

Gender mismatches with NP ellipsis in Bosnian/Croatian/Serbian

Overview

We investigated hybrid nouns in Bosnian/Croatian/Serbian such as *budala* 'fool', *varalica* 'cheater', *mušterija* 'customer', *propalica* 'loser, failure', and *pijanica* 'drunkard'. These nouns are invariant in form, but can trigger both masc. and fem. agreement with a masculine referent:

- (1) a. Milan nam je nov-**a** mušterija. (2) a. Marija nam je nov-**a** mušterija.
Milan us is new-**F** customer Marija us is new-**F** customer
'Milan is our new customer.' 'Marija is our new customer.'
b. %Milan nam je nov-**i** mušterija. b. *Marija nam je nov-**i** mušterija.
Milan us is new-**M** customer Marija us is new-**M** customer
'Milan is our new customer.' 'Marija is our new customer.'

There are various proposals for this variability in agreement, i.e. lexically pre-specified gender (Corbett 1991), semantic agreement (Sudo & Spathas to appear) or interaction of syntactic operations (Puškar 2015). These approaches differ with regard to the presence of conflicting gender features on the noun: unrealized features are either always present or only when reflected by agreement. We adopt the assumption of a syntactic identity condition on ellipsis (e.g. Merchant 2013) to try to distinguish between competing approaches to hybrid nouns in BCS.

Gender mismatches under 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 BCS, hybrid nouns allow for agreement with either natural or grammatical gender, but it is unclear to what extent one type of gender can license ellipsis of the other (3c,d).

- (3) a. Milan mu je star-**a** mušterija, a Jovan mu je nov-**a** (mušterija).
Milan him is old-**F** customer but Jovan him is new-**F** (customer)
'Milan is his old customer, and Jovan is his new one.'
b. %Milan mu je star-**i** mušterija, a Jovan mu je nov-**i** (mušterija).
Milan him is old-**M** customer but Jovan him is new-**M** (customer)
'Milan is his old customer, and Jovan is his new one.'
c. ?Milan mu je star-**i** mušterija, a Jovan mu je nov-**a** (mušterija).
Milan us is old-**M** customer but Jovan us is new-**F** (customer)
'Milan is his old customer, and Jovan is his new one.'
d. ?Milan mu je star-**a** mušterija, a Jovan mu je nov-**i** (mušterija).
Milan him is old-**F** customer but Jovan him is new-**M** (customer)
'Milan is his old customer, and Jovan is his new one.'

A = **gram. gender**, E = **gram. gender**

A = **nat. gender**, E = **nat. gender**

A = **nat. gender**, E = **gram. gender**

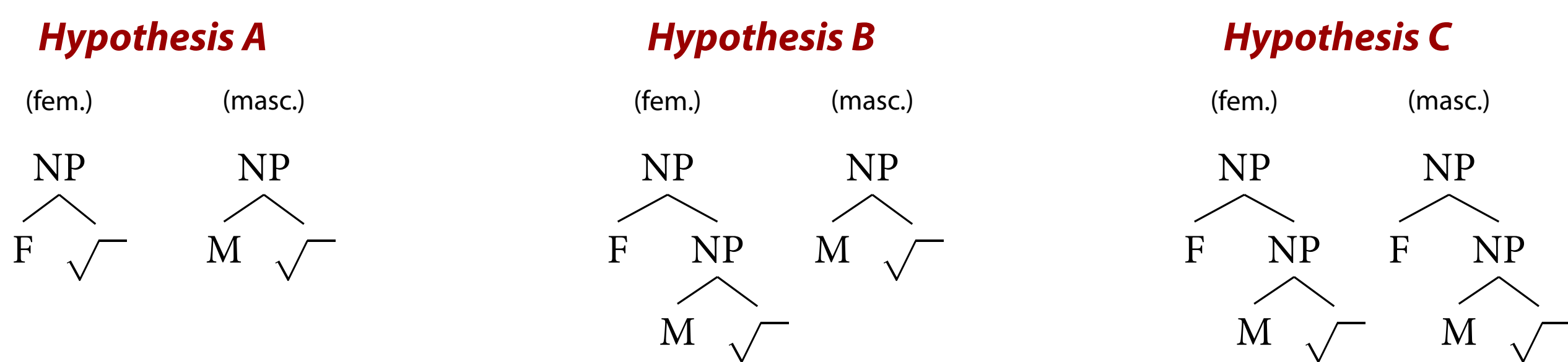
A = **gram. gender**, E = **nat. gender**

Theories of hybrid nouns

There are three general types of approach to nouns that have more than one gender feature:

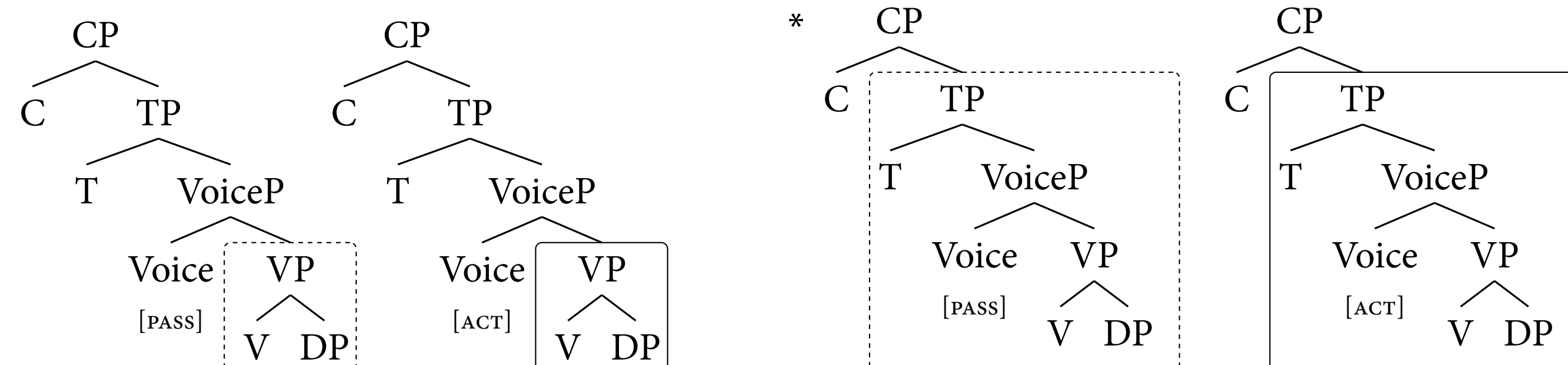
- Hypothesis A:** two distinct structures, distinct gender (M ≠ F) (cf. Corbett 1991, Merchant 2014)
Hypothesis B: two distinct structures, masculine underspecified (M ⊂ F) (cf. Pesetsky 2014, Kramer 2015)
Hypothesis C: one structure, both genders always present (M = F) (cf. Puškar 2015)

These can be represented schematically as follows:



Syntactic identity under ellipsis

Idea: Test hypotheses with a syntactic approach to ellipsis identity (cf. Merchant 2013, Murphy to appear).
Ellipsis licensing: no syntactic material in the ellipsis site not also present in the antecedent.



- Prediction for Hypothesis A:** Mismatches impossible: *nat ↔ gram (because M ≠ F).
Prediction for Hypothesis B: One way mismatches possible: *nat → gram, ✓gram → nat (because M ⊂ F).
Prediction for Hypothesis C: Two-way mismatches possible: ✓nat ↔ gram (because M = F).

Experiment

Design

- Task:** grammaticality judgement (7-point Likert scale)
- Factors:** gender of the subjects and type of agreement on adjectives in the first and second clause.
- Conditions:** combinations of gender on the subjects (NP₁ & NP₂) and adjectives (Adj₁ and Adj₂) in each clause.

CONDITION	AGREEMENT TYPE			
	NP ₁	Adj ₁	NP ₂	Adj ₂
1 MFMF	M	F	M	F
2 MMMM	M	M	M	M
3 MFMM	M	F	M	M
4 MMMF	M	M	M	F
5 FFFF	F	F	F	F
6 FMFM	F	M	F	M

- Test items:** 96 test items in total.
- 48 test items:** 4 conditions (MFMF, MMMM, MFMM, MMMF) × 12 items per condition. Only masculine subjects in both clauses.
- 48 control items:** 2 conditions (FFFF, FMFM) × 24 items per condition (good vs. bad baseline). Only feminine subjects in both clauses.

Stimuli

- (4) Jovan je star-**a** mušterija, a Marko potencijal-**a** ____ .
Jovan is old-**F** customer but Marko potential-**F**
'Jovan is an old customer and Marko a potential one.'
(5) Jovan je star-**i** mušterija, a Marko potencijal-**i** ____ .
Jovan is old-**M** customer but Marko potential-**M**
'Jovan is an old customer and Marko a potential one.'
(6) Jovan je star-**a** mušterija, a Marko potencijal-**i** ____ .
Jovan is old-**F** customer but Marko potential-**M**
'Jovan is an old customer and Marko a potential one.'
(7) Jovan je star-**i** mušterija, a Marko potencijal-**a** ____ .
Jovan is old-**M** customer but Marko potential-**F**
'Jovan is an old customer and Marko a potential one.'
(8) Slavica je tešk-**a** pričalica, a Bojana umerenij-**a** ____ .
Slavica is heavy-**F** talker but Bojana moderate-**F**
'Slavica is a big talker and Bojana is less of one.'
(9) Slavica je tešk-**i** pričalica, a Bojana umerenij-**i** ____ .
Slavica is heavy-**M** talker but Bojana moderate-**M**
'Slavica is a big talker and Bojana is less of one.'

Procedure

- The experiment was coded using LimeSurvey and run online via the LimeService platform.
- Sentences were presented one by one in a random order.
- Each participant saw all 96 sentences (*within-subject design*).
- The participant was asked to give a grammaticality judgement on a 7-point Likert scale (1= completely bad, 7=sounds excellent).



Fig. 3: Example test item

- A total of 50 participants: 12 male, 38 female (aged 15-55).
- Speakers who performed badly on the controls (*bad* (FMFM): ≥3, *good* (FFFF): ≤5) were removed for the analysis (n = 25).

Results

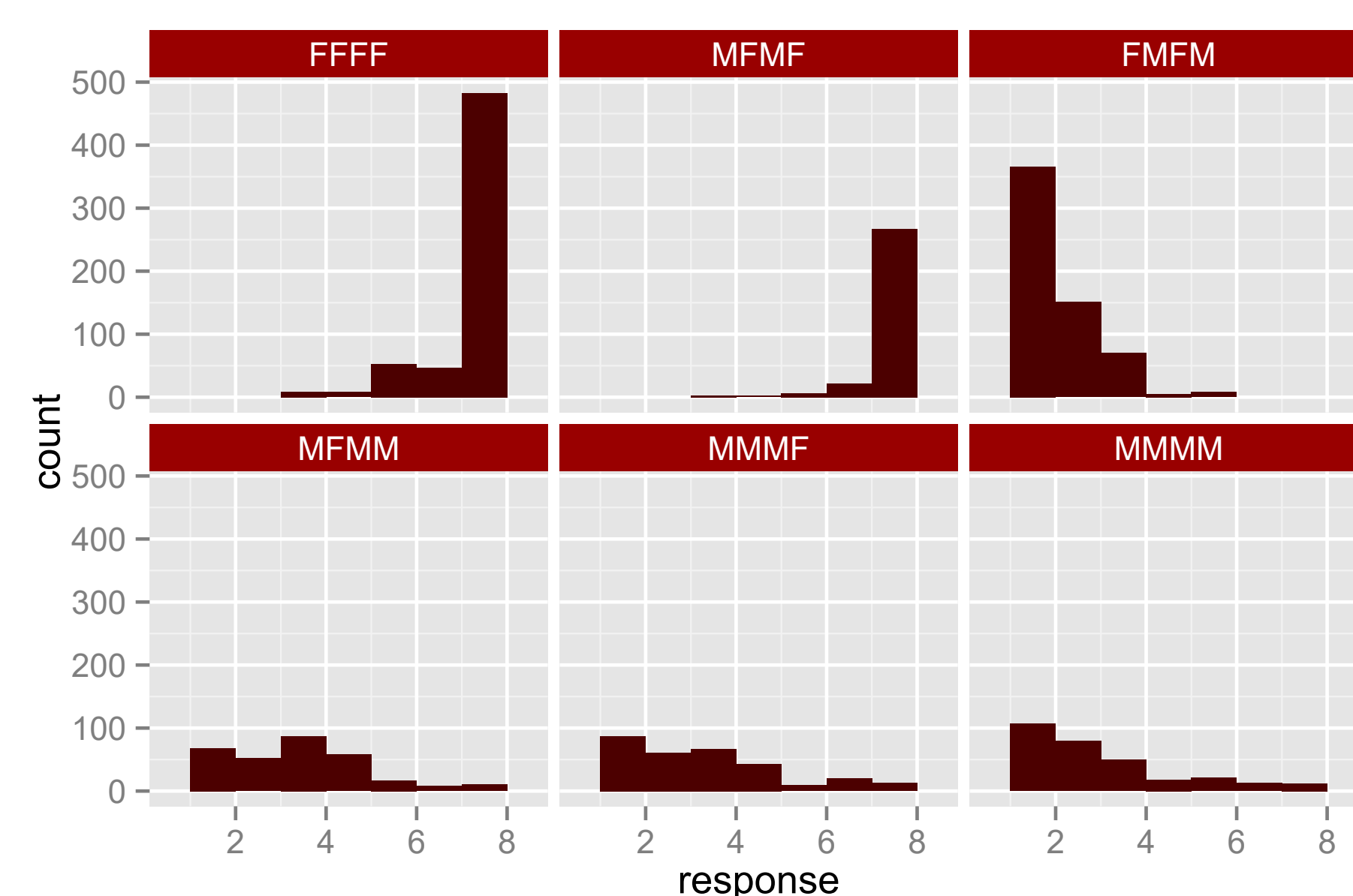


Fig. 4: Total responses per condition

	post.mean	l-95% CI	u-95% CI	pMCMC
(Intercept - MFMM)	2.89339	2.66968	3.10867	<0.001 **
combination FFFF	3.75542	3.61688	3.91681	<0.001 **
combination FMFM	-1.32946	-1.47508	-1.17964	<0.001 **
combination MFMF	3.93493	3.76130	4.11543	<0.001 **
combination MMMF	-0.09145	-0.29315	0.07554	0.347
combination MMMM	-0.37642	-0.54405	-0.19304	<0.001 **

Fig. 5: Formula response ~ combination + (1|participant) using MCMCglmm package (Hadfield 2010)

Discussion

Hypothesis A

- Prediction:** Mismatches impossible. No difference between FMFM (ungrammatical control) and test items.
- Result:** ☉ There was a significant difference between MFMM and FMFM (p < 0.001).

Hypothesis B

- Prediction:** One-way mismatches possible. Predicted difference between MFMM and MMMF (i.e. *MMMF).
- Result:** ☉ There was no significant difference between MFMM and MMMF (p = 0.347).

Hypothesis C

- Prediction:** Two-way mismatches possible. No difference between MFMM and MMMF.
- Result:** ☉ There was no significant difference between MFMM and MMMF (p = 0.347).

Conclusion

- If syntactic identity is correct, then our results are incompatible with the hypothesis that gender-variable nouns have two distinct syntactic structures (Hypotheses A & B).
- The assumption that hybrid nouns always contain both gender features can be maintained (Hypothesis C). Under this approach, variation in agreement can be attributed to other factors, e.g. relativized probing (cf. Puškar 2015).
- Possibly also compatible with semantic approaches to ellipsis identity (e.g. F-closure; Merchant 2001), depending on whether gender is encoded semantically or not.
- Future research:** Test mismatches varying the gender of the subject in the test items, e.g. MMFF vs. FFMM.

References

Bobaljik, J. D. & C. L. Zocca (2011). 'Gender Markedness: The anatomy of a counterexample', *Morphology* 21, 141-166. • Hadfield, Jarrod D (2010). 'MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package', *Journal of Statistical Software* 33. • Kramer, R. (2015). *The morphosyntax of gender*. OUP. • Merchant, J. (2001). *The Syntax of Silence*. OUP. • Merchant, J. (2013). 'Voice and ellipsis', *Linguistic Inquiry* 44 (1), 77-108. • Merchant, J. (2014). 'Gender Mismatches under Nominal Ellipsis', *Lingua* 151, 9-32. • Murphy, A. (to appear). 'Subset relations in ellipsis licensing', *Glossa*. • Nunes, J. & C. L. Zocca (2009). 'Lack of morphological identity and ellipsis resolution in Brazilian Portuguese', in *Minimalist essays on Brazilian Portuguese syntax*. Benjamins, 215-236. • Pesetsky, D. (2014). *Russian case morphology and the syntactic categories*. MIT Press. • Puškar, Z. (2015). 'Interactions of gender and number agreement: Evidence from Bosnian/Croatian/Serbian', Ms. Leipzig. • Sudo, Y. & G. Spathas (to appear). 'Nominal ellipsis and the interpretation of gender in Greek', in *Proceedings of Sub8.20*.