

Differences between the features: an experimental study on Russian

We present a series of experiments on Russian revealing curious differences between gender, number and case in processing. Previous studies on other languages compared gender and number agreement mismatches in processing and showed controversial results. Some authors reported no significant differences (e.g. Lukatela et al. 1987; Cole, Segui 1994; Bañón et al. 2012). In other studies, agreement mismatches in gender caused longer RTs (Faussart et al. 1999), larger P600 effects (Barber, Carreiras 2005) and P300 latencies (Barber, Carreiras 2003), suggesting a higher cost of repair for gender mismatches. Still others observed larger N400 and P600 effects for number mismatches (Popov, Bastiaanse 2018; Muralikrishnan, Idrissi 2021), pointing to a higher saliency and repair cost of number mismatches. Case has never been compared to gender and number in real-time processing. The above-mentioned studies used diverse stimuli (article-noun pairs, sentences with adjective-noun or subject-verb agreement) and diverse tasks (lexical decision task, rapid serial visual presentation with acceptability or grammaticality judgments). We decided to compare gender and number agreement in Russian in different constructions (attributive and predicative, with SV... and VS... word orders) and using different methods, and to add case to the picture.

We conducted five experiments using both behavioral methods and ERPs. Participants of all studies were adult native Russian speakers. In Exp. 1 (N=88), we compared gender (GEN), number (NUM) and case (CASE) mismatches in participle-noun agreement (attributive, ATTR). Participants read stimulus sentences in four conditions (see (1)) word-by-word in the self-paced reading (SPRT) mode. In the second SPRT study, Exp. 2 (N=68), we compared GEN and NUM in subject-verb agreement (predicative, PRED). Participants were presented with stimulus sentences as in (2). In Exp. 3 (N=150), we employed a speeded grammaticality judgment task (SGJT) with both ATTR and PRED stimuli sets from Exp. 1-2. Exp. 4 (N=49) was an ERP study in which we compared GEN, NUM and CASE in ATTR and GEN and NUM in PRED. Participants read sentences like (1) and (2) below in the rapid serial visual presentation mode and made grammaticality judgments for 25% of them while continuous EEG was recorded. In Exp. 5 (N=121), we compared gender and number in predicative agreement in constructions with SV... (affirmative sentences) and VS... (questions) word order, as in (3a-b), using the SPRT method. Thus, we could test the role of the target/controller order.

Table 1. Materials of Exp.1-5.

| Exp. 1 | Exp. 2 | Exp. 3 | Exp. 4 | Exp. 5 |
|-------------------------------------|-------------------------------------|--|---|--|
| 64 target (ATTR), 176 fillers | 48 target (PRED), 112 fillers | 112 target (ATTR+PRED), 48 fillers | 320 target (160 ATTR, 160 PRED), 80 fillers | 30 target (PRED, 15 affirmative, 15 questions), 90 fillers |

- (1) Conditions in Exp. 1, 3-4 (ATTR: correct / GEN / NUM / CASE; controller – target)
*Učenyj natknulsja na stat'ju, **soderžavš-uju/*-ij/*-ie/*-ej** interesnye fakty.*
scientist stumbled upon paper_{F.ACC.SG} **containing**_{F.ACC.SG/*M.ACC.SG/*ACC.PL/*F.LOC.SG} interesting facts
- (2) Conditions in Exp. 2-4 (PRED: correct / NUM / GEN; controller – target)
*Iva **stojal-a/*-ø/*-i** u pruda, ukrašavšego tenistyj park.*
willow_{F.SG} **stood**_{F.SG/*M.SG/*PL} near pond adorning shadowy park
- (3) Conditions in Exp. 5 (PRED: correct / NUM / GEN; controller – target vs. target – controller)
 - a. *Učebnik **vypal-ø/*-a/*-i** iz ruk mal'čika.*
textbook_{M.NOM.SG} **fell**_{PAST.M.SG/*PAST.F.SG/*PAST.PL} from hands (of) boy
 - b. ***Vypal-ø/*-a/*-i** li učebnik iz ruk mal'čika?*
fell_{PAST.M.SG/*PAST.F.SG/*PAST.PL} QPRT textbook_{M.NOM.SG} from hands (of) boy

RTs in Exp. 1, 2, 5 were analyzed using mixed-effects linear regressions and Tukey tests for multiple comparisons. For response accuracy analysis in Exp. 3, we used mixed-effects logistic regressions and Tukey tests. In Exp. 4, we analyzed mean ERP amplitudes for target words (participles in ATTR and verbs in PRED) averaged across 9 regions of interest (ROIs) in

selected time windows using mixed-effects linear regressions and pairwise contrasts. We will not be able to present EEG data in detail for space reasons and will focus on behavioral results.

CASE effects were the least salient in processing. The P600 amplitude was smaller for CASE than for NUM and GEN. RTs in CASE stimuli did not significantly differ from those in correct stimuli (CORR) (Fig. 1). The SGJT accuracy for CASE was significantly lower than for NUM and GEN (Fig. 1).

Figure 1. Average word-by-word reading times (in ms) for Exp. 1, 2 and 5.

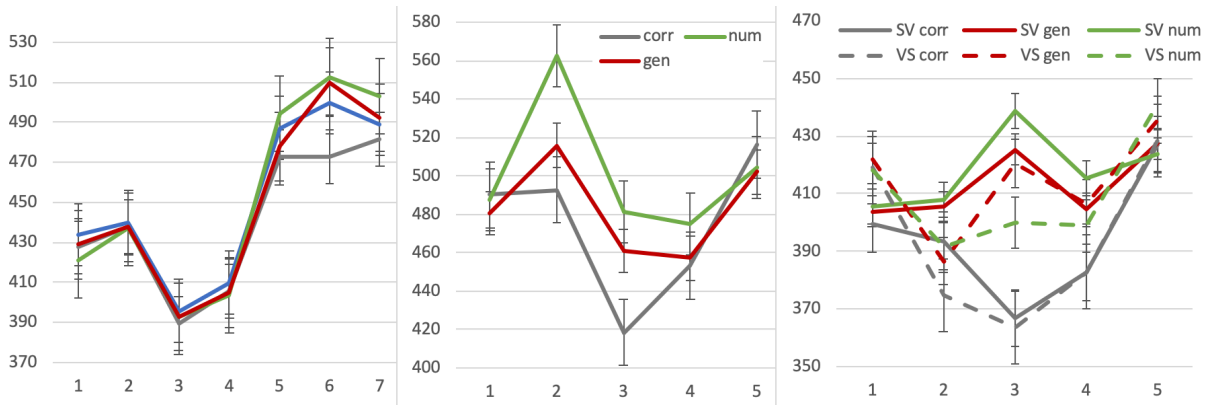


Table 2. Average response accuracy in Exp. 3.

| | Corr | Gen | Num | Case | PRED | Corr | Gen | Num |
|------|-----------|-----------|-----------|-----------|------|-----------|-----------|-----------|
| ATTR | .94 (.08) | .92 (.13) | .91 (.15) | .81 (.25) | | .98 (.06) | .91 (.14) | .90 (.14) |

Some differences between GEN and NUM were also observed. In ATTR, NUM effects on ERPs were the most salient. NUM elicited an N400, which was not observed for other mismatch types. In PRED, both GEN and NUM elicited a LAN/N400 and a P600. For GEN, LAN/N400 effects occurred earlier and had a broader distribution. NUM effects on RTs occurred earlier (on the verb) and persisted longer (up to the 2nd word after the verb) than GEN effects (observed only on the word after the verb). In Exp. 5, in affirmative sentences number errors caused larger delays than gender errors (i.e. NUM mismatches were more salient, as in Exp. 2). However, in questions the pattern was the opposite.

The observed differences between the features can be explained by their different nature. Gender is a property of the lexeme, while number and case are properties of the word form. Case depends on the syntactic context, i.e., it is defined ‘from outside’, unlike number and gender, which makes it the least salient. Case is not directly associated with semantics, gender is not semantically loaded in most nouns, while number is. In our experiments, number and gender mismatches were costlier than case mismatches. In attributive agreement, number was more salient in ERPs, as in (Popov, Bastiaanse 2018; Muralikrishnan, Idrissi, 2021). In predicative agreement, gender was more salient in ERPs, as in (Barber, Carreiras 2003, 2005), while number was more salient in RTs. We hypothesize that predicative agreement processing is more basic and immediate, so it allows to observe the earliest effect of gender mismatches on ERPs (gender is an inherent property of the noun, so the mismatch is noticed faster). But number mismatches become more salient later in the course of processing because number is semantically loaded. In questions (VS word order, the target before the controller) the readers probably can detect a gender error while they are still processing the stem, while number errors become evident only on the inflection. The influence of the word order is a possible explanation of the controversial results in previous studies, since they used different orders of the agreement controller and agreement target. We aim to investigate this further in attributive agreement.