

Anne Mickan, Maren Schiefke and Anatol Stefanowitsch\*

# Key is a llave is a Schlüssel: A failure to replicate an experiment from Boroditsky et al. 2003

**Abstract:** In this paper, we present two attempts to replicate a widely-cited but never fully published experiment in which German and Spanish speakers were asked to associate adjectives with nouns of masculine and feminine grammatical gender (Boroditsky et al. 2003). The researchers claim that speakers associated more stereotypically female adjectives with grammatically feminine nouns and more stereotypically male adjectives with grammatically masculine nouns. We were not able to replicate the results either in a word association task or in an analogous primed lexical decision task. This suggests that the results of the original experiment were either an artifact of some non-documented aspect of the experimental procedure or a statistical fluke. The question whether speakers assign sex-based interpretations to grammatical gender categories at all cannot be answered definitively, as the results in the published literature vary considerably. However, our experiments show that if such an effect exists, it is not strong enough to be measured indirectly via the priming of adjectives by nouns.

**Keywords:** sex, grammatical gender, object conceptualization, linguistic relativity, psycholinguistics

DOI 10.1515/gcla-2014-0004

## 1 Introduction

The semantics of the grammatical gender systems of the Indo-European languages is a contentious issue. The fact that, for human nouns, there is a close-to-perfect correlation between grammatical gender and the biological and/or social

---

\*Corresponding author: **Anatol Stefanowitsch:** Freie Universität Berlin.

E-mail: anatol.stefanowitsch@fu-berlin.de

**Anne Mickan:** Radboud Universiteit Nijmegen. E-mail: a.mickan@student.ru.nl

**Maren Schiefke:** Freie Universität Berlin. E-mail: maren.schiefke@fu-berlin.de

sex<sup>1</sup> of the referent in many of these languages is transparently reflected in the terminology (“masculine” and “feminine” gender), and some influential early linguistic thinkers assumed that speakers would project the semantics of sex quite naturally onto inanimate subjects (see, for example, Grimm 1843: 193–195). However, this view did not initially gain widespread acceptance among linguists; the dominant view was that the Indo-European gender systems were simply formal noun-class systems that happened to overlap with the sex of human (and some higher animal) referents in a small portion of the lexicon but that were otherwise largely arbitrary (see Kilarski 2007 for an overview) or at least not based on biological/social sex (cf. Köpcke and Zubin 1996).

The idea that sex provides a semantic motivation for grammatical gender even for non-animate referents (referred to as the *gender-as-sex hypothesis* in the following), made a comeback in the framework of linguistic relativity, and there is a substantial body of psycholinguistic research on the topic. This body of research has yielded somewhat mixed results. The existence of a gender-as-sex effect has been shown convincingly at least for some languages in research using stimuli referring to or depicting animals (e.g. Imai et al. 2010; Saalbach et al. 2012), or a mix of objects and animals (Vigliocco et al. 2005; Flaherty 2001; Phillips and Boroditsky 2000; Boroditsky and Schmidt 2000; Boroditsky et al. 2003). Studies focusing on object labels or depictions have found an effect under some experimental conditions (roughly speaking, conditions that encourage a strategic recourse to sex when interpreting grammatical gender offline; cf. e.g. Konishi 1993; Sera et al. 1994; Koch et al. 2007), but not under others (roughly speaking, conditions that reflect online processing relatively closely; most notably, Bender et al. 2011). Where such effects are found (for animals or for objects), these are present and/or consistent in some languages (notably the Romance languages), but not in others (for example, German) (Sera 2002; Vigliocco et al. 2005; Koch et al. 2007).<sup>2</sup>

One study that is widely cited in support of the gender-as-sex hypothesis is a study by Boroditsky, Schmidt and Phillips, summarized in four paragraphs in a

---

1 We use the term *sex* to refer to the social and/or biological properties that lead to the categorization of human referents as “male” or “female”. We are aware of the fact that the term *gender* is frequently used in the literature, often to stress the social nature of this categorization, but we would like to reserve the term *gender* for the grammatical categories “masculine”, “feminine” and “neuter”.

2 This variation across languages is interesting in itself, and a number of hypotheses have been put forth to account for it; notably, that the existence and/or strength of the effect is related to the complexity of the gender system or the transparency of morphological gender marking, cf. Sera et al. 2002, Milardovic and Stefanowitsch, submitted.

comprehensive survey of literature (Boroditsky et al. 2003), but never fully published. In this study, the authors investigate the hypothesis that native speakers of Spanish and German (two languages with potentially sex-based gender systems) would associate stereotypically female properties with inanimate nouns of feminine grammatical gender and stereotypically male properties with inanimate nouns of masculine grammatical gender. Their procedure is described as follows:

[The authors] created a list of 24 object names that had opposite grammatical genders in Spanish and German (half were masculine and half feminine in each language), and then asked a group of native Spanish speakers and another group of native German speakers to write down the first three adjectives that came to mind to describe each object on the list. The study was conducted entirely in English, and none of the participants were aware of the purpose of the study. [...] After all of the adjectives provided by Spanish and German speakers were collected, a group of English speakers (unaware of the purpose of the study) rated the adjectives as describing masculine or feminine properties of the objects [...]. (Boroditsky et al. 2003: 69–70).

They claim that the results confirmed their hypothesis:

As predicted, Spanish and German speakers generated adjectives that were rated more masculine for items whose names were grammatically masculine in their native language than for items whose names were grammatically feminine. Because all object names used in this study had opposite genders in Spanish and German, Spanish and German speakers produced very different adjectives to describe the objects. For items that were grammatically masculine in Spanish but feminine in German, adjectives provided by Spanish speakers were rated more masculine than those provided by German speakers. For items that were grammatically masculine in German but feminine in Spanish, adjectives provided by German speakers were rated more masculine than those provided by Spanish speakers. (Boroditsky et al. 2003: 70).

The authors do not report quantitative results, confining themselves to a few illustrative examples. For example, they report that for the word *key*, German speakers responded with male-rated adjectives like *hard*, *heavy*, *jagged*, *metal*, *serrated*, and *useful*, corresponding to the fact that *der Schlüssel* is grammatically masculine in German, while Spanish speakers responded with female-rated adjectives like *golden*, *intricate*, *little*, *lovely*, *shiny*, and *tiny*, corresponding to the fact that *la llave* is grammatically feminine in Spanish.

In this paper, we report two attempts to replicate Boroditsky et al.'s results. Experiment 1 is a relatively direct attempt at replication, following Boroditsky et al.'s procedure as closely as possible based on the published description (except as far as the object language of the experiment is concerned); the data we report here are reanalyzed from Schiefke (2011). Experiment 2 is a more

sophisticated redesign of Boroditsky et al.'s experiment as a primed lexical decision task (cf. Mickan 2013).<sup>3</sup>

There are two motivations for our replication study: First, if shown to be replicable, Boroditsky et al.'s results would constitute very convincing evidence of a connection between grammatical gender and the semantics of sex, with interesting consequences not just for cognitive linguistics and relativity research, but also for feminist linguistics (and sociolinguistics more generally). Their method would also be extremely useful for investigating the semantics of gender systems in other languages. Second, despite being essentially unpublished (except for the sketchy description in Boroditsky et al. 2003), the study has recently been widely cited as evidence for linguistic relativity in research articles (e.g. Percy et al. 2009; Fenko et al. 2010; Chen and Su 2011; Cubelli et al. 2011; Nicoladis and Foursha-Stevenson 2012; Vuksanović et al. in press) as well as encyclopedic overviews (e.g. Boroditsky 2006; Reines and Prinz 2009). It is unfortunate that such a widely-cited study has remained unpublished and we hoped that a successful replication would put the conclusions drawn from Boroditsky et al.'s study on a firmer empirical basis.<sup>4</sup>

## 2 Experiment 1

### 2.1 Aims and methods

The first experiment was an association task based closely on Boroditsky et al.'s experiment to the extent that it could be reconstructed from their description. Unlike in that experiment, however, the subjects were tested in their respective native language rather than English, mainly because native speakers of German and Spanish with an appropriately high command of English were not available in sufficient numbers.

Since Boroditsky et al.'s list of stimuli was not available, a completely new list of ten concrete nouns was created, five with masculine gender in German

---

<sup>3</sup> Schiefke's thesis was supervised by Thomas Berg of the University of Hamburg, whose support in designing her experiment is gratefully acknowledged; here, the corresponding parts of our paper reanalyze the data from the thesis but are not otherwise based on it. Mickan's thesis was supervised by the third author of the current paper; the corresponding parts of our paper are based directly on the thesis. Both theses are available from the respective authors on request.

<sup>4</sup> In fact, we believe that replications are important even where studies are published, especially if their results go against the mainstream opinion in the field; we hope that the publication of our results encourages other researchers in cognitive linguistics to publish replications.

and feminine gender in Spanish (*luna/Mond* ‘moon’, *calabaza/Kürbis* ‘pumpkin’, *manzana/Apfel* ‘apple’, *flecha/Pfeil* ‘arrow’, *ballena/Wal* ‘whale’), and five with feminine gender in German and masculine gender in Spanish (*reloj/Uhr* ‘clock’, *violín/Geige* ‘violin’, *tenedor/Gabel* ‘fork’, *ratón/Maus* ‘mouse’, *sol/Sonne* ‘sun’). The German and Spanish lists were then presented in the form of a questionnaire to 15 native speakers each of the respective languages. They were instructed to write down the first three adjectives that came to mind when thinking about the object referred to by each word.

This resulted in a list of 195 different Spanish adjectives and 143 different German adjectives. These were presented to a new set of native-speaker subjects (ten Spanish and ten German speakers respectively), who were unaware of the experiment. These subjects were asked to categorize the adjectives as “female” or “male” if possible, with the option of judging them as “neutral” if they could not decide. From their responses, we calculated a male/female score for each adjective by subtracting the number of subjects who judged it female from the number of subjects who judged it male. In order to avoid having to deal with negative numbers, we added 10 to the result, giving us scores ranging from 0 (female judgments only) to 20 (male judgments only). We then coded the responses from the association task according to these scores and derived two different versions of the dependent variable: first, and straightforwardly, the mean male/female score of all three adjectival responses (referred to as *total response value* below); second, since it is possible that the first response in the association task reflects subjects’ conceptualizations of nouns more closely than the second and third, the male/female score of the first adjectival response (referred to as *primary response value* below).

## 2.2 Results

Table 1 shows the means of the total response values by German and Spanish subjects to nouns with masculine and feminine gender in their respective language.

**Table 1:** Means of the total response values

		Gender	
		Feminine	Masculine
Language	German	8.89	9.67
	Spanish	9.88	10.20

The responses point in the expected direction; responses to grammatically masculine nouns on average have slightly more male ratings than responses to grammatically feminine nouns in both languages. However, the difference is very small and neither a by-subject nor a by-item repeated-measures ANOVA revealed a significant effect for gender (by-subject:  $F(1,28) = 3.194$ ,  $p = .0847$ ; by-stimulus:  $F(1,8) = 0.611$ ,  $p = .457$ ) or for the interaction between language and gender (by-subject:  $F(1,28) = .562$ ,  $p = .4598$ ; by-stimulus:  $F(1,8) = .025$ ,  $p = .879$ ).

Table 2 two shows the means of the primary response values.

**Table 2:** Means of the primary response values

		Gender	
		Feminine	Masculine
Language	German	9.07	9.65
	Spanish	10.48	9.76

Here, the differences are even smaller, and they point in the wrong direction in the case of Spanish. Again, a repeated-measures ANOVA revealed no significant effects for gender (by-subject:  $F(1,28) = .013$ ,  $p = .912$ ; by-stimulus:  $F(1,8) = .012$ ,  $p = .915$ ) or for the interaction between language and gender (by-subject:  $F(1,28) = 1.206$ ,  $p = .281$ ; by-stimulus:  $F(1,8) = .133$ ,  $p = .725$ ).

## 2.3 Discussion

The differences in the male/female scores of adjectival responses to grammatically masculine and feminine nouns are very small and statistically non-significant, both for the mean scores of all three adjectival responses (*total response values*) and for the scores of the first adjectival response (*primary response value*). Thus, there is no evidence for a spillover of the human semantics of grammatical gender to non-human/inanimate entities (including animals).

This could be due to the experimental design. An association task may be unable to uncover differences in the semantics of masculine and feminine nouns for at least two reasons: First, the adjectival responses by the subjects tend to be descriptive of the object named and this may override an existing tendency to associate nouns with gender-stereotypical adjectives. For example, responses to ‘clock’ (*Uhr/reloj*) in both languages tended to be adjectives like *exact/precise*, *temporal*, *punctual*, *important*, *useful* etc. Second, the very open nature of the design leads to a large, and perhaps noisy data set (recall that there

were almost 200 Spanish and almost 150 German different adjectives among the responses).

The question remains why Boroditsky et al. (2003) found a significant effect. One reviewer suggested that this might be due to the fact that they conducted the experiment in English rather than in the respective native languages of the subjects and that the use of a foreign language encourages stereotypical thought. This is an intriguing possibility worthy of further study. We are skeptical, however, first, because other experiments conducted in subjects' native languages but using different methods did find significant effects (e.g. Koch et al. 2007); second, because there is research suggesting that English-French and English-Spanish bilinguals resort to sex-stereotypical thinking *less* frequently and/or strongly when using gender-less English than when using the systematically-gendered Romance languages (Wasserman and Wesley 2009).

In order to exclude all potential problems with the design of our association study, we decided to construct an analogous experiment using a more constrained design.

## 3 Experiment 2

### 3.1 Aims and methods

We decided on a primed lexical decision task with pictures of objects as primes, and adjectives (both words and non-words) as targets. If the grammatical gender of words for inanimate objects is meaningful, subjects should respond faster to a target if they are primed with an object whose label has a grammatical gender that is congruent with the stereotypical connotation of the adjective than if they are primed with an object with a label whose gender is incongruent.

We created a list of primes, consisting of 16 familiar objects that are easy to represent as a line drawing and that have an obvious linguistic label. Eight of these labels are masculine in German and feminine in Spanish (*chair, key, spoon, table, broom, star, moon, arrow*), and eight are feminine in German and masculine in Spanish (*bridge, clock, fork, bench, toothbrush, socks, sun, paper clip, arrow*). In addition, a list of 32 distractor items was created, containing objects whose labels either have the same gender in both languages or which have neutral gender in German. Line drawings of all primes and distractor items were collected from publicly accessible databases.

Next, we created a list of targets, consisting of eight target adjectives with stereotypically male connotation (*rough, hard, aggressive, powerful, strong,*

*heavy, dangerous, large*) and eight with stereotypically female connotations (*beautiful, soft, lovely, fragile, sweet, elegant, tiny, light*). The selection procedure involved two steps: first, we chose 44 adjectives that are clearly male or female in terms of their semantic differential (Osgood et al. 1957; Heise 1965). These were then presented to 91 native speakers of English, Spanish or German who were asked to rate them on a 5-point scale from “very female” to “very male”. The 16 adjectives mentioned above were the ones most clearly rated “male” or “female” by these speakers. In addition, we created a list of distractor items consisting of eight neutral adjectives and 24 non-words.

The primes and targets were combined into 16 critical prime-target pairs, eight congruent ones (match between the grammatical gender of the prime and the stereotypical connotation of the target) and eight incongruent ones (mismatch between gender and connotation). We took care to avoid descriptive connections (e.g. the picture of an apple followed by the adjective *sweet*). The distractors were combined into 32 filler sequences. The critical items and the distractors were combined into a pseudo-random overall sequence in which no two critical prime-target pairs followed each other immediately, and where word and non-word targets did not occur in a predictable order.

The experiment was run on a MacBook Pro (Intel Core i5, 2,3 GHz, Mac OS X 10.6.8) using SuperLab 4.5.4 to present stimuli and record answers and reaction times. Participants were tested individually in our lab room with no visual distractions in their field of vision.

All German and most Spanish participants were students at the Freie Universität Berlin, 5 of the Spanish participants were full-time employees living in Berlin. After the data collection was complete, we had to exclude 4 participants (3 Spanish, 1 German) because they turned out to be native bilinguals with second languages that could potentially interfere with the congruency of the prime-target pairs. One participant was excluded because they answered correctly only twice in the 16 critical conditions (rejecting 14 English adjectives as non-words).

## 3.2 Results

Table 3 shows the mean response times of the German and Spanish subjects in the congruent and incongruent conditions.

Clearly, the effect goes against the prediction of the gender-as-sex hypothesis: the response times in the congruent conditions are actually longer, on average, than those in the incongruent conditions. However, a repeated-measures ANOVA revealed no significant effect of gender congruity (by subject:  $F(1,54) = 0, p = .98$ ;

**Table 3:** Mean response times

		Gender	
		Congruent	Incongruent
Language	German	667.6 ms	652.4 ms
	Spanish	677.6 ms	670.5 ms

by-item:  $F(1,14) = 1.76, p = .21$ ) or the interaction of gender congruity and language (by-subject:  $F(1,54) = 2.06, p = .16$ ; by item:  $F(1,14) = .49, p = .49$ ).<sup>5</sup>

### 3.3 Discussion

Again, the effects predicted by the gender-as-sex hypothesis failed to materialize, so the null hypothesis cannot be rejected. Given the careful selection of primes and targets, this seems unlikely to be due to confounding factors in the experimental design (although of course this possibility cannot be excluded). Also, it is unlikely to be due to insufficiently exact measurements, as the difference in reaction times between words and non-words is highly significant (subjects responded to words much faster than to non-words,  $F(1,54) = 37.83, p < .001$ ).

It seems more likely to us that the failure to elicit significant differences in response times for sex/gender congruent and incongruent conditions reflects the absence of a gender-to-sex priming effect. Incidentally, since the experiment was conducted in English, the null result constitutes another argument against the suggestion that the null result in Experiment 1 was due to the fact that it was conducted in the subjects' native languages.

## 4 General discussion

Our failure to replicate Boroditsky et al.'s findings using two different experimental designs strongly suggests that their results were either a statistical fluke or that they were due to some unreported aspect of their procedure (for example, the specific stimuli they used or the specific male/female ratings their subjects assigned to the adjectives).

<sup>5</sup> See Mickan (2013) for a more detailed quantitative analysis of the results using a general linear mixed effects model (the absence of an effect remains the same).

This does not mean that the gender-as-sex hypothesis is fundamentally wrong, i.e. that the grammatical gender of inanimate objects will never influence speaker's conceptualizations. After all, there are experiments that clearly show such an effect at least for some languages (including Spanish) and under certain conditions.

Looking at the body of previous research, one thing that is striking is that the experiments reported by Bender et al. (2011), which found no gender-as-sex effect whatsoever, are classic priming experiments designed specifically to prevent a strategic transfer of sex semantics onto inanimate objects, while the studies that do report positive results tend to use designs that encourage such a transfer (for example asking subjects explicitly to assign sex to inanimate objects). From this perspective, it is perhaps not surprising that our primed lexical decision variant of Boroditsky et al.'s experiment yielded a null result.

It is a little more surprising that the direct replication of Boroditsky et al.'s experiment in the form of an association task did not show a gender-as-sex effect. After all, an association task (especially a written one that does not impose time constraints on the subjects), offers subjects the chance to use sex-based stereotypes strategically. However, if there is such a strategic gender-as-sex effect at all, it seems to be either too weak or unstable to be measured indirectly by the connotation of adjectives as “male” and “female”. After all, such connotations are only one aspect of the meaning of adjectives, and perhaps not one that is particularly relevant compared to their denotations in the context of an association task.

## References

- Bender, Andrea, Sieghard Beller & Karl Christoph Klauer. 2011. Grammatical gender in German: A case for linguistic relativity? *The Quarterly Journal of Experimental Psychology* 64(9). 1821–35.
- Boroditsky, Lera & Lauren Schmidt. 2000. Sex, syntax and semantics. In Lila R. Gleitman & Arun K. Joshi (eds.), *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*, 42–47. Mahwah, NJ: Lawrence Erlbaum.
- Boroditsky, Lera, Lauren Schmidt & Webb Phillips. 2003. Sex, syntax, and semantics. In Dedre Gentner & Susan Goldin-Meadow (eds.), *Language in mind: Advances in the study of language and thought*, 61–79. Cambridge, MA: MIT Press.
- Boroditsky, Lera. 2006. Linguistic relativity. In Lynn Nadel (ed.), *Encyclopedia of cognitive science*, 917–21. Chichester: John Wiley & Sons.
- Chen, Jenn-Yeu & Jui-Ju Su. 2011. Differential sensitivity to the gender of a person by English and Chinese speakers. *Journal of Psycholinguistic Research* 40(3). 195–203.
- Cubelli, Roberto, Daniela Paolieri, Lorella Lotto & Remo Job. 2011. The effect of grammatical gender on object categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 37(2). 449–460.

- Fenko, Anna, Jacco J. Otten & Hendrik N.J. Schifferstein. 2010. Describing product experience in different languages: The role of sensory modalities. *Journal of Pragmatics* 42(12). 3314–3327.
- Flaherty, Mary. 2001. How a language gender system creeps into perception. *Journal of Cross-Cultural Psychology* 32(1). 18–31.
- Grimm, Jacob. 1843. *Grammatik der hochdeutschen Sprache unserer Zeit* [Grammar of the standard German language of our time]. Constance: Belle-Vue.
- Imai, Mutsumi, Lennart Schalk, Henrich Saalbach & Hiroyuki Okada. 2010. Influence of grammatical gender on deductive reasoning about sex-specific properties of animals. In: Stellan Ohlsson and Richard Catrambone (eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, 1160–1165. Austin, TX: Cognitive Science Society.
- Kilarski, Marcin. 2007. On grammatical gender as an arbitrary and redundant category. In Douglas A. Kibbee (ed.), *History of linguistics 2005: Selected papers from the Tenth International Conference on History of the Language Sciences*, 24–36. Amsterdam & Philadelphia: John Benjamins.
- Koch, Sabine C., Friederike Zimmermann, and Rocio Garcia-Retamero. 2007. El Sol – Die Sonne. Hat das grammatische Geschlecht von Objekten Implikationen für deren semantischen Gehalt? [Does the grammatical gender of objects have implications for their semantic content?]. *Psychologische Rundschau* 58(3). 171–182.
- Konishi, Toshi. 1993. The semantics of grammatical gender: a cross-cultural study. *Journal of Psycholinguistic Research* 22(5). 519–534.
- Köpcke, Klaus-Michael, and David Zubin. 1996. Prinzipien für die Genuszuweisung im Deutschen [Principles of gender assignment in German]. In Ewald Lang & Gisela Zifoun (eds.), *Deutsch – typologisch* [German – typological], 473–491. Berlin & New York: De Gruyter.
- Mickan, Anne. 2013. *The influence of L1 grammatical gender on object conceptualization in English as an L2. A comparison of German and Spanish native speakers*. Berlin: Freie Universität Berlin bachelor thesis.
- Milardovic, Kristina & Anatol Stefanowitsch (submitted). Why *el sol* is a man but *die Sonne* is just a star. Grammar, morphology and the gender-as-sex effect.
- Nicoladis, Elena & Cassandra Foursha-Stevenson. 2012. Language and Culture Effects on Gender Classification of Objects. *Journal of Cross-Cultural Psychology* 43(7). 1095–1109.
- Percy, Elise J., Steven J. Sherman, Leonel Garcia-Marques, André Mata & Teresa Garcia-Marques. 2009. Cognition and native-language grammar: The organizational role of adjective-noun word order in information representation. *Psychonomic Bulletin & Review* 16(6). 1037–1042.
- Phillips, Webb & Lera Boroditsky. 2003. Can quirks of grammar affect the way you think? Grammatical gender and object concepts. In Richard Alterman & David Kirsh (eds.), *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*, 928–933. Mahwah, NJ: Lawrence Erlbaum.
- Reines, Maria Francisca & Jesse Prinz. 2009. Reviving Whorf: The return of linguistic relativity. *Philosophy Compass* 4(6). 1022–1032.
- Saalbach, Henrik, Mutsumi Imai & Lennart Schalk. 2012. Grammatical gender and inferences about biological properties in German-speaking children. *Cognitive Science* 36(7). 1251–1267.

- Schiefke, Maren. 2011. *The influence of grammatical gender on the connotations of nouns. A cross-linguistic comparison*. Hamburg: Universität Hamburg master thesis.
- Sera, Maria, Christian A.H. Berge & Javier del Castillo Pintado. 1994. Grammatical and conceptual forces in the attribution of gender by English and Spanish speakers. *Cognitive Development* 9. 261–292.
- Sera, Maria, Chryle Elieff, James Forbes, Melissa Burch, Wanda Rodríguez & Diane Poulin-Dubois. 2002. When language affects cognition and when it does not: An analysis of grammatical gender and classification. *Journal of Experimental Psychology: General* 131(3). 377–397.
- Vigliocco, Gabriella, David P. Vinson, Federica Paganelli & Katharina Dworzynski. 2005. Grammatical gender effects on cognition: Implications for language learning and language use. *Journal of Experimental Psychology: General* 134. 501–520.
- Vuksanović, Jasmina, Jovana Bjekić, and Natalija Radivojević. in press. Grammatical gender and mental representation of object: The case of musical instruments. *Journal of Psycholinguistic Research*.