Sentence comprehension assessment in Russian

Daria Chernova[1], Artem Novozhilov[1,] Natalia Slioussar[1,2]
[1]Institute for Cognitive Studies, St Petersburg University, Russia
[2]HSE University, Moscow, Russia
https://doi.org/10.36505/ExLing-2021/12/0012/000485

Abstract

We developed a Sentence Comprehension Test in Russian for syntactic competence assessment. It includes 60 unambiguous grammatically complex sentences of several types with comprehension questions aimed to test effective syntactic processing. The test does not show ceiling effects with adult native speakers. The results of the test correlate with verbal working memory span.

Keywords: syntactic complexity, sentence comprehension, Russian language

Introduction

The ability to parse syntactically complex sentences efficiently is a crucial skill for text comprehension, and numerous studies demonstrated considerable by-subject variability in performance on syntactic processing tasks (Farmer et al. 2012). However, while tests on sentence processing are widely used with various neurologically impaired populations, tests for healthy adult native speakers are difficult to develop due to ceiling effects and have been created only for English so far (Acheson 2008; Dabrowska 2018). In the present paper, we present the Sentence Comprehension Test we developed for Russian.

Materials

We created 60 unambiguous grammatically complex sentences of six types. These types are difficult to process according to previous experimental studies on Russian and other languages, see (1)-(6). All sentences were semantically reversible and unbiased. For every sentence, we created a question with a choice of two answers aimed to assess syntactic structure comprehension.

(1) object relative clauses (see Price & Witzel 2017, Malyutina et al. 2018)
Svidetel', *kotorogo* *upomjanul* *v svoej rechi* *istec,*
witness-NOM whom-ACC mentioned in his speech claimant-NOM
vskochil so svoego mesta v zale suda.
jumped from his seat in room court
'The witness that the claimant mentioned jumped up from his seat in the courtroom.'
Question: Who was mentioned?
Response options: A) the witness B) the claimant

(2) spatial constructions (see Laurinavichyute et al. 2017)
Passazhir sprjatal v seryj jashhik kozhanyj chemodan.
passenger-NOM hid into grey box-ACC leather suitcase-ACC.
'The passenger hid the leather suitcase into the grey box.'
Question: What was hidden where?
Response options: A) the box in suitcase B) the suitcase in the box

(3) temporal constructions (see Fedorova 2005)
Pered tem kak Tolja propylesosit pol, Julja vyguljaet sobaku.
before Tolja vacuum-cleans floor Julia walks dog
'Before Tolja vacuum cleans the floor, Julia will walk the dog.'
Question: What happens first?
Response options: A) Tolja vacuum cleans the floor B) Julia walks the dog

(4) sentences with high adjunct attachment in a complex noun phrase *Konvert peredali pomoshhniku detektiva, sledivshemu*
envelope-ACC gave assistant-DAT detective-GEN following-DAT
za podozrevaemym.
after suspect-INST
'The envelope was given to the assistant$_i$ of the detective$_j$, following$_i$ the suspect.'
Question: Who followed the suspect?
Response options: A) the detective B) the assistant

(5) sentences with low adjunct attachment in a complex noun phrase *Notarius napisal naslediku millionera, zhivshego za granicej.*
notary wrote heir-DAT millionaire-GEN living-GEN abroad
'The notary wrote to the heir$_i$ of the millionaire$_j$ living$_j$ abroad.'
Question: Who lived abroad?
Response options: A) the millionaire B) the heir

(6) comparative constructions
Sherstjanaja jubka dlinnee shelkovoj, no koroche l'njanoj.
woolen-NOM skirt-NOM longer silk-INST but shorter linen-INST
'The woolen skirt is longer than the silk one, but shorter than the linen one.'
Question: Which skirt is longer?
Response options: A) the silk one B) the linen one

The test also included 40 filler sentences with a simpler syntactic structure (7).
(7) *Na ploshhadke ja vstretil brata moego druga s bol'shoj sobakoj*
at playground I met brother-ACC my friend-GEN with big dog
'At the playground I met my friend's brother with a big dog'.
Question: Who did he meet?
Response options: A) his brother's friend B) his friend's brother

Method

42 native speakers of Russian (29 female, age 19-32) took part in the experiment. The word-by-word self-paced reading methodology was used because otherwise the task would be too easy. The incorrect response was always a noun that was mentioned in the sentence, so the grammatical structure of the sentence had to be analyzed to give a correct answer. The order of the response options was random. Accuracy, word by word reading times, and time to give the answer were registered. We also measured participants' verbal working memory span using the Russian adaptation (Fedorova 2003) of the test by (Daneman&Carpenter 1980) as syntactic processing was shown to be affected by the characteristics of the working memory (Caplan & Waters 1999).

Results and discussion

For the statistical analysis, we used logistic and linear mixed-effects regressions with random intercepts and slopes by participant and by item and Tukey's tests for post hoc comparisons. Test sentences were significantly more difficult to process than fillers, both in terms of correct answers (80.6% vs. 92.6% on average; $\beta = 0.25$, $SE = 0.04$, $t = 6.01$, $p < 0.01$) and reading and response times (711.4 ms vs. 595.6 ms, $\beta = 0.25$, $SE = 0.04$, $t = 6.01$, $p < 0.01$; 3484.2 ms vs. 3096.3 ms, $\beta = 0.25$, $SE = 0.04$, $t = 6.01$, $p < 0.01$). This proves the validity of the test.

Another important proof comes from the fact that we detected significant variation between participants. In target sentences, they made from 1 to 24 errors (98%-60% correct answers). We found a significant correlation between answer accuracy and working memory span test scores ($r=0.59$, $p<0.01$). The number of errors in filler sentences did not vary that much (from 0 to 8, which means 100%-80% correct answers, with ¾ participants making no more than two errors).

There were also significant differences between most construction types. High/low attachment and comparative constructions had longer word-by-word reading times than other target sentence types ($t > 6.74$, $p<0.01$ for all pairwise comparisons). Spatial and comparative constructions had longer response times than other sentence types ($t > 5.58$, $p<0.01$ for all pairwise comparisons).

Finally, high and low attachment sentences triggered the largest number of incorrect responses (74.3% and 62.6%, respectively). Low attachment sentences were significantly different from all other types (expect for high attachment, $t > 3.78$, $p<0.01$ for all pairwise comparisons), while for high attachment sentences, only some comparisons gave significant results. Importantly, every target type was significantly different from fillers, except for temporal one ($t > 3.61$, $p<0.01$ for all pairwise comparisons), which further proves the validity of the test.

Thus, sentence types that take more time to read and especially to answer are not the ones in which participants make more mistakes. This may point to two

different aspects of syntactic complexity. In some cases, arriving at any coherent interpretation is difficult; in the other cases, one arrives at some interpretation easily, but often this is not the correct one. The generalizations we made about different constructions may be useful for further processing studies. To conclude, our pilot study showed that the test we developed is far from trivial for the participants and does not show ceiling accuracy. After validation, it can be used in various studies as a tool to measure syntactic processing efficiency.

Acknowledgements

The study is supported by RSF, project № 21-78-00064.

References

Acheson, D. J., et al. 2008. New and updated tests of print exposure and reading abilities in college students. Behavior Research Methods, 40, 278–289.

Caplan, D., Waters G. 1999. Verbal Working Memory and Sentence Comprehension. Behavioral and Brain Sciences, 22, 1, 77-94.

Dąbrowska, E. 2018. Experience, aptitude and individual differences in native language ultimate attainment. Cognition, 178, 222–235.

Daneman, M., Carpenter P. 1980. Individual Differences in Working Memory and Reading. Journal of Verbal Learning and Verbal Behavior, 19, 4, 450–466.

Farmer, T.A., Misyak, J B., Christiansen, M.H. 2012. Individual differences in sentence processing. In M. J. Spivey, K. McRae, & M. F. Joanisse (Eds.), Cambridge handbooks in psychology. The Cambridge handbook of psycholinguistics (p. 353–364). Cambridge University Press.

Fedorova, O.V. 2003. Test po opredeleniju ob'ema operativnoj pamjati: istorija i sovremennoe sostojanie (in Russian, 'Survey of the State of the Art in Verbal Span Test'). Computational Linguistics and Intellectual Technologies.

Fedorova, O.V. 2005. Pered or posle: chto proshche. (in Russian, 'Before or after: What is easier'). Voprosy Jazykoznanija, 2005, 6, 44-58.

Laurinavichyute, A., Chrabaszcz, A., Farizova, N., Tolkacheva, V., Dragoy, O. 2017. Vlijanie sensomotornyh stereotipov na ponimanie prostranstvennyh konstrukcij: dannye dvizhenij glaz (in Russian, 'The influence of sensorimotor stereotypes on the understanding of spatial constructions: eye-tracking evidence'). Voprosy Jazykoznanija, 3, 99-109.

Malyutina, S., Laurinavichyute, A., Terekhina, M., Lapin, Y. 2018. No evidence for strategic nature of age-related slowing in sentence processing. Psychology and Aging, 33(7), 1045–1059.

Price, I.K., Witzel J. 2017. Sources of Relative clause processing difficulty: Evidence from Russian. Journal of Memory and Language, 97, 208–244.