The structure of hybrid nouns in Serbian: Evidence from mismatches under NP ellipsis
Andrew Murphy & Zorica Puškar (Universität Leipzig)

Background: Hybrid nouns in Serbian such as budala ‘fool’, varalica ‘cheater’, mušterija ‘customer’, propalica ‘loser, failure’, pijanica ‘drunkard’, etc. are grammatically feminine, but can trigger both masc. (natural gender) and fem. (grammatical) agreement with a masculine referent:

(1) Milan je star-a/%star-i mušterija.
‘Milan is an old customer.’

(2) Marija je star-a/%star-i mušterija.
‘Marija is an old customer.’

There are various proposals for this variability in agreement, i.e. two different lexical entries for hybrid nouns (Merchant 2014), semantic agreement (Sudo & Spathas to appear) or the interaction of syntactic operations (Puškar 2015). These approaches make different assumptions about the presence/absence of conflicting gender features on the noun. We use syntactic identity under ellipsis (Merchant 2013) as a diagnostic to choose between these hypothesis by investigating the possibility of gender mismatches under NP ellipsis.

Evidence from NP ellipsis: Recent literature on NP ellipsis (e.g. Nunes & Zocca 2009, Bobaljik & Zocca 2011, Merchant 2014) shows that gender-variable nouns allow for various types of mismatches. In Serbian, while ellipsis of a noun with grammatical (F) or natural gender (M) (indicated by the agreement on the adjective) is licensed by an antecedent noun of the same type (3a,b), it is unclear to what extent mismatches are possible (3c,d).

(3) a. Milan mu je stara mušterija, a Jovan mu je nova (mušterija).
‘Milan him is old.F customer but Jovan him is new.F (customer)’

b. %Milan mu je stari mušterija, a Jovan (mu je) novi (mušterija).
‘Milan him is old.M customer but Jovan him is new.M (customer)’

c. ?Milan mu je stari mušterija, a Jovan mu je nova (mušterija).
‘Milan him is old.M customer but Jovan him is new.F (customer)’

d. ?Milan mu je stara mušterija, a Jovan mu je novi (mušterija).
‘Milan him is old.F customer but Jovan him is new.M (customer)’

Main hypotheses: We identify three broad classes of approaches to the variability of agreement with hybrid nouns: Hypothesis A: Hybrid nouns are listed in the lexicon twice, with two different gender features on the two entries (lexicalist approach, e.g. Corbett 1991, Merchant 2014) (4).

Hypothesis B: Grammatical gender is always present on the noun and natural gender is only there when it is reflected by agreement/concord (cf. Steriopolo & Witschko 2010, Pesetsky 2014, Landau 2015). Thus, the structure of feminine-agreeing nouns contains the structure of the masculine-agreeing noun (assuming masculine is the underspecified form) (4).

Hypothesis C: Both the natural and grammatical gender are always present on hybrid nouns, and variability in agreement is achieved by some other aspect of the syntax, e.g. by the structure of the probe and syntactic operations (cf. Puškar 2015).

(4) Hypothesis A
masc: fem: NP NP
M \(\sqrt{\ }) F \(\sqrt{\ })

(5) Hypothesis B
masc: fem: NP NP
M \(\sqrt{\ }) F M \(\sqrt{\ })

(6) Hypothesis C
masc: fem: NP NP
M \(\sqrt{\ }) F M \(\sqrt{\ })

We assume that the requirement for syntactic identity under ellipsis (cf. Chung 2013, Merchant 2013) can be used to decide between these hypotheses. For example, Merchant (2013) shows
convincingly for voice mismatches with VP/TP ellipsis that conditions on ellipsis require that there be no syntactic material in the ellipsis site that is not present in the antecedent, i.e., the ellipsis site must constitute a syntactic subset of the antecedent (E ⊆ A). The three hypotheses above make different predictions with regard to the possibility of gender mismatches under NP ellipsis. **Prediction for Hypothesis A:** Mismatches impossible: *nat ↔ gram ((3c) & (3d) should be ungrammatical because M≠F). **Prediction for Hypothesis B:** One way mismatches possible: √nat → gram, √gram → nat *(3c) & (3d) because M⊂F, but *F⊂M. **Prediction for Hypothesis C:** Two-way mismatches possible: √nat ↔ gram ((3c) & (3d) both acceptable because M=F).

**Experiment:** **Aim:** Test the whether grammatical gender can license ellipsis of natural gender (5c), and *vice versa* (5d), i.e., whether there is a significant different in acceptability. **Task:** A grammaticality judgement task (7-point Lickert scale, 1 = ‘bad’, 7 = ‘good’). Each sentence contained two clauses, 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, as in the following:

(7) [NP₁ Jovan] je [Adj₁ star-a] mušterija, a [NP₂ Marko] [Adj₂ potencijaln-a] ________.

‘Jovan is a customer but Marko a potential one.’ MFMF

Factors in the design included gender of the subjects (M(asc) vs. F(em)) and agreement on adjectives in the first and second clause (M vs. F), while the conditions were the combinations of gender on the subjects (NP₁ & NP₂) and adjectives (Adj₁ and Adj₂) in each clause (e.g. MFMF = NP₁=M, Adj₁=F; NP₂=M is Adj₂=F, as in (7)). The stimuli included the combinations MFMF, MMMM, MFMM and MMMF. Participants were presented with 96 test items total: 48 test items; 4 conditions above x 12 items per condition. The subject in each condition was masculine. 48 control items; 2 conditions x 24 items per condition. These included feminine subjects, FFFF (grammatical control), FMFM (ungrammatical control). The experiment was coded using LimeSurvey and completed by 50 volunteers.

**Results and conclusions:** The crucial two combinations with gender mismatches, MFMM and MMMF, were almost identical in distribution. They were significantly better than the completely ungrammatical control (FMFM) (p < 0.001), contra Hypothesis A. There was no significant difference between the two combinations (p < 0.347). This is unexpected for Hypothesis B, which predicts a one-way mismatch due to the asymmetric structures of hybrid nouns. However, Hypothesis C predicted that mismatches should be permitted (i.e. better than the ungrammatical cases), but there should be no directional preference for the mismatch. The present results support this conclusion. These findings support the analysis in Puškar (2015) where variability in agreement is not due to different underlying structures, but rather the nature of the Agree. A follow-up study is planned that will also test the effect of mismatches in natural gender of the referent (masc. vs. fem) as well as number mismatches (singular vs. plural).