Chapter 14

Gender encoding on hybrid nouns in Bosnian/Croatian/Serbian: Experimental evidence from ellipsis

Andrew Murphy
University of Leipzig

Zorica Puškar
Leibniz-Zentrum Allgemeine Sprachwissenschaft, Berlin

Matías Guzmán Naranjo
Heinrich Heine University Düsseldorf

In this paper, we report the results of an experimental study on the possibility of gender mismatches with ellipsis of a particular type of hybrid nouns in Bosnian / Croatian / Serbian (henceforth: BCS). Using agreement mismatches under NP ellipsis as a diagnostic for gender feature specification of hybrid nouns, we show that these nouns disallow agreement mismatches under NP ellipsis if natural gender (and a presupposition that it introduces) is present either in the antecedent or the ellipsis site. We argue that natural gender feature (masculine in our study) can optionally be present on the hybrid noun, and that its inclusion in ellipsis contexts leads to a violation of the standard identity requirement between the antecedent and the ellipsis site, namely Merchant's (2001) e-ɢɪᴠᴇɴness.

Keywords: gender features, ellipsis, agreement, mismatches

1 Introduction: Hybrid nouns in BCS

Hybrid nouns have been a challenge for theories of agreement and NP structure (see Corbett 1991; Wechsler & Zlatić 2003; Alsina & Arsenijević 2012b,a; Peset-

sky 2013; Kramer 2015; Landau 2016; Smith 2015; 2016; Arsenijević & Gračanin-Yuksek 2015; Despić 2017), as it seems that they can simultaneously bear two types of gender specification: (i) natural gender (reflecting the gender of the referent) and (ii) grammatical gender (assigned arbitrarily). In this paper, we focus only on the Class II hybrid nouns in BCS (ending in *-a*) which have grammatical feminine gender, but variable natural gender, which depends on the discourse referent.

Table 1: Some hybrid nouns in BCS

-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'

For instance, with a masculine referent, the noun *mušterija* ('customer') can trigger either grammatical feminine (1a)/(1b) or natural masculine agreement (1c)/(1d) (subject to speaker variability).

(1) 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, however, such nouns can only trigger feminine agreement:

(2) Marija nam je nov-a / *nov-i   mušterija.
     Marija us   is new-F   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 simply listed in the lexicon twice, each with a different gender feature (e.g. Merchant 2014: 19). However, such an approach to BCS hybrid nouns seems to be problematic, as there is evidence that these nouns can *simultaneously* bear natural and grammatical gender. For example, in (3) the adjective and determiner target different gender values of the noun, while (4) shows the same for the nominal modifier and the participle.[1]

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

(4) Osm-a   budala je bio-∅   mnogo kul tip ali su ga se
     eighth-F.SG fool.SG is been-M.SG very   cool guy but are him REFL
     drugi malo plašili.[2]
     others little feared
     'The eighth fool was a very cool guy, but others were a bit afraid of him.'

These examples raise the question about the structural representation of such nouns: whether they do contain both types of gender feature simultaneously, and how exatly they should be represented. In what follows, we will answer these questions by using NP ellipsis as a diagnostic for the gender feature specification of BCS hybrid nouns.

---

[1]Relative pronouns can also show different gender values than the attributive modifiers of a hybrid noun, as shown in (i), obtained at http://vukajlija.com/seoski-fudbalski-tim-iz-beton-lige <accessed 26.11.2016>.

(i) 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...'

However, as pointed out by a reviewer, whether the relative pronoun agrees directly with the hybrid noun, or whether agreement is more indirect (since we are dealing with a non-restrictive relative clause in which the relative pronoun is more akin to regular pronouns, c.f. de Vries 2006) is an issue that requires further investigation. See Arsenijević & Gračanin-Yuksek (2015) for further discussion on hybrid agreement with relative pronouns.

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

2 Gender mismatches and NP ellipsis

There is a growing body of literature on the permissibility of gender mismatches under NP ellipsis (e.g. Nunes & Zocca 2010; Bobaljik & Zocca 2011; Merchant 2014; Sudo & Spathas 2016; Barrie 2016). Based on whether gender mismatches under NP ellipsis are allowed or not, previous literature has identified three classes of nouns. The first type are the TWO-WAY ALTERNATING NOUNS (henceforth: the *doctor*-class), where a masculine antecedent can license deletion of a noun with a feminine referent, and *vice versa:*

(5) **Two-way alternating nouns** (*doctor*-class)

    a. O   Petros ine kalos   jatros, ala i   Maria ine mia kakia ⟨jatros⟩.
       the Petros is   good.M doctor but the Maria is   a.F bad.F 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.F doctor but the Petros is   a.M bad.M doctor
       'Maria is a good doctor, but Petros is a bad one.'

                                            (Greek; Merchant 2014: 15)

The second type are the NON-ALTERNATING NOUNS (henceforth: *brother*-class), which do not allow mismatches in either direction:

(6) **Non-alternating nouns** (*brother*-class)

    a. *O   Petros ine kalos   adherfos, ala i   Maria ine mia kakia
       the Petros is   good.M brother   but the Maria is   a.F bad.F
       ⟨adherfi⟩.
        sister
       Intended: 'Petros is a good brother, but Maria is a bad one (sister).'

    b. *I   Maria ine kali   adherfi, ala o   Petros ine enes kakos
       the Maria is   good.F sister   but the Petros is   a.M bad.M
       ⟨adherfos⟩.
        brother
       Intended: 'Petros is a good brother, but Maria is a bad one (sister).'
                                            (Greek; Merchant 2014: 12)

Finally, ONE-WAY ALTERNATING NOUNS (or the *actor*-class) allow a masculine noun to antecede an elided feminine noun (7a), but a mismatch in the opposite direction is ungrammatical (7b).

(7) **One-way alternating nouns** (*actor*-class)

    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
       Intended: 'Fernanda is an actress and Paulo is (an actor) too.'

                      (Brazilian Portuguese; Bobaljik & Zocca 2011: 142)

With regard to their tolerance for gender mismatches, these three classes can be summarized as in Table 2:

Table 2: Classes of predicative nouns under ellipsis (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. However, the additional complication with hybrid nouns is that there is evidence for the simultaneous presence of two gender features. Thus, we will try to gain some insight into their structure by testing the acceptability of gender mismatches with mismatched referents in either direction (8b)/(8c).

(8)   a. Milan mu je star-**a** mušterija, a   Marija mu je nov-**a** ⟨mušt.⟩.
        Milan him is old-[F] customer and Marija him is new-[F]   cust.
        'Milan is his old customer and Marija a new one.'

    b. ? Milan mu je star-**i** mušterija, a   Marija mu je nov-**a** ⟨mušt.⟩.
        Milan him is old-[M] customer and Marija him is new-[F]   cust.
        'Milan is his old customer and Marija a new one.'

    c. ? Marija mu je star-**a** mušterija, a   Milan mu je nov-**i** ⟨mušt.⟩.
        Marija him is old-[F] customer and Milan him is new-[M]   cust.
        'Marija is his old customer and Milan a new one.'

3 Experiment

The aim of the experiment was to discover whether an agreement mismatch was tolerated when masculine agreement was found either the antecedent (8b) or the ellipsis site (8c), i.e. whether the hybrid nouns under study were two-way alternating, one-way alternating, or allowed no alternation at all when the noun had natural masculine gender (indicated by the agreement on the adjective). Considering previous theories (Nunes & Zocca 2010; Bobaljik & Zocca 2011; Merchant 2014; Sudo & Spathas 2016), the (im)possibility of a mismatch should be an indicator of (i) the difference in the quality of gender features (i.e. differences in ther semantics or morphosyntactic representation) and (ii) the licensing conditions on ellipsis (i.e. identity requirements between the antecedent and the ellipsis site). Finally, sentences in which both adjectives show feminine agreement, regardless of the gender of the subject, were expected to be grammatical, as the hybrid nouns in question have grammatical gender as a formal feature.

In order to verify the grammaticality of gender mismatches under NP ellipsis, we ran an online acceptability judgement study.[3] The experimental design involved the factors in (9). The first factor involves the type of agreement the adjective has with a masculine subject and has two levels (grammatical agreement MF vs. natural agreement MM) (9a), which should serve as an indicator of the type of gender on the hybrid noun (grammatical vs. natural). The second factor pertains to agreement with feminine subjects and has only a single level (FF, both grammatical and natural agreement) (9b). The final factor regards the position of the masculine referent and has two levels: in the first clause, or in the second clause (the one with NP ellipsis) (9c), which should indicate whether natural gender can function as an antecedent for ellipsis and whether it can be found in the ellipsis site. The $2 \times 2 \times 1$ combination of factors in (9) yield the four experimental conditions in Table 3.

(9) **Factors** (manipulated within items)

    a. AGREEMENT WITH MASCULINE SUBJECT; two levels: grammatical (MF) and natural (MM)

    b. AGREEMENT WITH FEMININE SUBJECT; one level: FF

    c. CLAUSE; two levels: first and second

Some example test sentences for each condition are given below. Each contains two clauses coordinated by the conjunction *a* 'and' and an elided noun in the

---

[3]For a detailed description of the design and for all the materials see https://osf.io/r3npz.

Table 3: Experimental conditions

	Factors and levels			Conditions
	M.SUBJ.AGR	CLAUSE	F.SUBJ.AGR	
1	MF	first	FF	MFFF
2	MM	first	FF	MMFF
3	MF	second	FF	FFMF
4	MM	second	FF	FFMM

second conjunct. Sentences (10) and (12) were expected to be perceived as grammatical, while (11) and (13) were the relevant mismatching cases.

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

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

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

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

Since masculine agreement with hybrid nouns is not accepted by all speakers, a control sentence involved only masculine referents and agreement (i.e. the condition MMMM) (14).

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

Furthermore, we included the following additional controls: a grammatical baseline with feminine agreement and all feminine referents FFFF (15), and an ungrammatical baseline FMFM, involving masculine agreement with feminine referents (16).
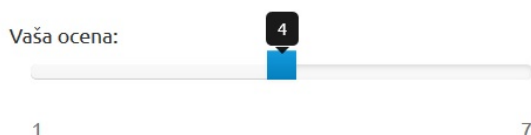
❋ Jovan je redovna mušterija, a Milica povremena.

Vaša ocena: 4

1   7

Figure 1: Example experimental item

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

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

A total of 18 controls were included, 6 per combination (all of them the same in every list, see below). There were 96 test sentences altogether, with 24 sentences per condition. All lexical items were balanced for proper names (24 male, 24 female), adjectives (48) and hybrid nouns (6). We used a Latin square design where the 96 test sentences 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. The experiment was coded using LimeSurvey[4] and run online via the LimeService platform. Sentences were presented one-by-one in a random order. Each participant saw 62 sentences (24 test items + 18 controls + 20 fillers) and was asked to give a grammaticality judgement on a 7-point Likert scale (1 = completely bad, 7 = sounds excellent) by dragging a centered slider either to the left or right to indicate their response (see Figure 1).

The experiment was performed by 164 volunteers, 131 female and 33 male, aged 16–66. Participants reportedly spoke different varieties of BCS: Bosnian (22 speakers), Croatian (5 speakers) and Serbian (136 speakers). None of the participants were paid or otherwise compensated for their participation.
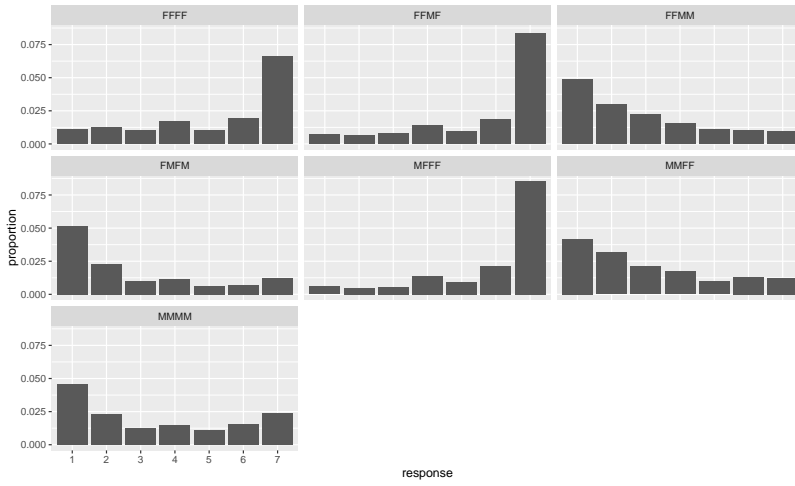
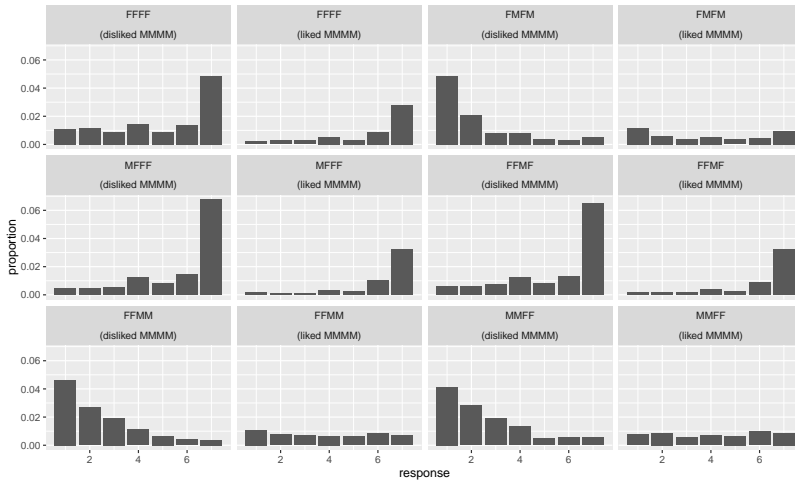Figure 2: All responses by all participants.

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

4 Results

Figure 2 shows the distribution of responses under each condition from all participants for the baselines (first column) and our four experimental conditions (second and third column). The baselines show strong grammaticality effects: feminine agreement with feminine subjects was rated as grammatical, while masculine agreement was unacceptable. A u-shaped type of distribution for MMMM (masculine agreement with masculine subjects) suggests that although the majority of speakers dispreferred it, some speakers clearly did find it grammatical. With mismatching subjects, feminine agreement (FFMF and MFFF) was rated as grammatical, as shown in the second column. More gradient (un)acceptability was found for FFMM and MMFF (third column).

We compared the responses for all conditions based on whether speakers liked the MMMM combination (median rating ≥ 4) or disliked it (Figure 3). In total, 51 speakers found the MMMM combination grammatical. The overall picture shows that the distributions concerning the grammatical patterns are fundamentally the same, regardless of whether the speaker liked MMMM or not (the first two rows). However, there was in fact a difference for conditions with low acceptability scores (FFMM and MMFF), as it can be seen in the final row of Figure 3. Speakers who liked MMMM showed no clear preference or dispreference for either of the mismatching combinations – they were perceived as equally bad.

Figure 4 shows the four crucial conditions just for those speakers who rated the MMMM baseline higher than 4. We see that the patterns with feminine agreement throughout were overwhelmingly acceptable and grammatical to these speakers, while the conditions with mismatches show more variation and received comparably lower scores.

To clarify whether the differences in these responses were statistically significant, we fitted an ordinal regression model with only `condition` as a dependent variable, and `participant` and `hybrid_noun` as random effects (see the Appendix for further details about the model).[5]

As the plot in Figure 5 shows, the factors with overlapping confidence intervals are not statistically different from each other. We see that FFMF and MFFF slightly overlap with 0, which means that they are not statistically different from the intercept (the grammatical baseline FFFF). On the other hand, FFMM and MMFF

---

[4]https://limesurvey.org

[5]Ordinal regression assumes an ordered discrete response variable. This is exactly the kind of data one obtains from grammaticality judgement tasks. In the models, `gender` and `region` did not play a role.

are statistically worse than the intercept, but not different from each other or the ungrammatical baseline FMFM.

The conclusion we can draw from this is that gender mismatches are possible in either direction as long as the adjectival agreement is feminine. This is shown in (17) and (18).

(17)   **Two-way mismatches possible with feminine agreement**

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

(18)   **No mismatch possible with masculine agreement**

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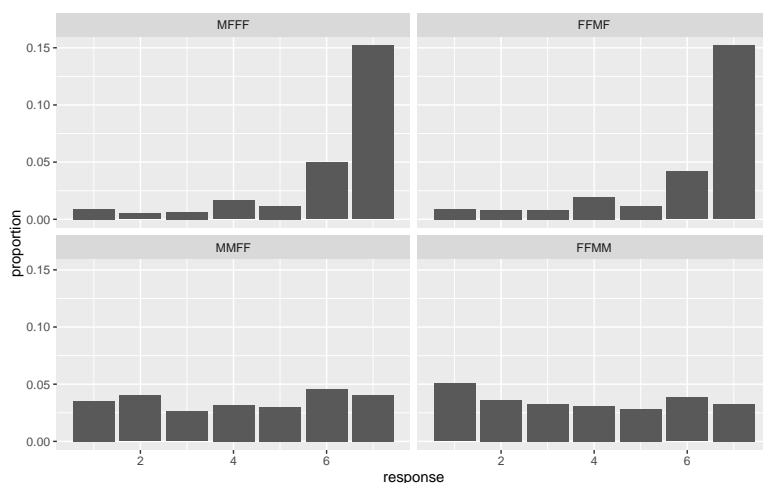'Jovan is a regular customer and Milica an occasional one.'   MMFF

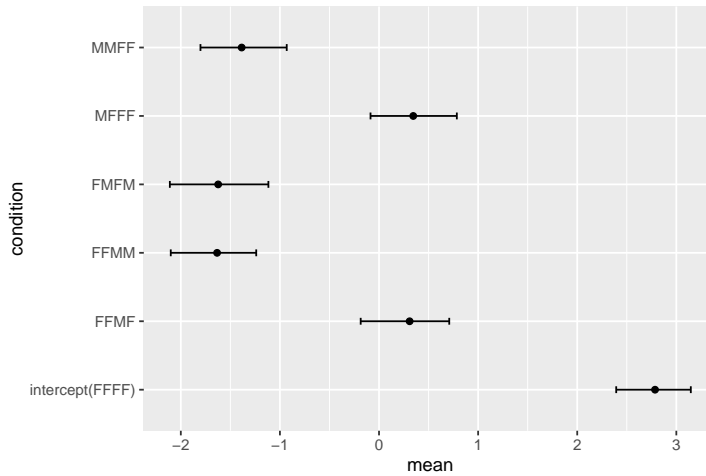Figure 4: Responses by participants who liked MMMM (median rating ≥ 4)

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

Thus, with regard to the noun classes in Table 2, our hybrid nouns behave like nouns of the two-way alternating *actor*-class when there is agreement in grammatical gender (feminine), but they behave like non-alternating nouns of the *brother*-class if at least one of them bears natural gender (masculine).

5 Analysis

First, let us assume that the identity requirement for elided material involves Merchant's (2001) e-GIVENness (i.e. mutual entailment), as defined in (19), where existential (∃)-type shifting is "a type-shifting operation that raises the expressions of type t and existentially binds unfilled arguments" while F-closure of α "is the result of replacing F-marked parts of α with ∃-bound variables of the appropriate type (modulo ∃-type shifting)" (Merchant 2001: 14):

(19) **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 condition requires that mutual entailment holds between the antecedent and the ellipsis site. This prevents semantically-equivalent, non-matching ellipsis sites (20b) Merchant (2001: 27).

(20) a. Abby *called Ben an idiot* after Mary did ⟨call Ben an idiot⟩.
         ($\exists x . x$ called Ben an idiot $\leftrightarrow \exists x . x$ called Ben an idiot)

   b. # Abby *called Ben an idiot* after Mary did ⟨insult Ben⟩.
         ($\exists x . x$ called Ben an idiot $\nleftrightarrow \exists x . x$ 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). In Greek, non-alternating nouns of the *brother*-class have been claimed to contain a presupposition about gender of the referent (21) (Merchant 2014: 19, Sudo & Spathas 2016: 715).

(21) a. ⟦adherfos⟧ $= \lambda x_e :$ MALE$(x) .$ SIBLING(x)
    b. ⟦adherfi⟧ $= \lambda x_e :$ FEMALE$(x) .$ SIBLING(x)

(22) a. * O   Petros ine kalos   adherfos, ala i   Maria ine mia kakia
           the Petros is   good.M brother   but the Maria is   a.F bad.F
           ⟨adherfi⟩.
            sister
           Intended: 'Petros is a good brother, but Maria is a bad one (sister).'

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

                                                     (Greek; Merchant 2014: 12)

The elided and antecedent noun have conflicting gender presuppositions, not mutually-entailing:

(23) $\exists x :$ MALE$(x) .$ SIBLING$(x) \leftrightarrow \exists x :$ FEMALE$(x) .$ SIBLING(x)

In the analysis of two-way alternating "epicene" nouns, however, it is often assumed that they do not contain any lexical presuppositions about gender (24), and thus ellipsis is licensed (26).

(24) ⟦jatros⟧ $= \lambda x_e .$ DOCTOR(x)

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

                                                    (Greek; Merchant 2014: 15)

(26) $\exists x \,.\, \text{DOCTOR}(x) \leftrightarrow \exists x \,.\, \text{DOCTOR}(x)$

If we adopt a similar approach for hybrid nouns in BCS, then hybrid nouns agree-
ing in grammatical gender (feminine) do not contain any presupposition about
gender, while the hybrid nouns that agree in natural gender (masculine) should
have an additional gender presupposition. Assuming that the grammatical femi-
nine gender of hybrid nouns does not contribute any gender presupposition, this
would make it compatible with male and female referents:

(27) $[\![n_{[\text{F}]}]\!] = \lambda P \lambda x \,.\, P(x)$

For masculine-agreeing hybrid nouns, let us assume that the denotation of natu-
ral gender is a partial identity function restricting the set of customers to the set
of male customers (Cooper 1983).

(28) $[\![n_{[\text{M}]}]\!] = \lambda P \lambda x : \text{MALE}(x) \,.\, P(x)$

The introduction of a gender presupposition with masculine agreeing hybrid
nouns makes them incompatible with female referents.[6]

(29) 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.'

(30) $[\![\text{novi mušterija}]\!] = \lambda x \,.\, \text{CUSTOMER}(x)$, defined only if x is male
   a. $[\![(29a)]\!] = \text{CUSTOMER}(\text{MARIJA})$

---

[6]While other theories rule out such examples based on competition with the feminine agreeing
  form (e.g. Maximize Presupposition; Bobaljik & Zocca 2011: 148f. or Principle of Gender Com-
  petition; Sudo & Spathas 2016: 722), we argue that the syntactic presence of both features is
  necessary based on instances of mixed agreement, such as (3).

b. ⟦(29b)⟧ = CUSTOMER(MARIJA), defined only if Marija is male
(presupposition failure!)

Returning to gender mismatches under ellipsis, in an example such as (31), involving feminine agreement in both clauses, both instances of *mušterija* will lack natural gender feature and the corresponding presupposition.

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

In order to analyze the ellipsis patterns, we adopt the assumptions of Distributed Morphology that nouns are built up from a category neutral root and a head *n* that categorizes this root as nominal (Halle & Marantz 1993; Harley & Noyer 1999; Kihm 2005; Acquaviva 2009; Kramer 2015). Following Kramer (2015), we will treat gender for now as a feature introduced by *n* (although see below for more detail about the different possibilities of simultaneous separate structural encoding of natural and grammatical gender on hybrid nouns). We assume further that number features of nouns are introduced on a NumP. These assumptions yield the structure for (31) as given in (32), where ellipsis is triggered by an [E] feature on Num (e.g. Merchant 2014; Saab & Lipták 2016; Saab 2019).

(32) **Antecedent**:   **Ellipsis site**:
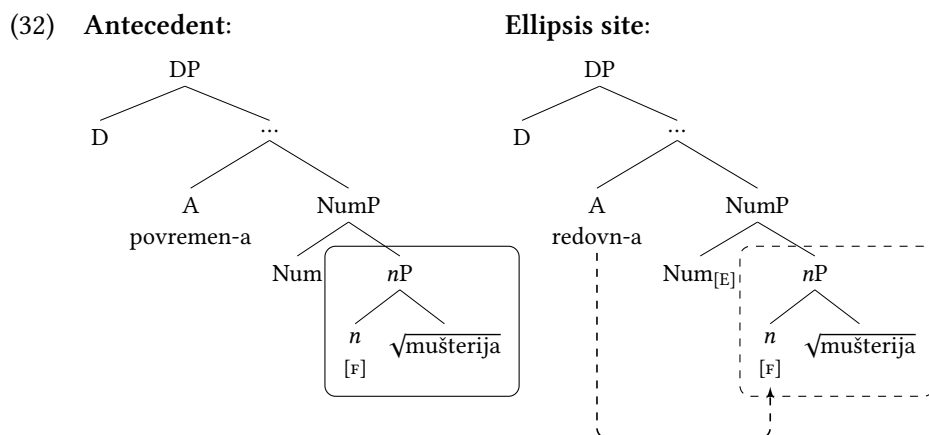
Importantly, from the point of view of the e-GIVENness condition in (19), mutual entailment is trivially satisfied since the elided material is identical.

(33)

However, as soon as we have masculine agreement on one of the conjuncts, we necessarily have the relevant natural gender feature [M] of the referent, making the ellipsis unacceptable (34). The relevant structure, demonstrating the lack of feature identity, is given in (35).

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

(35) **Antecedent**:   **Ellipsis site**:

Recall that the masculine feature also introduces the presupposition that the referent is male. If there is no corresponding feature in the antecedent, then e-GIVENness is violated because there being a customer does not entail there being a male customer:

(36)

The same situation holds if masculine agreement obtains in the antecedent:

(37) * 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

Since *mutual* entailment is required, the existence of masculine gender at only one of the conjuncts results in ungrammaticality, since the e-GIVENness requirement is not met.

(38)

$$
\left[\!\!\left[
\begin{array}{c}
n\text{P} \\
\overbrace{\underset{[\text{M}]}{n} \quad \sqrt{\text{mušterija}}}
\end{array}
\right]\!\!\right] \quad \nleftarrow \quad
\left[\!\!\left[
\begin{array}{c}
n\text{P} \\
\overbrace{\underset{[\text{F}]}{n} \quad \sqrt{\text{mušterija}}}
\end{array}
\right]\!\!\right]
$$

$\exists x : \textsc{male}(x).\ \textsc{customer}(x) \qquad\qquad \exists x.\ \textsc{customer}(x)$

This accounts for why mismatches in referent gender are not tolerated if the adjective agrees in masculine in one conjunct only. As we expect, if we have two masculine referents (39), then mutual entailment is restored (40).

(39) Uroš je redovn-i mušterija, a   Tomislav povremen-i ⟨mušterija⟩.
      Uroš is regular-M customer and Tomislav occasional-M   customer
      'Uroš is a regular customer and Tomislav an occasional one.'   MMMM

(40)

$$
\left[\!\!\left[
\begin{array}{c}
n\text{P} \\
\overbrace{\underset{[\text{M}]}{n} \quad \sqrt{\text{mušterija}}}
\end{array}
\right]\!\!\right] \quad \leftrightarrow \quad
\left[\!\!\left[
\begin{array}{c}
n\text{P} \\
\overbrace{\underset{[\text{M}]}{n} \quad \sqrt{\text{mušterija}}}
\end{array}
\right]\!\!\right]
$$

$\exists x : \textsc{male}(x).\ \textsc{customer}(x) \qquad\quad \exists x : \textsc{male}(x).\ \textsc{customer}(x)$

Exactly why masculine agreement with hybrid nouns (39) is not possible for all speakers is an open issue, and we suggest it could be due to not all speakers having the variant of *n* containing the additional [M] feature. We leave further examination of interspeaker variation to future research.
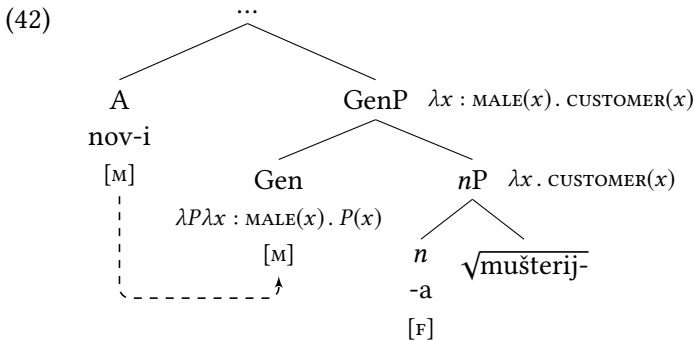
5.1 The encoding of natural and grammatical gender features

Since the nouns addressed in this paper have the possibility to simultaneously encode both the grammatical feminine and the natural masculine gender (as illustrated by (3), see also Wechsler & Zlatić 2003; Despić 2017; Puškar 2017), a question that remains open is where exactly these features are encoded in the DP structure.

Some accounts (Matushansky 2013; Pesetsky 2013; Landau 2016) assume grammatical gender to be encoded low in the noun's structure, as a property related to the nominal stem. Natural gender is optionally introduced on a higher functional projection. Under this approach, we would treat the grammatical feminine gender of our hybrid nouns as a property of *n*, while natural gender would be introduced at a higher functional projection (e.g. GenP, cf. Picallo 1991). The lower gender would not introduce any gender presuppositions, while the denotation of the gender on Gen would be a partial identity function restricting the set of customers to the set of male customers (Cooper 1983). If the male-referring hybrid noun triggers feminine agreement, adjectival concord then targets the grammatical gender feature on *n* (41).

(41)
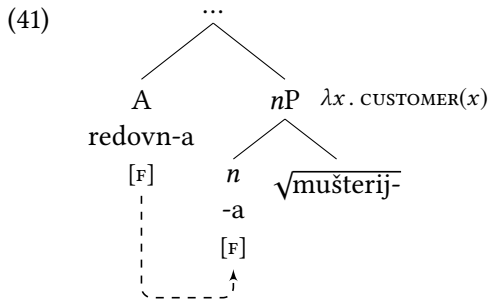
For masculine-agreeing hybrid nouns, the closest target for Agree will be the higher masculine gender (42).

(42)

$$\llbracket n\text{P} \rrbracket \quad = \quad \lambda x \,.\, \text{CUSTOMER}(x)$$

$$
\begin{aligned}
\llbracket \text{GenP} \rrbracket \quad &= \quad \llbracket \text{Gen} \rrbracket (\llbracket n\text{P} \rrbracket) \\
&= \quad [\lambda P \lambda x : \text{MALE}(x)\,.\, P(x)](\lambda x \,.\, \text{CUSTOMER}(x)) \\
&= \quad \lambda x : \text{MALE}(x)\,.\, [\lambda x' \,.\, \text{CUSTOMER}(x')](x) \\
&= \quad \lambda x : \text{MALE}(x)\,.\, \text{CUSTOMER}(x)
\end{aligned}
$$

Ellipsis would be licensed if neither the antecedent nor the elided noun project the GenP, or if both of them do. The existence of GenP and the masculine feature in either the antecedent or the ellipsis site would lead to a lack of mutual entailment and the concomitant impossibility of ellipsis, in the manner proposed above.

Another theoretical option would be to make the GenP the locus of grammatical gender, and the *n* of the natural one for the hybrid nouns of this particular type (see Puškar 2017; 2018). A hybrid noun with no natural gender would be represented as (43), while the noun with the natural gender feature would have the structure in (44). Under this approach, the natural gender feature of *n* would have a more complex syntactic representation than the grammatical one. The gender probe is relativized towards the more complex gender. If this gender is present on the *n*, it will be the preferred goal for Agree (44). If it is absent, Agree would copy the higher grammatical gender (43) (see Puškar 2017 for further detail).
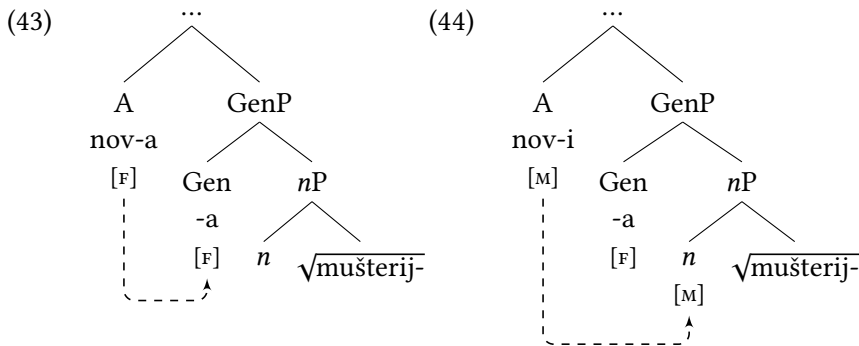
If this analysis were to be adopted, the ungrammatical sentences would be ruled out by prohibiting any mismatches between the features of antecedent and those of the ellipsis site (cf. Merchant 2013).

To sum up, both types of approaches to the structural position of gender features would in principle be compatible with the results of our experiment. Future

research should then be tasked with teasing apart the two and defining the exact locus of the natural and grammatical gender features. However, what the results undoubtedly reveal is the necessity to represent the two types of features separately, as well as that they differ in quality, such that natural gender has an additional meaning component.

6 Conclusion

This paper has presented new experimental data from NP ellipsis showing that hybrid nouns in BCS show a split behaviour with regard to gender ellipsis: if they agree in feminine gender, then mismatches in referent gender are permitted for either the antecedent or elided noun. However, masculine agreement in either conjunct blocks gender mismatches. We have linked this to the optional presence of a natural [M] gender feature which represents a target for adjectival concord and introduces a gender presupposition. It is this gender presupposition that destroys the mutual entailment relation required by Merchant's (2001) e-GIVENness requirement on ellipsis licensing. Consequently, this study suggest that natural masculine gender is syntactically absent when the adjective agrees in grammatical gender.

Abbreviations

F	feminine	REFL	reflexive
M	masculine	SG	singular
PL	plural		

Acknowledgements

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 4 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 5.

Table 4: Coefficients for the MCMC model with confidence intervals and cutpoints.

	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 5: Random effects for the MCMC model.

	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

References

Acquaviva, Paulo. 2009. Roots and lexicality in Distributed Morphology. In Alexandra Galani, Daniel Redinger & Norman Yeo (eds.), *York-Essex Morphology Meeting 2 York*, 1–21.

Alsina, Aleks & Boban Arsenijević. 2012a. The two faces of agreement. *Language* 88(2). 369–379.   DOI:10.1353/lan.2012.0039

Alsina, Aleks & Boban Arsenijević. 2012b. Hierarchies and competing generalizations in Serbo-Croatian hybrid agreement. In *17th International Lexical Functional Grammar Conference (LFG2012)*. Udayana University, Bali.

Arsenijević, Boban & Marina Gračanin-Yuksek. 2015. *Agreement and the structure of relative clauses*. Manuscript, University of Niš & Middle East Technical University.

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(2). 141–166. DOI:10.1007/s11525-010-9156-3

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

Corbett, Greville. 1991. *Gender*. Cambridge: Cambridge University Press.

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

de Vries, Mark. 2006. The syntax of appositive relativization: On specifying coordination, false free relatives, and promotion. *Linguistic Inquiry* 37(2). 229–270. DOI:10.1162/ling.2006.37.2.229

Despić, Miloje. 2017. Investigations in mixed agreement: Polite plurals, hybrid nouns and coordinate structures. *Morphology* 27(3). 253–310. DOI:10.1007/s11525-017-9301-3

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. http://www.jstatsoft.org/v33/i02/.

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The View from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT Press.

Harley, Heidi & Rolf Noyer. 1999. State-of-the-article: Distributed Morphology. *Glot International* 3. 3–9.

Heim, Irene. 2008. Features on bound pronouns. In Daniel Harbour, David Adger & Susana Bejar (eds.), *Phi Theory: Phi-features across modules and interfaces*, 35–56. Oxford: Oxford University Press.

Kihm, Alain. 2005. Noun class, gender and the lexicon-syntax-morphology interfaces: A comparative study of Niger-Congo and Romance languages. In Guglielmo Cinque & Kayne Richard S. (eds.), *The Oxford handbook of comparative syntax*, 459–512. Oxford: Oxford University Press.

Kramer, Ruth. 2015. *The Morphosyntax of gender*. Oxford: Oxford University Press.

Kratzer, Angelika. 2009. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40(2). 187–237. DOI:10.1162/ling.2009.40.2.187

Landau, Idan. 2016. DP-internal semantic agreement: A configurational analysis. *Natural Language & Linguistic Theory* 34(3). 975–1020. DOI:10.1007/s11049-015-9319-3

Matushansky, Ora. 2013. Gender confusion. In Lisa Cheng & Norbert Corver (eds.), *Diagnosing syntax*, 271–294. Oxford: Oxford University Press.

Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.

Merchant, Jason. 2013. Voice and ellipsis. *Linguistic Inquiry* 44(1). 77–108. DOI:10.1162/LING_a_00120

Merchant, Jason. 2014. Gender mismatches under nominal ellipsis. *Lingua* 151(A). 9–32. DOI:10.1016/j.lingua.2014.01.008

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

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

Picallo, M. Carme. 1991. Nominals and nominalizations in Catalan. *Probus* 3(3). 279–316. DOI:10.1515/prbs.1991.3.3.279

Puškar, Zorica. 2018. Interactions of gender and number agreement: Evidence from Bosnian/Croatian/Serbian. *Syntax* 21(3). 275–318. DOI:10.1111/synt.12154

Puškar, Zorica. 2017. *Hybrid agreement: Modelling variation, hierarchy effects and phi-feature mismatches*. University of Leipzig dissertation.

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

Saab, Andrés. 2019. Nominal ellipsis. In Jeroen van Craenenbroeck & Tanja Temmerman (eds.), *The Oxford handbook of ellipsis*, 526–561. Oxford: Oxford University Press.

Saab, Andrés & Anikó Lipták. 2016. Movement and deletion after syntax: Licensing by inflection reconsidered. *Studia Linguistica* 70(1). 66–108. DOI:10.1111/stul.12039

Sauerland, Uli. 2003. A new semantics for number. In Robert B. Young & Yuping Zhou (eds.), *Proceedings of SALT 13*, 258–275. Ithaca, NY: Cornell Linguistics Club. DOI:10.3765/salt.v13i0.2898

Sauerland, Uli. 2008. On the semantics markedness of phi-features. In David Adger, Daniel Harbour & Susana Bejar (eds.), *Phi Theory: Phi-features across modules and interfaces*, 57–82. Oxford: Oxford University Press.

Smith, Peter. 2015. *Feature mismatches: Consequences for syntax, morphology and semantics*. University of Connecticut dissertation.

Smith, Peter. 2016. *The Agreement Hierarchy and AGREE*. Manuscript, Goethe-Universität Frankfurt. http://ling.auf.net/lingbuzz/003194.

Spathas, Giorgos. 2010. *Focus on anaphora*. Utrecht University dissertation. https://www.lotpublications.nl/Documents/264_fulltext.pdf.

Sudo, Yasutada. 2012. *On the semantics of phi features on pronouns*. Cambridge, MA: MIT dissertation.

Sudo, Yasutada & Giorgos Spathas. 2016. Nominal ellipsis and the interpretation of gender in Greek. In Nadine Bade, Polina Berezovskaya & Anthea Schöller (eds.), *Proceedings of Sinn und Bedeutung 20*, 712–724. https://semanticsarchive.net/sub2015/SeparateArticles/Sudo-Spathas-SuB20.pdf.

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