Arseny Anisimov

fclm10@yandex.ru

Daria Sidorkina

sidorckina.darya@yandex.ru

Between methods and theories – Experimental evidence on semantics of Russian number

In this study we present two parallel experiments comparing the semantic properties of grammatical number and contexts of scalar implicature, such as quantifiers and disjunction, in Russian using grammaticality and veridicality judgements.

▲ Introduction

Among the formal theories of number one of the most prominent is the implicature theory (Spector 2007; Ivlieva 2013). According to this theory, singular number (SG) encodes strict singularity and plural (PL) is undefined, while the strict plurality meaning is derived through the pragmatic competition between the stronger SG form and the weaker PL form. Roughly, if SG is not used, then the entity must be non-singular.

Other theories include the ambiguity theory (Farkas & de Swart 2010) and the homogeneity approach (Križ 2017), all of which make analogous predictions. A plausible way to distinguish between these approaches is to test the behavior of number experimentally. (Tieu et al. 2020) use the assumption that preschool-aged children apparently compute fewer implicatures than adults to compare children's and adults' computation of strict plurality. They observe that children are less likely to derive the multiplicity inference and conclude that acquisition patterns allow to conclude that the strict multiplicity meaning of PL is an implicature.

The aim of our research was to test the nature of multiplicity inferences experimentally on the basis of Russian. We avoided appealing to acquisition patterns, since there is no guarantee that implicatures are the only semantic mechanism to be acquired relatively late. In principle, the semantic mechanisms implemented in number theories could be acquired later as well.

▲ Grammaticality judgements

Using experimental evidence on English, (Cummins & Katsos 2010; Katsos & Bishop 2011) observe that, when grammaticality is being rated on a Likert scale, violations of an implicature are judged to be more grammatical than logical contradictions. For example, with quantifiers a logical contradiction is to say "all apples" describing the situation where only some of the apples are involved. A violation of an implicature, then, is to say "some apples" in a situation where all of the apples are involved. For Russian, the quantifiers are *neskol'ko* 'some' and *vse* 'all'. Russian *neskol'ko* is somehow similar to *several*, since it states the existence of more than one entity. Therefore, we used *one apple* as a logical contradiction to *several apples.* The types of violation for quantifiers and grammatical number are shown in Table 1.

Table 1. Violation types

violation type	*uttered*	*"reality"*
logical contradiction	some X	one X
implicature violation	some X	all X
violation of SG	X-SG	X-PL
violation of PL (of strict plurality)	X-PL	X-SG

In our first experiment we sought to prove a similar contrast for Russian in upward-entailing environments and then compare the contrast between logical contradiction and implicature violation with the contrast between violations of PL and violations of SG. We used quantifiers, disjunction and past tense to test the initial assumption. We also moved from the discrete Likert scale to a toggle-bar from 'Very bad' (1) to 'Very good' (100) to simplify the calculations. If the initial assumption proved to be true, we would expect the

violation of multiplicity with SG, as the multiplicity inference is considered to be an implicature within the implicature approach.

An example of a stimulus with SG-violation is shown in (1).

(1) *Ya ostavil   v saraje lopat-u   a   tochnee   tri   lopaty*
    I left   in shed **spade-ACC.SG** and more.precisely three **spade-ACC.PL**
    'I've left a spade in the shed. Three spades, to be precise'

▲ Veridicality judgements

In the second experiment, we took a step back and tried to use veridicality judgements instead. In the original works where the above methodology was proposed, the stimuli were framed as single utterances containing an inner contradiction of some sort, something of the form "A, in fact/specifically B". Such contexts appeared somehow unnatural to us, so we ran a parallel experiment using additional contexts and dialogues where one of the participants would not see the quantity of described objects, and asked the participants to present veridicality judgements using the same scale. For this experiment we have also measured the reading time using the self-paced reading method and reaction time for the grammaticality judgement. The structure of all sentences was identical to keep the reading time stable. The keyword with the number marker was always the third and it was followed by at least three words to avoid hesitation at the end of the sentence.

An example of a stimulus with SG-violation is presented in (2). First, the context was shown on the screen, and it would remain there. Then, participants would read a stimulus sentence with self-paced regulation, then see the entire sentence and then provide a veridicality judgement using a toggle bar. The violated and non-violated variants of the same item differed only in their contexts, while the stimulus remained the same.

(2)   {Maria Ivanovna asked Borya to tidy up the garden: "If you see tools lying around, take (them) and
       bring into the shed". Borya stumbled upon three spades and brought all of them into the shade. He
       finished the cleaning and came to give a report. He said:}

    Ya ostavil lopat-u   v   saraje   vozle   vhoda
    I   left   **spade-ACC.SG** in shed   near   entrance
    'I've left a spade in the shed near the entrance'

▲ Procedure and results

We used PCIbex farm (Zehr & Schwarz 2022) and distributed our experiments through Yandex.Toloka. For each experiment we included three trial-tasks at the beginning and used them to exclude people who appeared to answer randomly. Each experimental item on number had two inner variables – whether SG or PL was used and whether it was violated. Each item on implicatures had a single inner variable – it could be either a contradiction, an implicature violation or a control norm. For the veridicality judgement experiment, there were 8 items for number and 8 items for other implicatures, combined into 8 experimental sheets divided between 117 people. For the grammaticality judgement experiment, there were 4 items for number and 6 items for other implicatures divided between 39 people. We analyzed the data using *lmerTest, lme4* and *multcomp* packages in R (Bates et al. 2015). A single item on number was later removed from each experiment, as it was suspected to exhibit mass noun-like effects.

For the grammaticality judgements, the initial assumption proved to be true for quantifiers and past tense – logical contradictions were judged lower on the grammaticality scale than violations of an implicature (β = -31.214, SE = 7.192, z = 7.192, p < 0.001 for quantifiers). As for grammatical number, there appeared to be a considerable difference between judgements on SG violation and PL violation, but it was contrary to our expectations. While the expectation was that PL violation would be deemed more grammatical, the real

picture was reverse and in our case SG violations were judged to be more grammatical (β = 20.001, SE = 5.927, z = 3.374, p < 0.001 with mean judgement being 46.92).

For the truth-judgements, the initial assumption did not hold – implicatures were not judged better than logical contradictions. Judgements on number, however, have shown a slight difference between the SG violation and PL violation, with SG violations being rated as less truthful (β = -13.7, SE = 5.622, z = -2.437, p < 0.05 with mean judgement being 36.62). There was no significant difference in reading time for neither the keyword, nor the following words.

▲ Conclusion

Our findings show that the violation of PL is judged to be worse than the violation of SG, which contradicts the predictions made by an implicature theory of number. The data suggests that, if number is an implicature, then it should be the singularity meaning and not the strict plurality that is computed via pragmatic mechanisms. Grammaticality judgements are supported by self-paced reading data, which shows that violations of PL are grave enough to stagger the reading process, while violations of SG fail to do so.

▲ Links

■ Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software* 67(1). doi:10.18637/jss.v067.i01.　■　Cummins, Chris & Napoleon Katsos. 2010. Comparative and superlative quantifiers: Pragmatic effects of comparison type. *Journal of Semantics* 27(3). Oxford University Press, 271–305.　■　Ivlieva, Natalia. 2013. Scalar implicatures and the grammar of plurality and disjunction. Massachusetts Institute of Technology PhD Thesis.■　Katsos, Napoleon & Dorothy V.M. Bishop. 2011. Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition* 120(1), 67–81. doi:10.1016/j.cognition.2011.02.015.　■ Križ, Manuel. 2017. Bare plurals, multiplicity, and homogeneity.　■ Spector, Benjamin. 2007. Aspects of the Pragmatics of Plural Morphology: On Higher-Order Implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presupposition and Implicature in Compositional Semantics*, 243–281. London: Palgrave Macmillan UK. doi:10.1057/9780230210752_9.　■　Swart, Henriëtte de & Donka Farkas. 2010. The semantics and pragmatics of plurals. *Semantics and pragmatics* 3, 6–1.　■　Tieu, Lyn, Cory Bill, Jacopo Romoli & Stephen Crain. 2020. Testing theories of plural meanings. *Cognition* 205, 104307. doi:10.1016/j.cognition.2020.104307. ■ Zehr, Jérémy & Florian Schwarz. 2022. PennController for Internet Based Experiments (IBEX). Open Science Framework. doi:10.17605/OSF.IO/MD832. https://osf.io/md832/ (26 March, 2023).