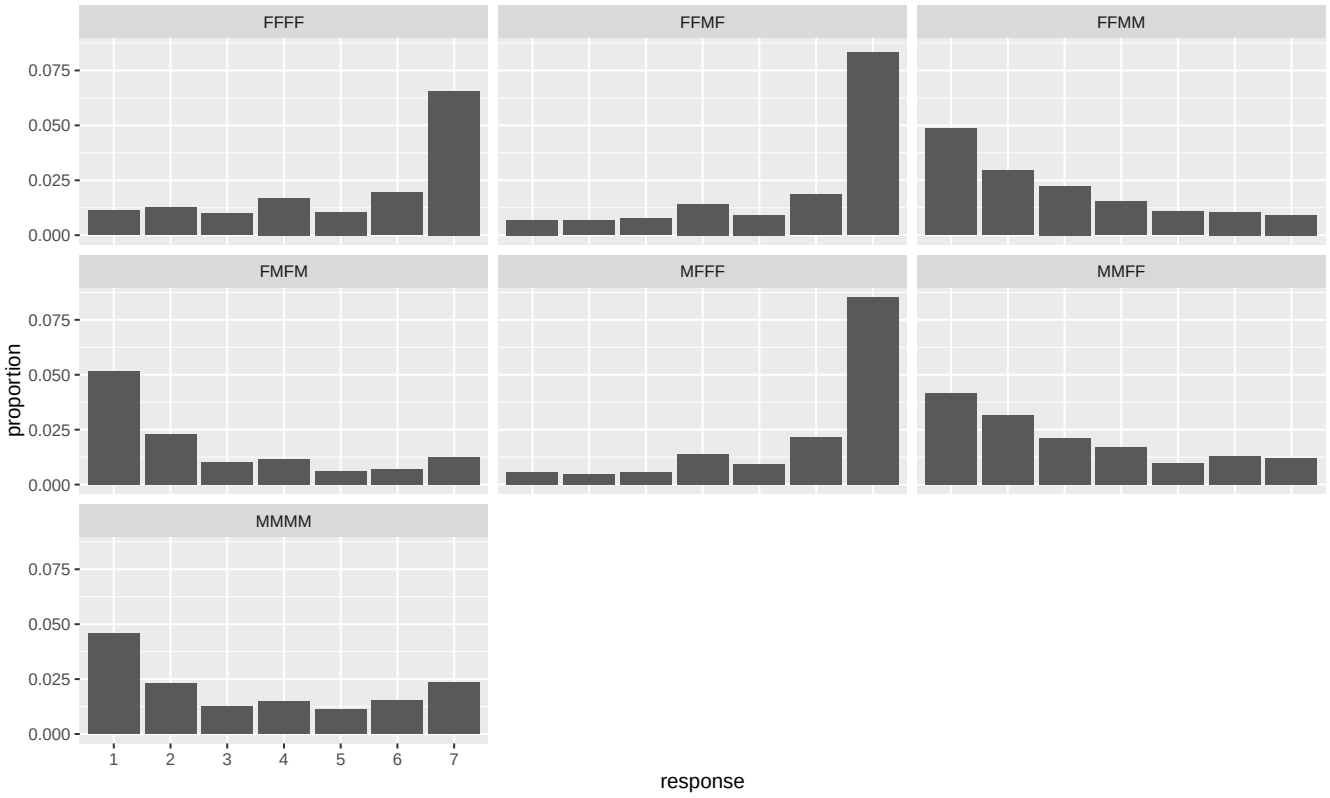


Figure 2: All responses by all participants.

- We compared and contrasted the responses for all conditions based on whether speakers liked or disliked (median rating >= 4) the MMMM combination (Figure 3).
- 51 speakers found the MMMM combination grammatical.
- Distributions for the grammatical patterns are fundamentally the same, whether the speaker liked MMMM or not.

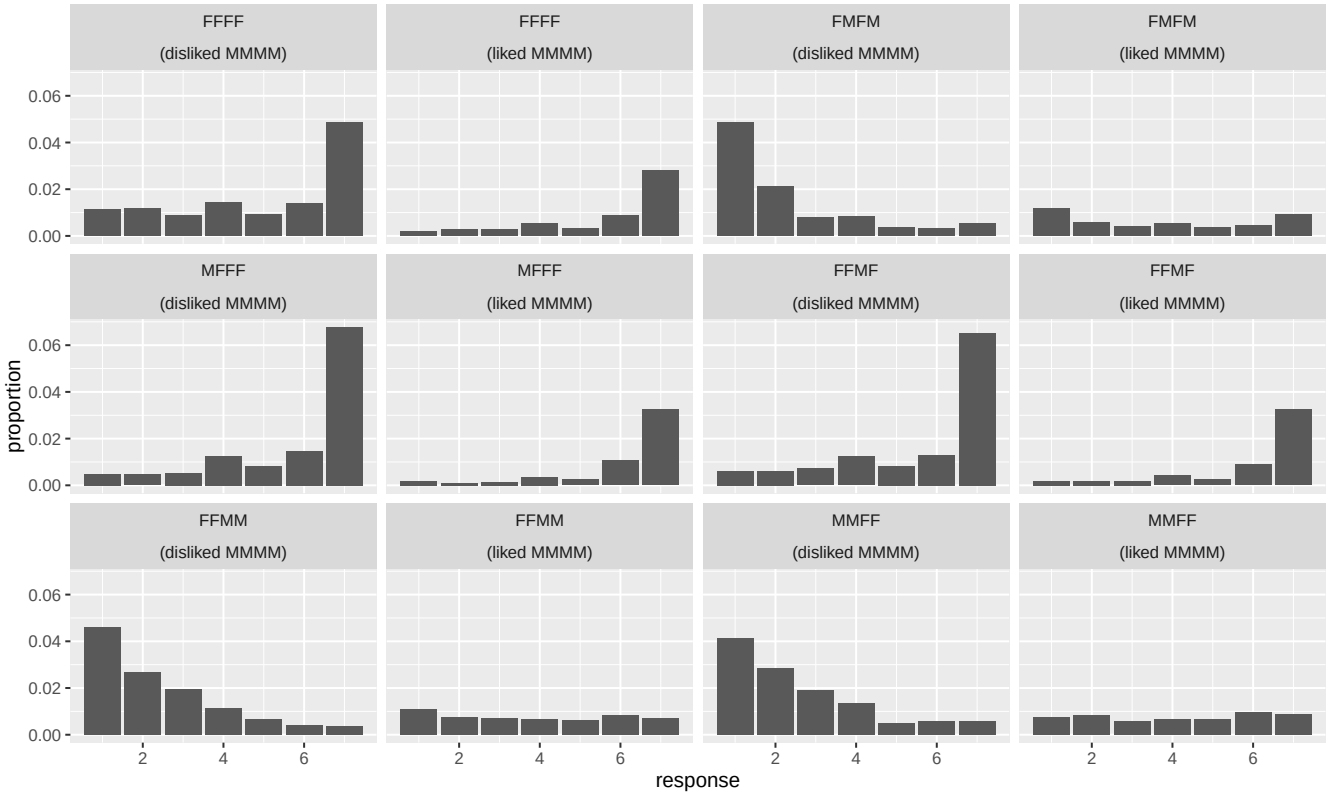


Figure 3: All responses by all participants according to whether the speaker liked or disliked the sentences in the MMMM condition.

- However, in the case of the low acceptability patterns (FFMM and MMFF), speakers who liked MMMM showed no clear preference or dispreference.
- As Figure 4 shows, the patterns with feminine agreement throughout were acceptable and grammatical to these speakers, while the patterns with mismatches were highly dispreferred.

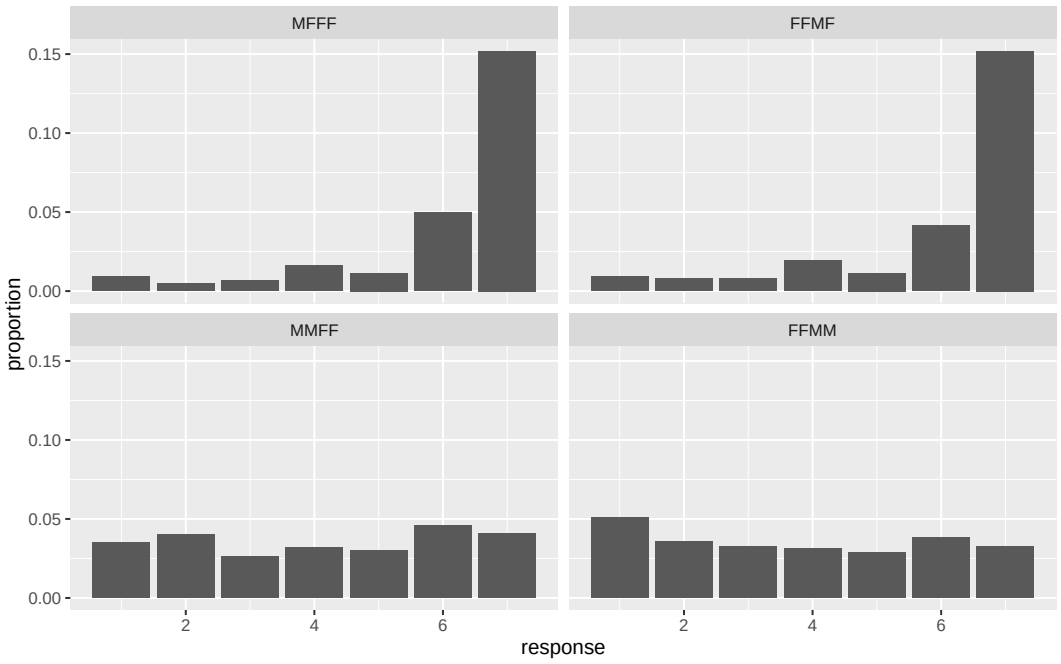


Figure 4: All responses by all participants according to whether the speaker liked or disliked the sentences in the MMMM condition.

- To test whether the responses for the different conditions were significantly different from each other we fitted an ordinal[3] regression model with only `condition`[4] as a dependent variable, and `participant` and `hybrid_noun` as random effects.

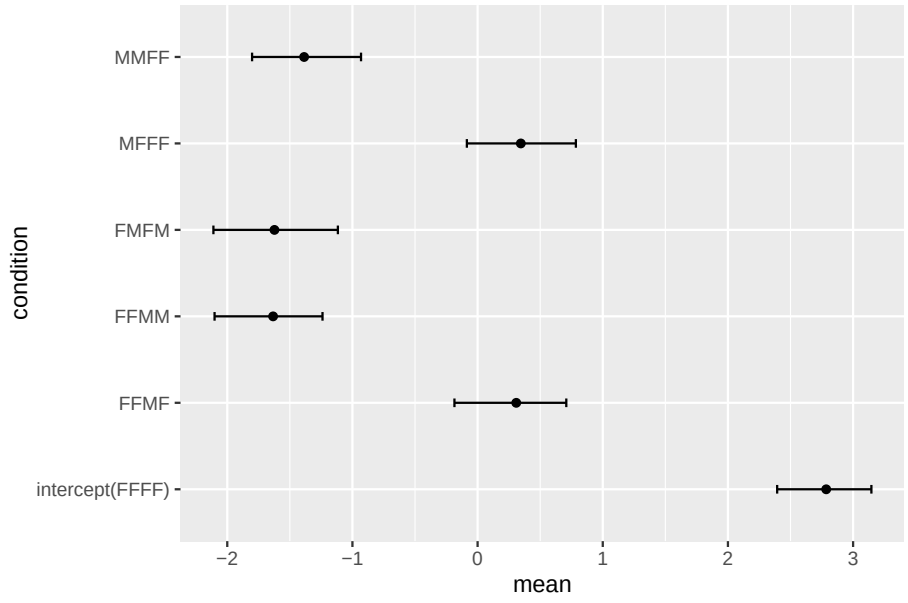


Figure 5: Posterior means and 95% confidence intervals for the Bayesian regression model.

[3]Ordinal regression assumes an ordered discrete response variable. This is exactly the kind of data one obtains from grammaticality judgment tasks.

[4]In the models, `gender` and `region` did not play a role.

- The factors with overlapping confidence intervals are not statistically different from each other.
- We see that FFMF and MFFF slightly overlap with 0, this means that they are not statistically different from the intercept (FFFF).
- Meanwhile, FMFM, FFMM and MMFF are statistically worse than the intercept, but not different from each other.

5 Analysis

- Summary of the results:

(30) *Two-way mismatches possible with feminine agreement*:

a. Jovan je redovn-a mušterija, a Milica povremen-a ⟨mušterija⟩.
Jovan is regular-F customer but Milica occasional-F customer
'Jovan is a regular customer and Milica an occasional one.' (MFFF)

b. Milica je povremen-a mušterija, a Jovan redovn-a ⟨mušterija⟩.
Milica is occasional-F customer but Jovan regular-F customer
'Milica is an occasional customer and Jovan a regular one.' (FFMF)

(31) *No mismatch possible with masculine agreement*:

a. *Milica je povremen-a mušterija, a Jovan redovn-i ⟨mušterija⟩.
Milica is occasional-F customer but Jovan regular-M customer
'Milica is an occasional customer and Jovan a regular one.' (FFMM)

b. *Jovan je redovn-i mušterija, a Milica povremen-a ⟨mušterija⟩.
Jovan is regular-M customer but Milica occasional-F customer
'Jovan is a regular customer and Milica an occasional one.' (MMFF)

- It seems that when the adjective agrees in feminine, hybrid nouns pattern behave like the two-way alternating *doctor*-class nouns, such as *jatros* in Greek:

(32) a. O Petros ine kalos jatros, ala i Maria ine mia kakia ⟨jatros⟩.
the Petros is good.MASC doctor but the Maria is a.FEM bad.FEM doctor
'Petros is a good doctor, but Maria is a bad one.'

b. I Maria ine kali jatros, ala o Petros ine enas kakos ⟨jatros⟩.
the Maria is good.FEM doctor but the Petros is a.MASC bad.MASC doctor
'Maria is a good doctor, but Petros is a bad one.'

(*Greek*; Merchant 2014:15)

- But when the noun triggers masculine agreement, these nouns seem to behave like the non-alternating *brother*-type nouns:

(33) a. *O Petros ine kalos adherfos, ala i Maria ine mia kakia ⟨adherfi⟩.
the Petros is good.MASC brother but the Maria is a.FEM bad.FEM sister
'*Int.* Petros is a good brother, but Maria is a bad one (sister).'

b. *I Maria ine kali adherfi, ala o Petros ine enes kakos ⟨adherfos⟩.
the Maria is good.FEM sister but the Petros is a.MASC bad.MASC brother
'*Int.* Petros is a good brother, but Maria is a bad one (sister).'

(*Greek*; Merchant 2014:12)

5.1 Ellipsis identity

- Following Merchant (2001), let us assume that the identity requirement for ellipsis involves *e-GIVENness* (i.e. mutual entailment):

(34) *e-GIVENness* (Merchant 2001:26):
An expression E counts as e-GIVEN iff E has a salient antecedent A and, modulo ∃-type shifting
(i) A entails F-clo(E), and
(ii) E entails F-clo(A)

- This prevents non-matching ellipsis sites Merchant (2001:27):

(35) a. Abby *called Ben an idiot* after Mary did ⟨call Ben an idiot⟩.
($\exists x.x$ called Ben an idiot ↔ $\exists x.x$ called Ben an idiot)
b. #Abby *called Ben an idiot* after Mary did ⟨insult Ben⟩.
($\exists x.x$ called Ben an idiot ↮ $\exists x.x$ insulted Ben)

- Following Cooper (1983), it is often assumed that (natural) gender features can introduce presuppositions (also see Sauerland 2003, 2008; Heim 2008; Kratzer 2009; Spathas 2010; Sudo 2012).

- Non-alternating nouns in the *brother*-class contain a presupposition about gender (Merchant 2014:19, Sudo & Spathas 2016:715).

(36) a. $[\![\text{adherfos}]\!] = \lambda x_{\langle e \rangle}$: x is male. sibling(x)
b. $[\![\text{adherfi}]\!] = \lambda x_{\langle e \rangle}$: x is female. sibling(x)

(37) a. *O Petros ine kalos adherfos, ala i Maria ine mia kakia ⟨adherfi⟩.
the Petros is good.MASC brother but the Maria is a.FEM bad.FEM sister
'*Int.* Petros is a good brother, but Maria is a bad one (sister).'
b. *I Maria ine kali adherfi, ala o Petros ine enes kakos ⟨adherfos⟩.
the Maria is good.FEM sister but the Petros is a.MASC bad.MASC brother
'*Int.* Petros is a good brother, but Maria is a bad one (sister).'
(*Greek*; Merchant 2014:12)

(38) $\exists x$: x is male. sibling(x) ↮ $\exists x$: x is female. sibling(x)

- In the analysis of two-alternating 'epicene' nouns, however, it is often assumed that they do not contain any lexical presuppositions about gender (39), and thus ellipsis is licensed (40).

(39) $[\![\text{jatros}]\!] = \lambda x_{\langle e \rangle}$. doctor($x$)

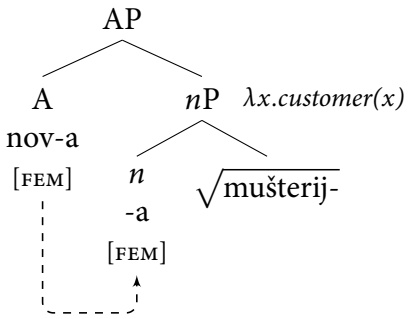
(40) $\exists x$. doctor(x) ↔ $\exists x$. doctor(x)

(41) a. O Petros ine kalos jatros, ala i Maria ine mia kakia ⟨jatros⟩.
the Petros is good.MASC doctor but the Maria is a.FEM bad.FEM doctor
'Petros is a good doctor, but Maria is a bad one.'
b. I Maria ine kali jatros, ala o Petros ine enas kakos ⟨jatros⟩.
the Maria is good.FEM doctor but the Petros is a.MASC bad.MASC doctor
'Maria is a good doctor, but Petros is a bad one.'
(*Greek*; Merchant 2014:15)

5.2 The structure of hybrid nouns

- Recall that we saw that although two-way mismatches were possible, this depended on agreement type (i.e. feminine agreement).
- Importantly, our findings are not directly compatible with analyses such as Merchant (2014) and Sudo & Spathas (2016), as they assume that the adjective and determiner agrees with the referent rather than the elided noun.
- We assume the following basic representations for hybrid nouns such as *mušterija* ('customer'):

(42)



- Since the denotation of the *n* head does not contribute anything about gender, this *n*P is compatible with both female and male referents.

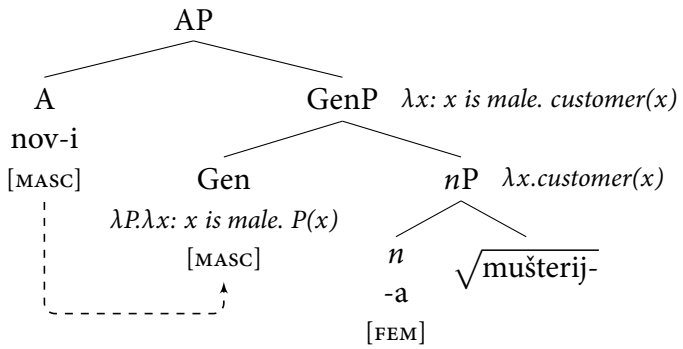
(43) a. Marija nam je nov-a mušterija.
Marija us is new-F customer
'Marija is our new customer.'

b. Milan nam je nov-a mušterija.
Milan us is new-F customer
'Milan is our new customer.'

- For masculine-agreeing cases, we assume an additional projection for (natural) gender GenP (cf. Picallo 1991) that is a partial identity function constraining the set of customers to the set of male customers (Cooper 1983).

(44) $[\![\text{Gen}_{[\text{MASC}]}]\!] = \lambda P.\lambda x\text{: } x \text{ is male. } P(x)$

(45)



$[\![nP]\!] = \lambda x.\text{customer}(x)$

$$\begin{aligned}[\![\text{GenP}]\!] &= [\![\text{Gen}]\!]([\![nP]\!]) \\ &= [\lambda P.\lambda x\text{: } x \text{ is male. } P(x)](\lambda x.\text{customer}(x)) \\ &= \lambda x\text{: } x \text{ is male. } [\lambda x'.\text{customer}(x')](x) \\ &= \lambda x\text{: } x \text{ is male. customer}(x)\end{aligned}$$

- The introduction of the presupposition with male agreement means that it is incompatible with female referents.

(46) Marija nam je nov-a mušterija.
Marija us is new-F customer
'Marija is our new customer.'

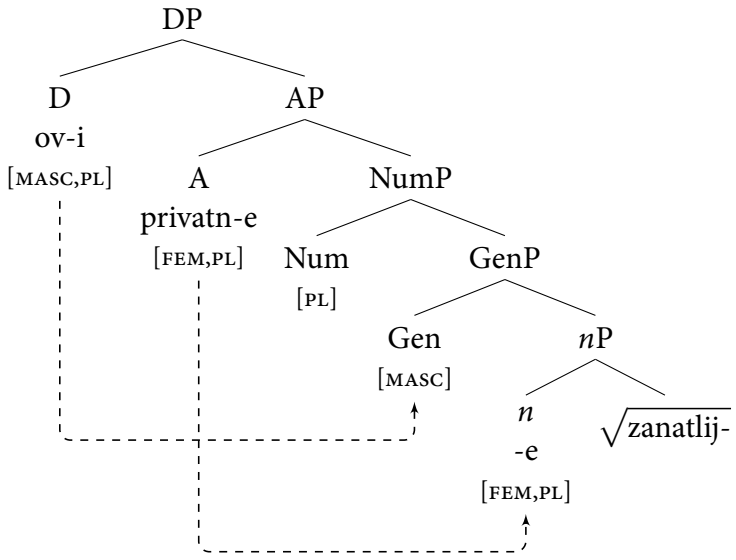
(47) *Marija nam je nov-i mušterija.
Marija us is new-M customer
'Marija is our new customer.'

(48) ⟦novi mušterija⟧ = λx. customer(x), defined only if x is male
a. ⟦(46)⟧ = customer(Marija)
b. ⟦(47)⟧ = customer(Marija), defined only if Marija is male (*presupposition failure!*)

- While other theories rule out such examples based on competition with the feminine agreeing form (e.g. *Maximize Presupposition*; Bobaljik & Zocca 2011:148f., *Principle of Gender Competition*; Sudo & Spathas 2016:722), we argue that the syntactic presence of both features is necessary based on instances of mixed agreement, such as (49).

(49) ov-i privatn-e zanatlije
these-**M.PL** private-**F.PL** artisan.PL
'these private artisans' (Corbett 2006:206)

(50)



- As such, the masculine agreeing variant of *mušterija* is built on top of the feminine variant.

5.3 Markedness

- In the proposed structure, feminine is properly contained in the representation of masculine.
- For a number of authors, this corresponds to saying that (with these hybrid nouns) masculine is the more marked gender (see Caha 2009, Bobaljik 2012 and Smith et al. 2015).
- Some of the markedness diagnostics utilized by Bobaljik & Zocca (2011); Merchant (2014); Sudo & Spathas (2016) in fact seem to indicate that [feminine] is the unmarked gender in hybrid nouns.
- For instance, with feminine agreement, these nouns in the plural can denote male, or female, or mixed gender groups.

(51) a. Mušterije su se posvađale.
customers are REFL argued.F.PL
'Customers (male group/female group/mixed group) had an argument.'

b. Mušterije su se posvađali.
customers are REFL argued.M.PL
'Customers (male group) had an argument.'[5]

- In negative existential sentences, the unmarked gender should not restrict the domain of quantification only on particular individuals. As we see in (52), feminine gender shows no restrictions, while in (53), masculine restricts the referents to males.

(52) Petar nema novu mušteriju.
Petar not.has new.F.SG customer
⇒ 'Petar does not have a male customer.'
⇒ 'Petar does not have a female customer.'

(53) Petar nema novog mušteriju.
Petar not.has new.M.SG customer
⇒ 'Petar does not have a male customer.'
⇏ 'Petar does not have a female customer.'

- Finally, Sudo & Spathas (2016:719) also apply the test with focus alternatives. Under the assumption that ϕ-features are presuppositions, they should not be ignored in focus alternatives.

(54) a. Only John is a male athlete. ⇏ Mary is not a female athlete.
b. Only Mary is a female athlete. ⇏ John is not a male athlete.

- The focus test also seems to indicate that feminine gender is the less marked one with these hybrid nouns.

(55) Jedino Petar je nova mušterija.
only Petar is new.F.SG customer
'Only Petar is a new customer.'
⇒ Marko is not a new customer.
⇒ Marija is not a new customer.

(56) Jedino Petar je novi mušterija.
only Petar is new.M.SG customer
'Only Petar is a new customer.'
⇒ Marko is not a new customer.
?⇏ Marija is not a new customer.

5.4 Mismatches under ellipsis

- With these assumptions in place, we now analyse our findings for gender mismatches with NPE.
- *Result*: gender mismatches are possible in either direction as long as the adjectival agreement is feminine.

(57) *Two-way mismatches possible with feminine agreement*:
a. Jovan je redovn-a mušterija, a Milica povremen-a ⟨mušterija⟩.
Jovan is regular-F customer but Milica occasional-F customer
'Jovan is a regular customer and Milica an occasional one.' (MFFF)
b. Milica je povremen-a mušterija, a Jovan redovn-a ⟨mušterija⟩.
Milica is occasional-F customer but Jovan regular-F customer
'Milica is an occasional customer and Jovan a regular one.' (FFMF)

[5] Theoretically, masculine agreement here could also denote that the group is mixed, in which case the given agreement would not be agreement in natural gender, but rather default agreement. We leave the issue of default agreement for further research, while noting that masculine agreement seems to be less restricted than feminine with these nouns.

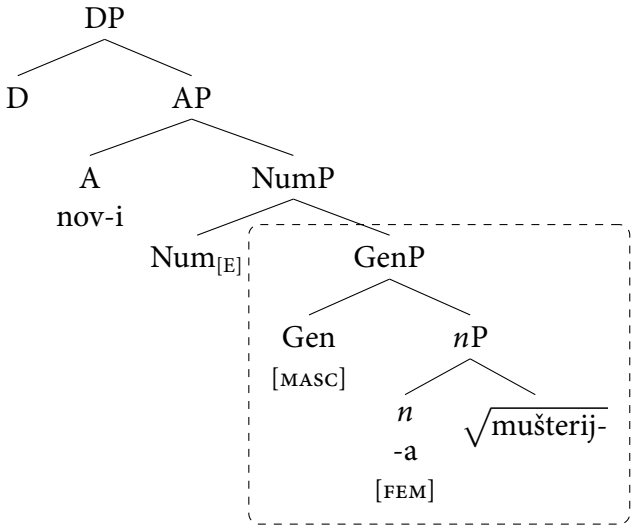
(58) *No mismatch possible with masculine agreement*:

a. *Milica je povremen-a mušterija, a Jovan redovn-i ⟨mušterija⟩.
Milica is occasional-F customer but Jovan regular-M customer
'Milica is an occasional customer and Jovan a regular one.' (FFMM)

b. *Jovan je redovn-i mušterija, a Milica povremen-a ⟨mušterija⟩.
Jovan is regular-M customer but Milica occasional-F customer
'Jovan is a regular customer and Milica an occasional one.' (MMFF)

- Under present assumptions, this means that having the masculine gender projected (a GenP projection) in either the antecedent or ellipsis site results in ungrammaticality.
- First, let us assume that NP ellipsis is triggered by an [E]-feature on the head of NumP (e.g. Merchant 2014; Lipták & Saab 2014; Saab & Lipták 2016; Saab to appear):

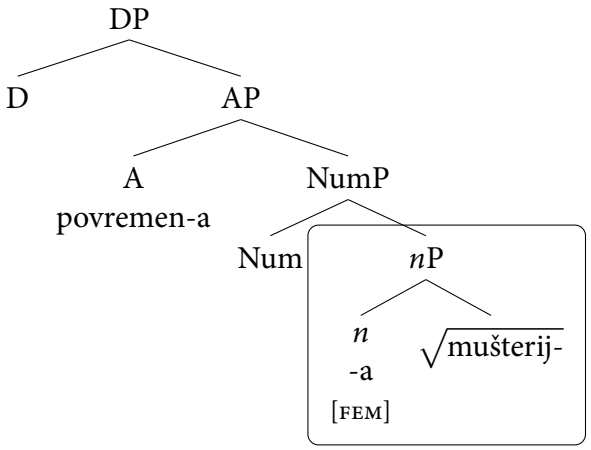
(59)



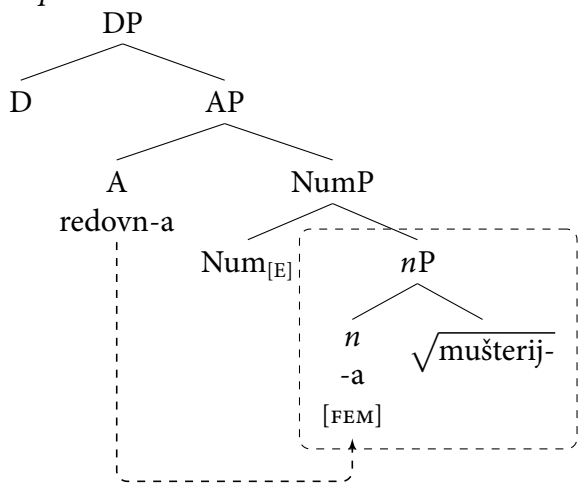
- In an example such as (60), the NP in the ellipsis site agrees in feminine – it must therefore lack a GenP projection.

(60) Milica je povremen-a mušterija, a Jovan redovn-a ⟨mušterija⟩.
Milica is occasional-F customer but Jovan regular-F customer
'Milica is an occasional customer and Jovan a regular one.' (FFMF)

(61) *Antecedent*:

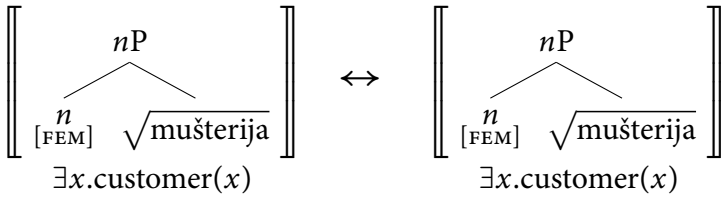


(62) *Ellipsis site*:



- Importantly, mutual entailment (e-GIVENness) is satisfied here since the elided material is identical.

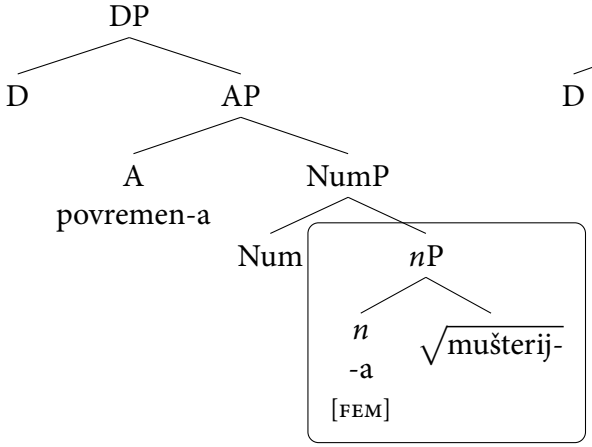
(63)



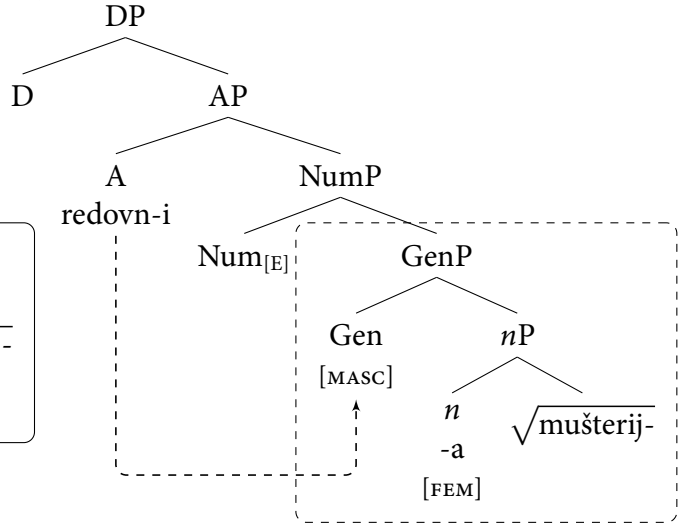
- However, as soon as we have masculine agreement in one of the conjuncts, we introduce a new masculine feature:

(64) *Milica je povremen-a mušterija, a Jovan redovn-i ⟨mušterija⟩.
Milica is occasional-F customer but Jovan regular-M customer
'Milica is an occasional customer and Jovan a regular one.' (FFMM)

(65) *Antecedent*:

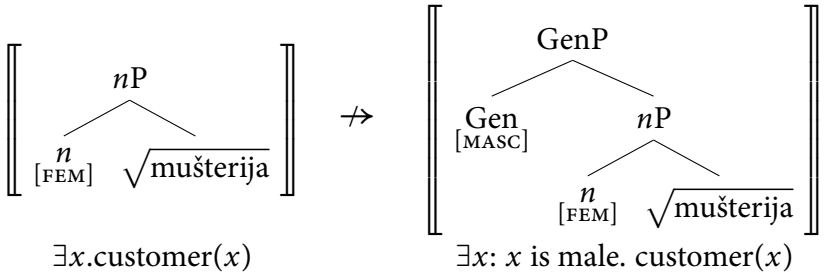


(66) *Ellipsis site*:



- Recall that the Gen head introduces the presupposition that its referent is male.
- If there cannot be a corresponding Gen head in the other conjunct (i.e. if there is a feminine referent), then e-GIVENness is violated and ellipsis is not licensed:

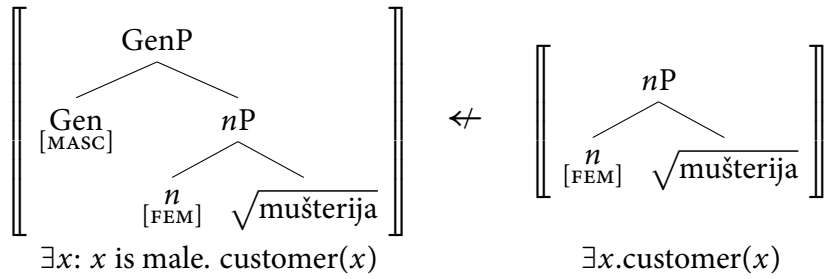
(67)



- The same is true if masculine agreement is in the antecedent:

(68) *Jovan je redovn-i mušterija, a Milica povremen-a ⟨mušterija⟩.
Jovan is regular-M customer but Milica occasional-F customer
'Jovan is a regular customer and Milica an occasional one.' (MMFF)

(69)



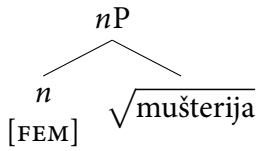
$\exists x$: x is male. customer(x) $\nleftarrow$ $\exists x$.customer(x)

- This is why mismatches in referent gender are not tolerated if the adjective agrees in masculine in one conjunct only.

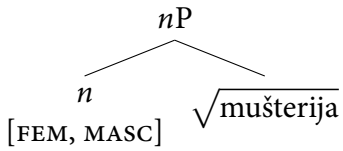
6 Conclusion

- We have shown that hybrid nouns in BCS seem to contain both masculine and feminine gender features.
- Using the (im)possibility of gender mismatches under ellipsis as a diagnostic tool, we argue that hybrid nouns are inherently feminine (grammatical gender) and can be additionally specified for masculine (natural gender) under certain circumstances.
- Two-way gender mismatches under NPE are tolerated only if there is feminine agreement. Introducing masculine agreement into either the antecedent or the ellipsis site results in ungrammaticality.
- We suggest that this is because masculine gender features introduce an additional gender presupposition that destroys mutual entailment.
- Specifically, we have argued for an additional structural GenP projection, however it seems that one could achieve the same basic result with two variatins of *n*:

(70)



(71)



- We leave it up to future research to tease apart the difference between these two types of representation.

7 Acknowledgements

We would like to thank Ana Bosnić, Peter Staroverov and Joanna Zaleska for their help in the experimental design, all the participants for their help and participation in the experiment, as well as Jelena Stojković, Boban Arsenijević, Andrew Nevins, Jana Willer-Gold and Jason Mechant for useful comments and feedback.

References

Barrie, Michael (2016): 'Gender mismatches and ellipsis in Cayuga', *Journal of Cognitive Science* **17**(3), 361–388.

Bobaljik, Jonathan D. & Cynthia Levart Zocca (2011): 'Gender Markedness: The anatomy of a counterexample', *Morphology* **21**, 141–166.

Bobaljik, Jonathan David (2012): *Universals in Comparative Morphology: Suppletion, Superlatives and the Structure of Words*. MIT Press, Cambridge, MA.

Caha, Pavel (2009): The Nanosyntax of Case. PhD thesis, University of Tromsø.

Cooper, Robin (1983): *Quantification and Syntactic Theory*. Reidel, Dordrecht.

Corbett, Gerville (1991): *Gender*. Cambridge University Press, Cambridge.

Corbett, Greville (2006): *Agreement*. Cambridge University Press, Cambridge.

Hadfield, Jarrod D (2010): 'MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package', *Journal of Statistical Software* **33**(2), 1–22.
URL: *http://www.jstatsoft.org/v33/i02/*

Heim, Irene (2008): Features on bound pronouns. *In:* D. Harbour, D. Adger & S. Bejar, eds, *Phi Theory: Phi-Features Across Modules and Interfaces*. Oxford University Press, Oxford, pp. 35–56.

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

Kratzer, Angelika (2009): 'Making a Pronoun: Fake Indexicals as Windows into the Properties of Pronouns', *Linguistic Inquiry* **40**(2), 187–237.

Landau, Idan (to appear): 'DP-internal semantic agreement: A configurational analysis', *Natural Language and Linguistic Theory* .

LimeSurvey Project Team, Carsten Schmitz (2012): *LimeSurvey: An Opeen Source survey tool*. LimeSurvey Project, Hamburg, Germany. URL: http://www.limesurvey.org.

Lipták, Anikó & Andrés Saab (2014): 'No N-raising out of NPs in Spanish: Ellipsis as a diagnostic of head movement', *Natural Language and Linguistic Theory* **32**(1247–1271).

Merchant, Jason (2001): *The Syntax of Silence: Sluicing, Islands and the Theory of Ellipsis*. Oxford University Press, Oxford.

Merchant, Jason (2014): 'Gender mismatches under nominal ellipsis', *Lingua* **151**, 9–32.

Nunes, Jairo & Cynthia Levart Zocca (2010): Lack of morphological identity and ellipsis resolution in Brazilian Portugese. *In:* J. Nunes, ed., *Minimalist essays on Brazilian Portuguese syntax*. John Benjamins, Amsterdam, pp. 215–236.

Pesetsky, David (2013): *Russian case morphology and the syntactic categories*. MIT Press, Cambridge, MA.

Picallo, M. Carme (1991): 'Nominals and Nominalizations in Catalan', *Probus* **3**, 279–316.

R Core Team (2016): *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
URL: *https://www.R-project.org/*

Saab, Andrés (to appear): Nominal ellipses. *In:* J. van Craenenbroeck & T. Temmerman, eds, *The Oxford Handbook of Ellipsis*. Oxford University Press, Oxford.

Saab, Andrés & Anikó Lipták (2016): 'Movement and deletion after syntax: Licensing by inflection reconsidered', *Studia Linguistica* **70**(1), 66–108.

Sauerland, Uli (2003): A new semantics for number. *In:* R. B. Young & Y. Zhou, eds, *Proceedings of SALT 13*. Cornell Linguistics Club, Ithaca, NY, pp. 258–275.

Sauerland, Uli (2008): On the semantics markedness of phi-features. *In:* D. Adger, D. Harbour & S. Bejar, eds, *Phi Theory: Phi-Features across Modules and Interfaces*. Oxford University Press, Oxford, pp. 57–82.

Smith, Peter (2015): Feature Mismatches: Consequences for Syntax, Morphology and Semantics. PhD thesis, University of Connecticut.

Smith, Peter (2016): The Agreement Hierarchy and AGREE. Ms. Goethe-Universität Frankfurt. lingbuzz/003194.

Smith, Peter W., Beata Moskal, Ting Xu, Jungmin Kang & Jonathan D. Bobaljik (2015): Pronominal Suppletion: Case and Number. *In:* T. Bui & D. Özyıldız, eds, *Proceedings of NELS 45*. .

Spathas, Giorgos (2010): Focus on Anaphora. PhD thesis, Universiteit Utrecht.

Sudo, Yasutada (2012): On the Semantics of Phi Features on Pronouns. PhD thesis, MIT, Cambridge,

MA.

Sudo, Yasutada & Giorgos Spathas (2016): Nominal ellipsis and the interpretation of gender in Greek. *In:* N. Bade, P. Berezovskaya & A. Schöller, eds, *Proceedings of Sinn und Bedeutung 20*. pp. 712–724.

Wechsler, Stephen & Larisa Zlatić (2003): *The many faces of agreement*. Center for the Study of Language and Information, Stanford.

8 Appendix

We use an ordinal bayesian regression model with the package `MCMCglmm` (Hadfield 2010) in R (R Core Team 2016). We used non-informative priors (an inverse gamma with V=1 and nu=0.002). Table 2 presents the posterior mean estimates, the confidence intervals and equivalent of a bayesian p value. The corresponding posterior estimates of the random effects can be are shown in Table 3

	post.mean	l-95% CI	u-95% CI	effect sample	pMCMC	
(Intercept)	2.7552	2.3839	3.1941	354.0	<0.002	**
condition FFMF	0.3172	-0.1073	0.7484	600.0	0.1433	
condition FFMM	-1.6084	-2.0599	-1.2233	487.7	<0.002	**
condition FMFM	-1.6095	-2.0645	-1.0990	600.0	<0.002	**
condition MFFF	0.3723	-0.0990	0.7610	516.1	0.0933	.
condition MMFF	-1.3541	-1.8228	-0.9690	514.2	<0.002	**
Cutpoints:						
	post.mean	l-95% CI	u-95% CI	effect sample		
cutpoint trait value.1	0.6284	0.5391	0.7129	90.48		
cutpoint trait value.2	1.0485	0.9316	1.1437	62.66		
cutpoint trait value.3	1.5598	1.4347	1.6683	30.25		
cutpoint trait value.4	1.9236	1.8068	2.0575	29.95		
cutpoint trait value.5	2.6803	2.5527	2.8159	43.93		

Table 2: Coefficients for the MCMC model with confidence intervals and cutpoints.

	post.mean	l-95% CI	u-95% CI	effect sample
participant	0.3557	0.2037	0.512	478.8
	post.mean	l-95% CI	u-95% CI	effect sample
hybrid noun	0.06491	0.004233	0.1571	600

Table 3: Random effects for the MCMC model.