



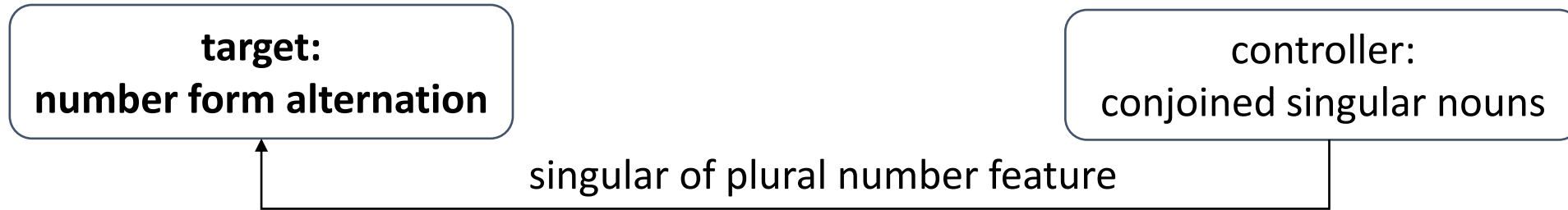
Night Whites 2022

Parametrizing
number variation:
evidence from Russian noun phrases
with coordinated modifiers

Kseniia Studenikina
Lomonosov Moscow State University

This work was supported
by Non-commercial Foundation
for support of Science and
Education «INTELLECT»

Agreement with coordinated structures



- Parametrization: morphological, syntactic and semantic factors
- inanimate abstract nouns → singular agreement

(1) *mezhdunarodn-ogo/ ?mezhdunarodn-ykh* *mir-a* *i* *bezopasnost-i*
international-GEN.SG international-GEN.PL peace-GEN.SG and security-GEN.SG
'of international peace and security'

- the proper names → plural agreement

(2) *perspektivn-yje/* *?perspektivn-yj* *Pimenov* *i* *Izmajlov*
promising-NOM.PL *promising-NOM.SG* *Pimenov.NOM.SG* and *Izmajlov.NOM.SG*
'promising Pimenov and Izmajlov'

Russian noun phrases with coordinated modifiers

- Number variation of agreement controller
- Two conjoined singular adjectives with split interpretation → singular or plural noun

(3) *v imenitel'n-om i vinitel'n-om padezh-e/ padezh-akh*
in nominative-LOC.SG and accusative-LOC.SG case-LOC.SG case-LOC.PL
'in nominative and accusative cases'

How to reveal the factors parametrizing variation?

Previous studies:

- Introspection (Kodzasov 1987)
- Corpus (Pekelis 2013)



Our study:

- Experimental method 😊
- Language modeling 🤖



Experimental studies

- controlled data from many speakers
 - subtle differences in acceptability
 - data about both acceptable and unacceptable sentences

 - offline data: acceptability judgements (Likert scale 1-7)
 - z-normalized scores
 - online data: self-paced reading
 - reading time
- 
- 

Language modeling

- acceptability judgments in natural language processing
 - language models' robustness
 - acquisition of grammatical phenomena
- generate texts almost indistinguishable from human texts
- trained on datasets without grammar
- sentence probability predicted by ruGPT-3
 - perplexity (Lau et al. 2020, Mikhailov et al. 2022)

Goals of the study:

- parametrizing variation
- comparing human judgements and model probabilities

1. The effect of noun morphology: language material

- Bulgarian: the irregular number morphology causes → preference for the singular noun (Harizanov & Gribanova 2015)

- 64 sentences:

- 32 stimuli

- 32 fillers

- 16 grammatical

(4) *slab-yj* *vostochn-yj* *veter* *nagnal* *gust-oj* *tuman.*
weak-M.SG east-M.SG wind.M.SG bring.PST.M thick-SG.ACC fog.SG.ACC

'A weak east wind brought a thick fog.'

- 16 ungrammatical

(5) *parizhsk-aja* *nacional'n-yj* *oper-a* *otkryl-a* *nov-yj* *sezon.*
Paris.ADJ-F.SG national-M.SG opera-F.SG open.PST-F new-SG.ACC season.SG.ACC

'The Paris National Opera has opened a new season.'

1. The effect of noun morphology: factors

Design: 2 × 4

- the noun number (singular / plural)
- the morphological noun type:
 - suppletion and stem alternations

(6) *vysok-ij* *i* *nizk-ij* ***chlovek/***
tall-SG and short-SG man

ljudi *perebezhali*
people run.PST-PL

cherez *dorog-u.*
across road-SG.ACC

‘A tall and a short man ran across the road.’

1. The effect of noun morphology: factors

Design: 2 × 4

- the noun number (singular / plural)
- the morphological noun type:
 - suppletion and stem alternations
 - suffix alternations

<i>(7) bur-yj</i>	<i>i</i>	<i>bel-yj</i>	<i>medvezhonok/ medvezhata</i>	<i>adaptirovalis'</i>	<i>v</i>	<i>zoopark-e.</i>
brown-SG	and	white-SG	bear.SG bear.PL	adapt.PST.PL	in	ZOO-SG.LOC

'The brown and the white bears adapted in the zoo.'

1. The effect of noun morphology: factors

Design: 2 × 4

- the noun number (singular / plural)
- the morphological noun type:
 - suppletion and stem alternations
 - suffix alternations
 - syncretic forms
 - singularia tantum

(8) *a. tvorcheska-ja i nauchna-ja molodjoz spel-i na prazdnik-e.*
creative-SG and scientific-SG youth sing.PST-PL at festival-SG.LOC

‘Creative and scientific youth sang at the festival.’

- pluralia tantum

b. bolsh-ie i malen’k-ije ochki zavalilis’ pod krovat’.
big-PL and small-PL glasses fall.PST.PL under bed.SG.ACC

‘Big and small glasses fell under the bed.’

1. The effect of noun morphology: factors

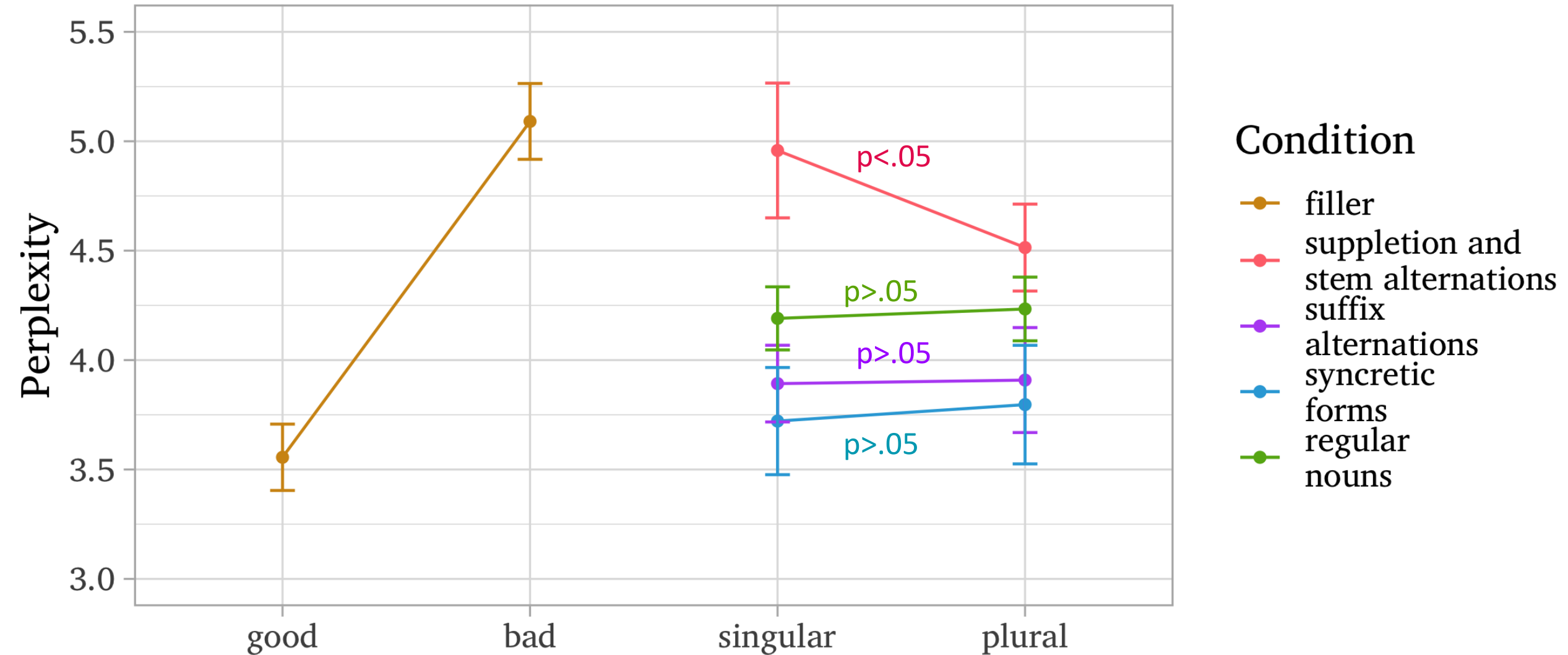
Design: 2 × 4

- the noun number (singular / plural)
- the morphological noun type:
 - suppletion and stem alternations
 - suffix alternations
 - syncretic forms
 - singularia tantum
 - pluralia tantum
 - regular nouns

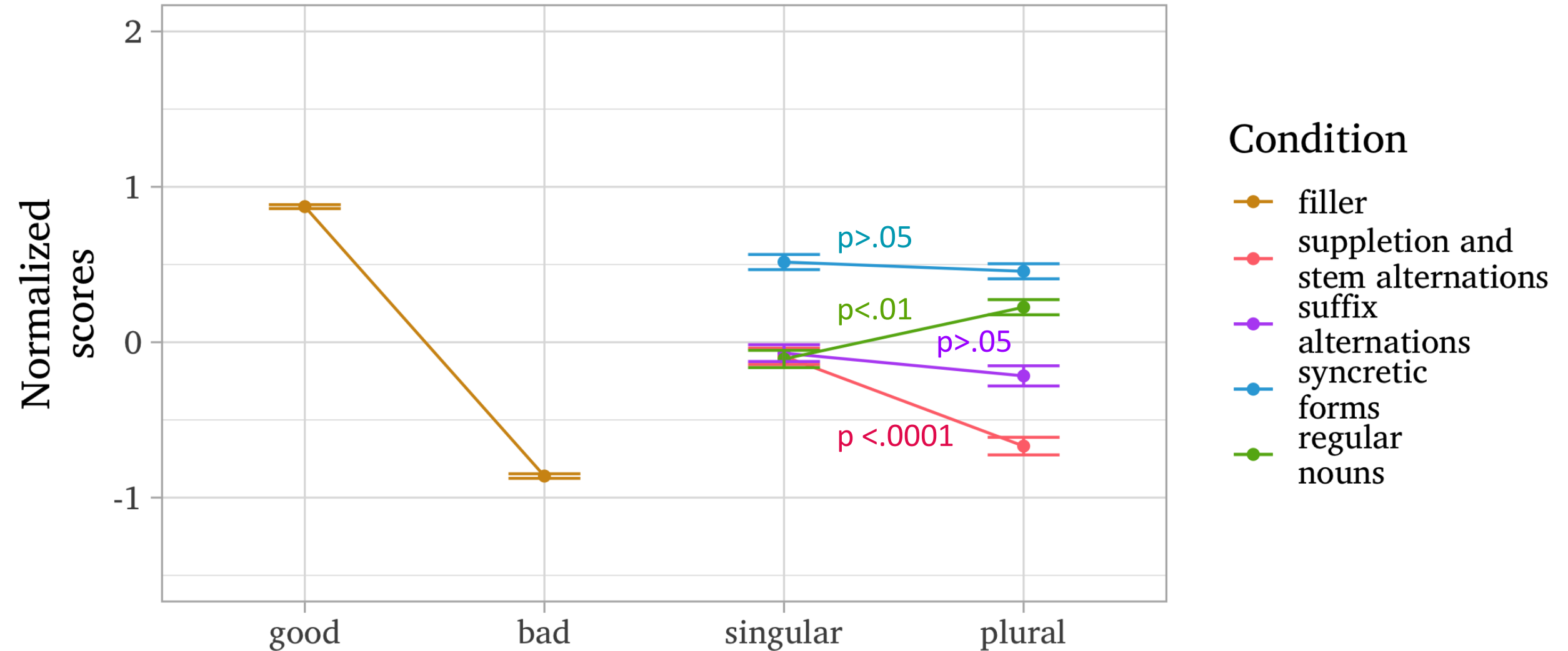
(9) *pozhl-oj i molod-oj prepodavatel'/ prepodavatel-i voshl-i v auditori-ju.*
elderly-SG and young-SG teacher.SG teacher-PL go.PST-PL in classroom-SG.ACC

'An elderly and a young teachers entered the classroom.'

1. The effect of noun morphology: language modeling

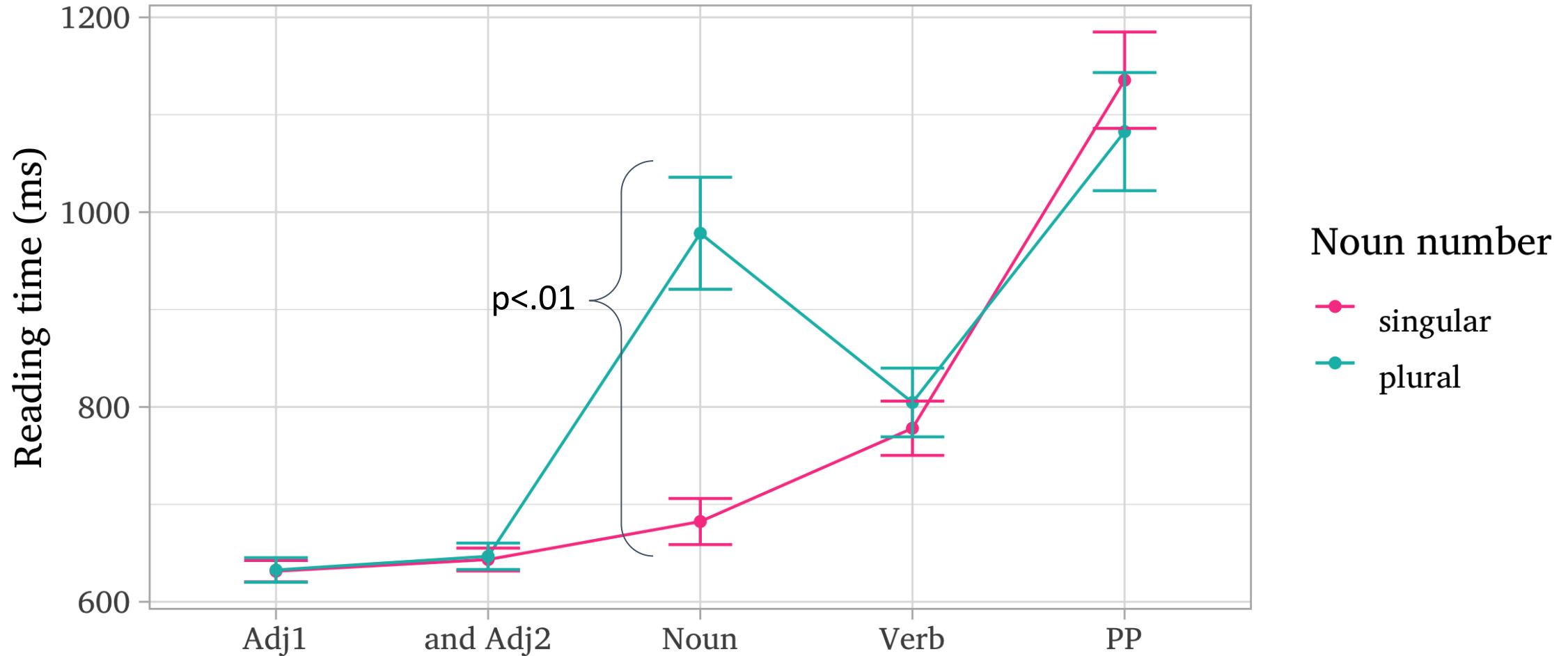


1. The effect of noun morphology: acceptability judgements



1. The effect of noun morphology: reading time

Suppletion and stem alternations



1. The effect of noun morphology: results

😊 Human results

- Stem alternations: singular number → higher acceptability scores, less reading time
- Suffix alternations: free variation
- It is stem rather than suffix alternation that influences the number choice
- Syncretic nouns: no number choice → the highest scores, no reading delay

🤖 Language model results

- Syncretic nouns: the highest probability
- Suffix alternations, syncretic nouns: equal probability for singular and plural forms
- Stem alternations: preference of singular forms is not predicted

2. The effect of premodifier agreement: language material

- Predicate plural agreement → preference for the plural noun (Kodzasov 1987)

- 64 sentences:

- 32 stimuli
- 32 fillers
 - 16 grammatical

(10) *Ljuba pobyval-a v znamenit-om moskovsk-om cirk-e v kompani-i rodstvennik-ov*

Ljuba visit.PST-F.SG in famous-LOC.SG Moscow.ADJ-LOC.SG circus-LOC.SG in company-LOC.SG relative-GEN.PL

‘Lyuba visited the famous Moscow circus in the company of relatives.’

- 16 ungrammatical

(11) *Larisa pobedil-a v mezhdunarodn-ykh muzykal'n-ykh konkurse v konc-e zim-y.*

Larisa win.PST-F.SG in international-LOC.PL music-LOC.PL competition-LOC.SG in end-LOC.SG winter-GEN.SG

‘Larisa won international music competitions at the end of winter.’

2. The effect of premodifier agreement: factors

Design: 2 × 2 × 2

- the premodifier number (singular / plural)
- the noun number (singular / plural)
- the case of the noun phrase (direct: accusative / oblique: instrumental)

(12) Prodavec predložil dva kostjum-a na vybor. (context)
assistant.SG offer.PST two suit-PL for choice

‘The assistant offered two suits for choice.’

(13) Vanja primeril **et-ot/** **et-i** sin-ij i korichnev-yj **kostjum/ kostjum-y**
Vanja try_on.PST this-ACC.SG/ this-ACC.PL blue-ACC.SG and brawn-ACC.SG suit.ACC.SG/ suit-ACC.PL

v magazine-e odezhd-y.

in store-LOC.SG clothes-GEN.SG

‘Vanya tried on these blue and brown suits in a store.’

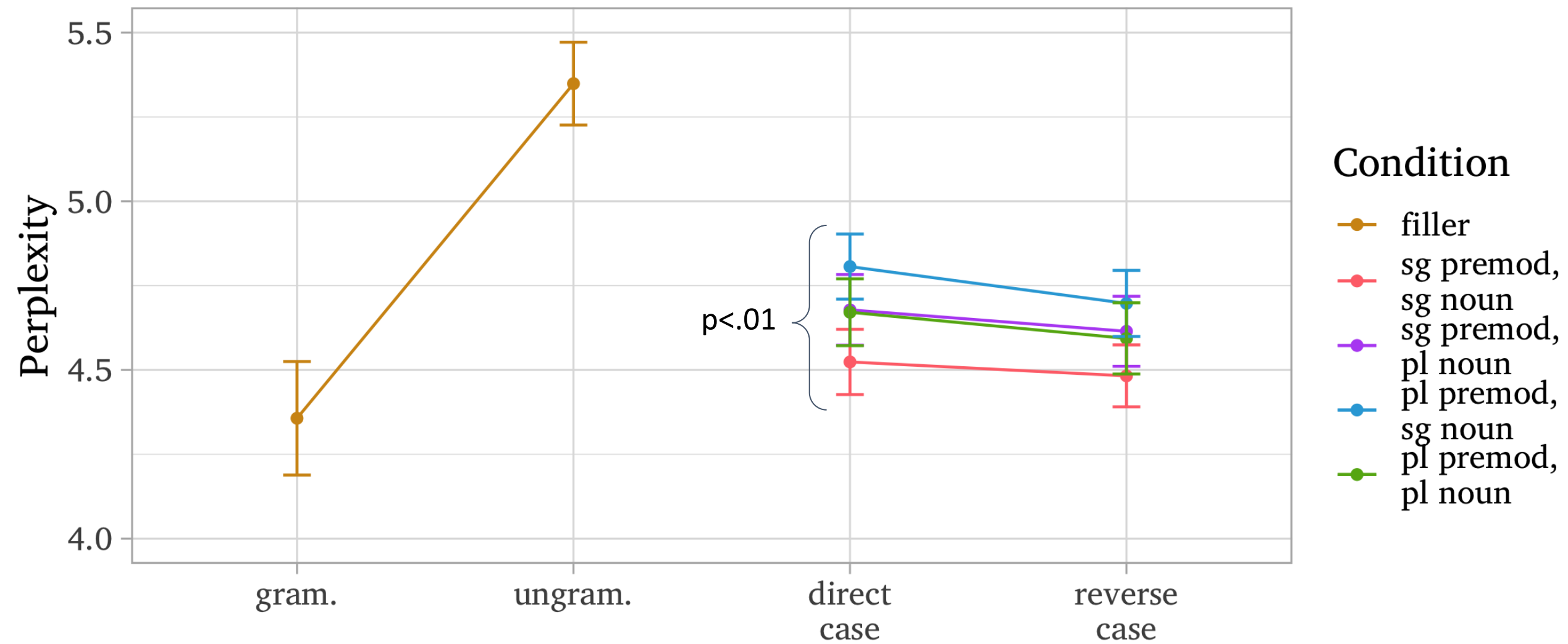
(14) Vanja zaljubovalsja **et-im/** **et-imi** sin-im i korichnev-ym **kostjum-om/ kostjum-ami**
Vanja admire.PST this-INSTR.SG/ this-INSTR.PL blue-INSTR.SG and brawn-INSTR.SG suit-INSTR.SG/ suit-INSTR.PL

v magazine-e odezhd-y.

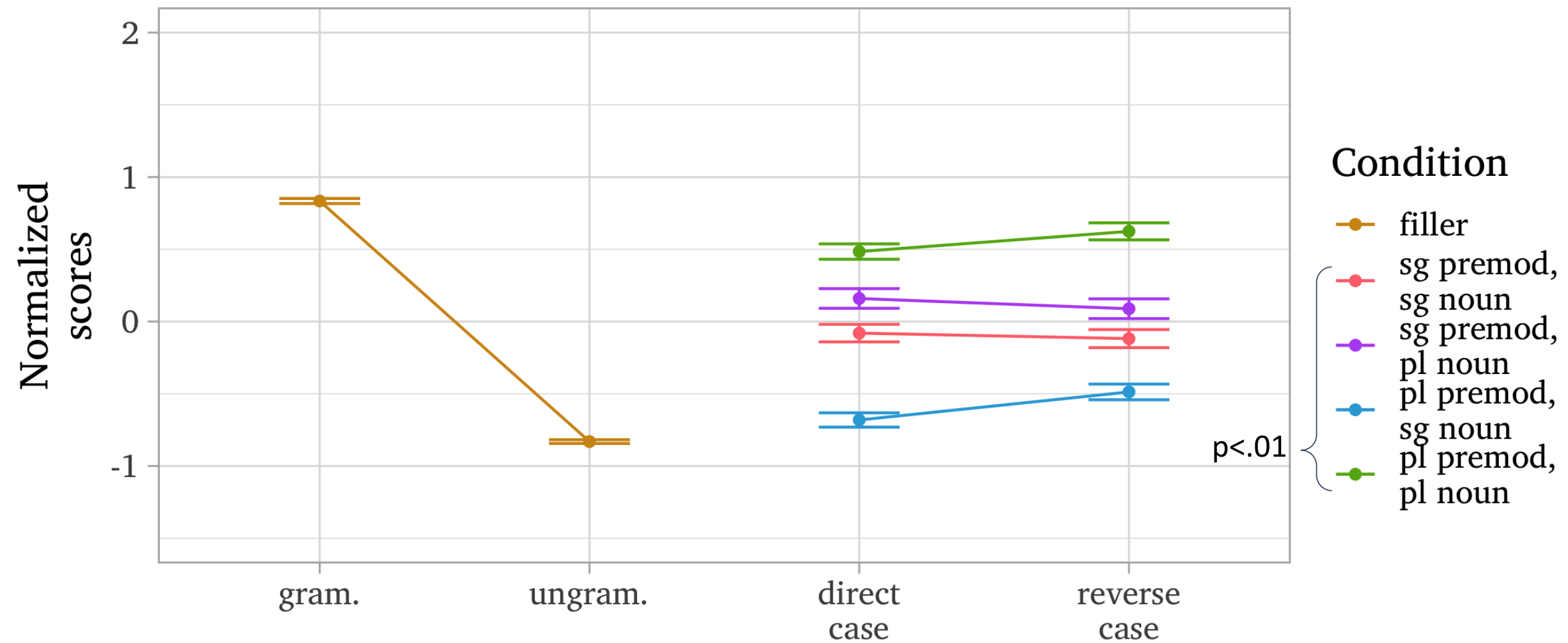
in store-LOC.SG clothes-GEN.SG

‘Vanya admired these blue and brown suits in a store.’

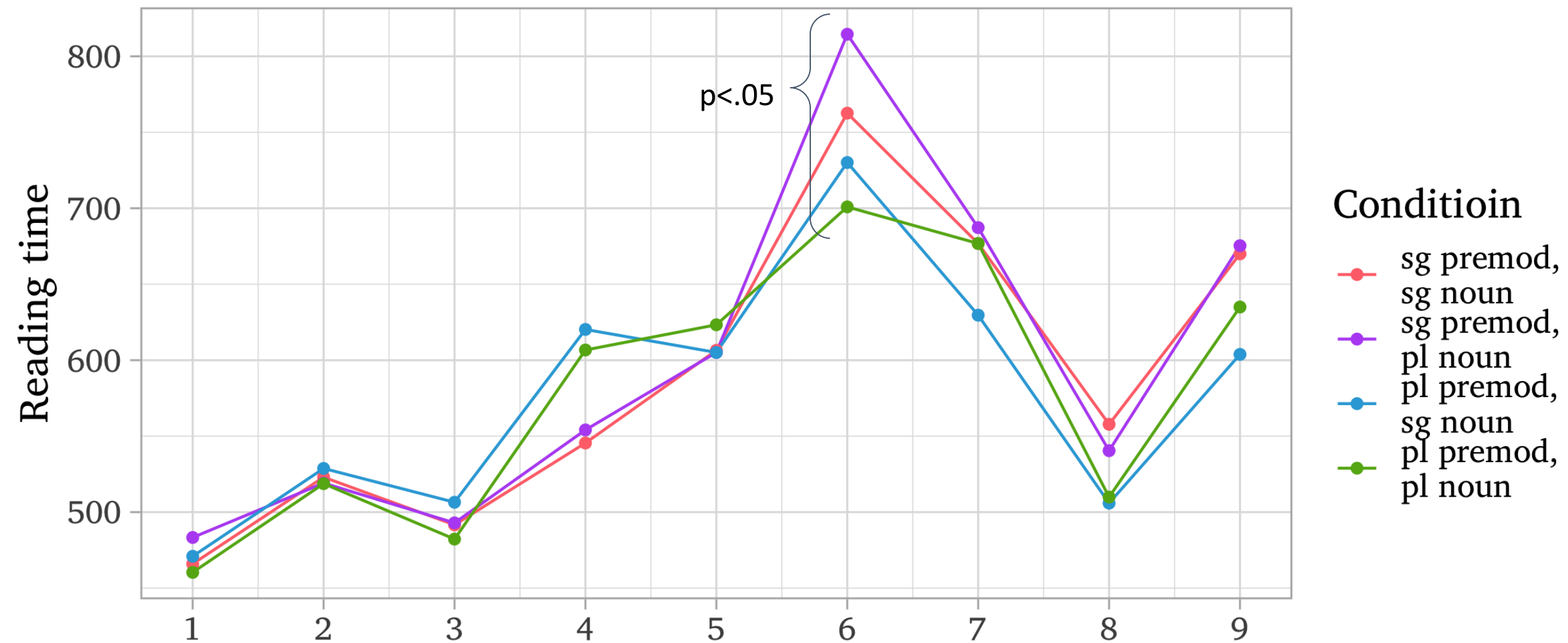
2. The effect of premodifier agreement: language modeling



2. The effect of premodifier agreement: acceptability judgements



2. The effect of premodifier agreement: reading time



2. The effect of premodifier agreement: results

😊 Human results

- Same number features of premodifier and noun: both singular or both plural are preferable
- Number mismatch: acceptable for plural noun (reading delays), unacceptable for singular noun

🤖 Language model results

- Plural premodifier and singular noun: the lowest probability
- Same number features of premodifier and noun: high probability
- Plural noun: preference for plural premodifier is not predicted

Conclusions

😊 Experimental studies

- We have collected offline and online data about the perception of Russian noun phrases with coordinated modifiers.
- Observed morphological and syntactic factors are significant and should be considered for parametrizing number variation.

🤖 Language modeling

- ruGPT-3 predicts correct probabilities for highly acceptable and highly unacceptable sentences (fillers and stimuli).
- The probabilities of variation cases do not correspond with the data about human perception.

References

Harizanov B., Gribanova V. How across-the-board movement interacts with nominal concord in Bulgarian // Proceedings from the Annual Meeting of the Chicago Linguistics Society 49. University of Chicago, IL: Chicago Linguistics Society, 2015. Pp. 115–129.

Lau J. H., Armendariz C., Lappin S., Purver M., Shu C. How furiously can colorless green ideas sleep? sentence acceptability in context // Transactions of the Association for Computational Linguistics, 8, 2020. Pp. 296-310.

Kodzasov S. V. Chislo v sochinitel'nykh konstruktsiyakh. [Number in coordinated constructions]. // Modelirovanie yazykovoï deyatel'nosti v intellektual'nykh sistemakh / A.E. Kibrik, A.S. Narin'yani (eds.). Moscow: Nauka, 1987. Pp. 204–219.

Mikhailov V., Shamardina T., Ryabinin M., Pestova A., Smurov I., Artemova E. Russian Corpus of Linguistic Acceptability // arXiv preprint arXiv:2210.12814. 2022.

Pekelis O. E. Sochinenie [Coordination]. [Materials for the project of Russian corpus grammar](#). Manuscript copyright, 2013.



Thank you for your attention!



Appendix



ruGPT-3 generated text


Напиши отзыв о стиральной машине.

Данная модель стиральной машины оснащена встроенной функцией самоочистки.

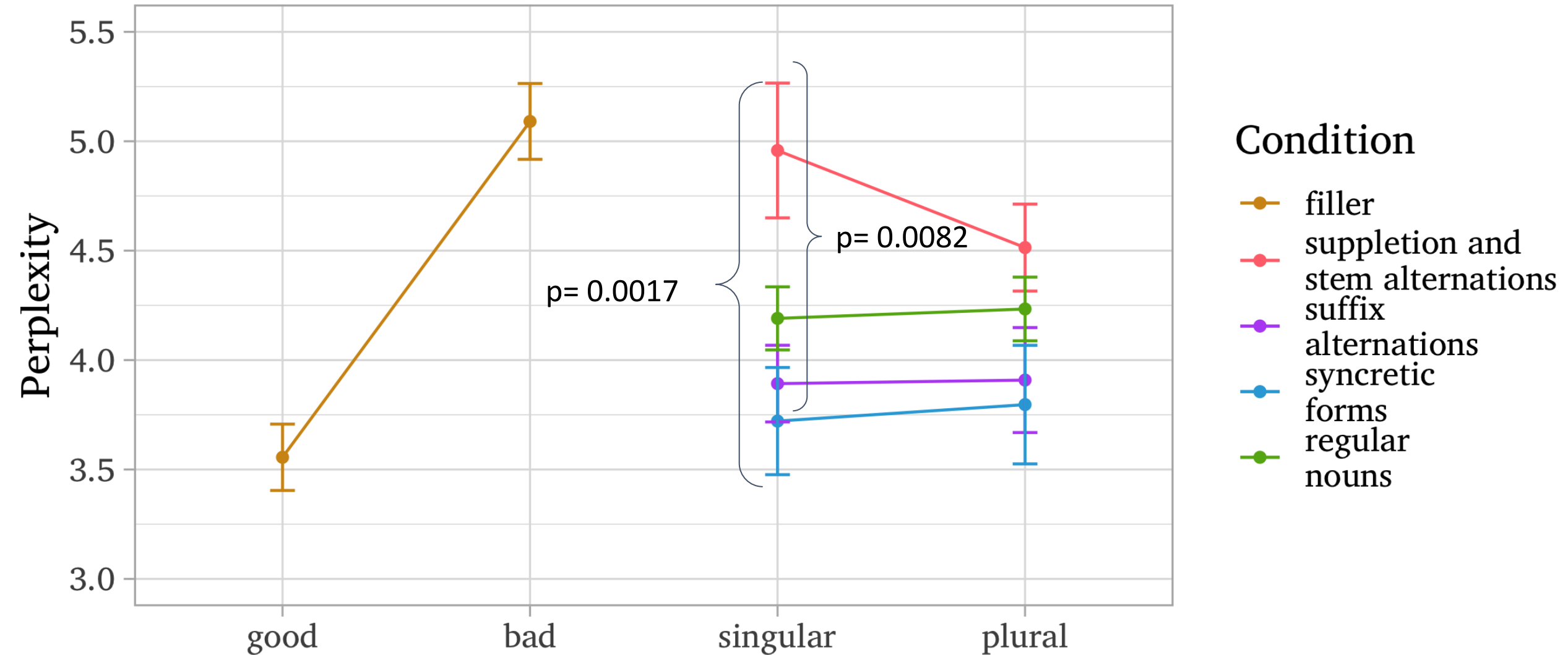
В конструктивном решении стиральной машины предусмотрено устройство автоматического затирания воды, которая способна удалять накипь и прочие загрязнения даже в труднодоступных местах.

Write a review about the washing machine.

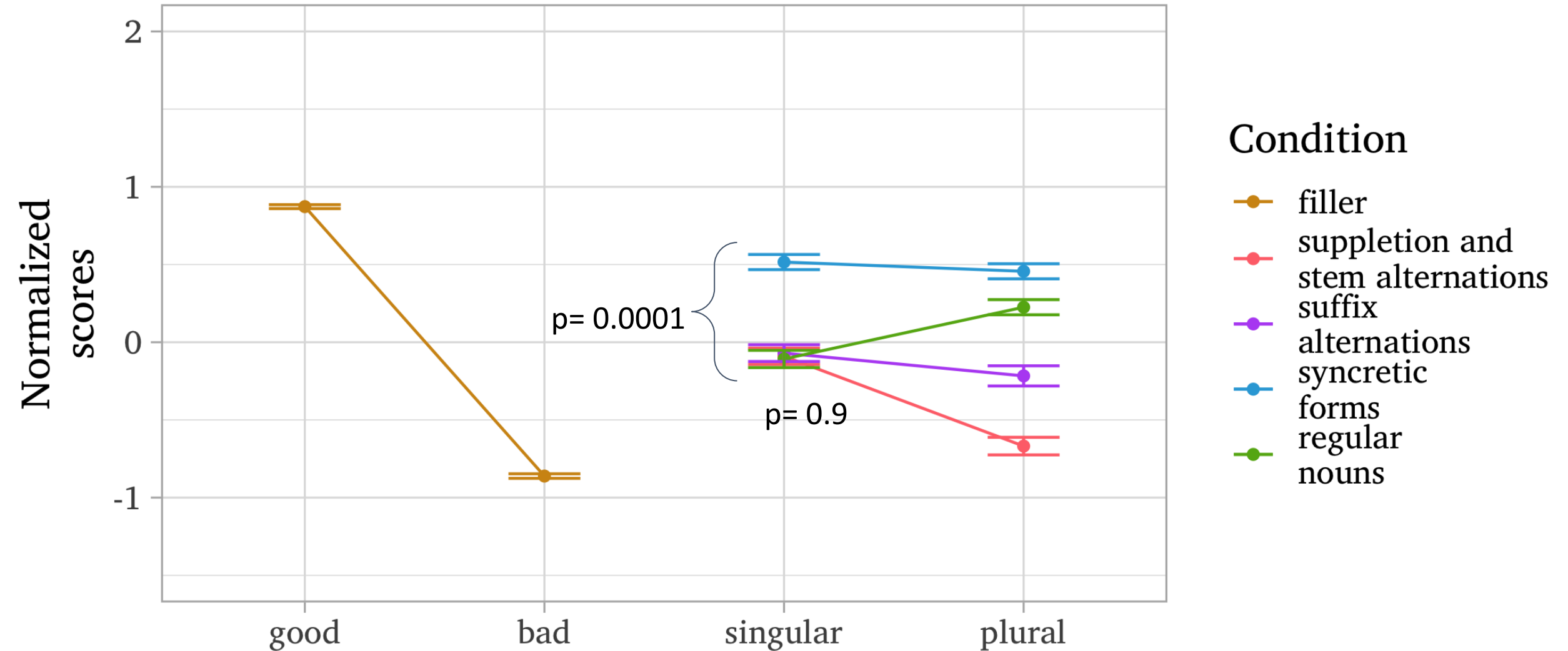
This model of washing machine is equipped with a built-in self-cleaning function. The constructive solution of the washing machine provides an automatic water mashing device, which is able to remove scale and other contaminants even in hard-to-reach places.



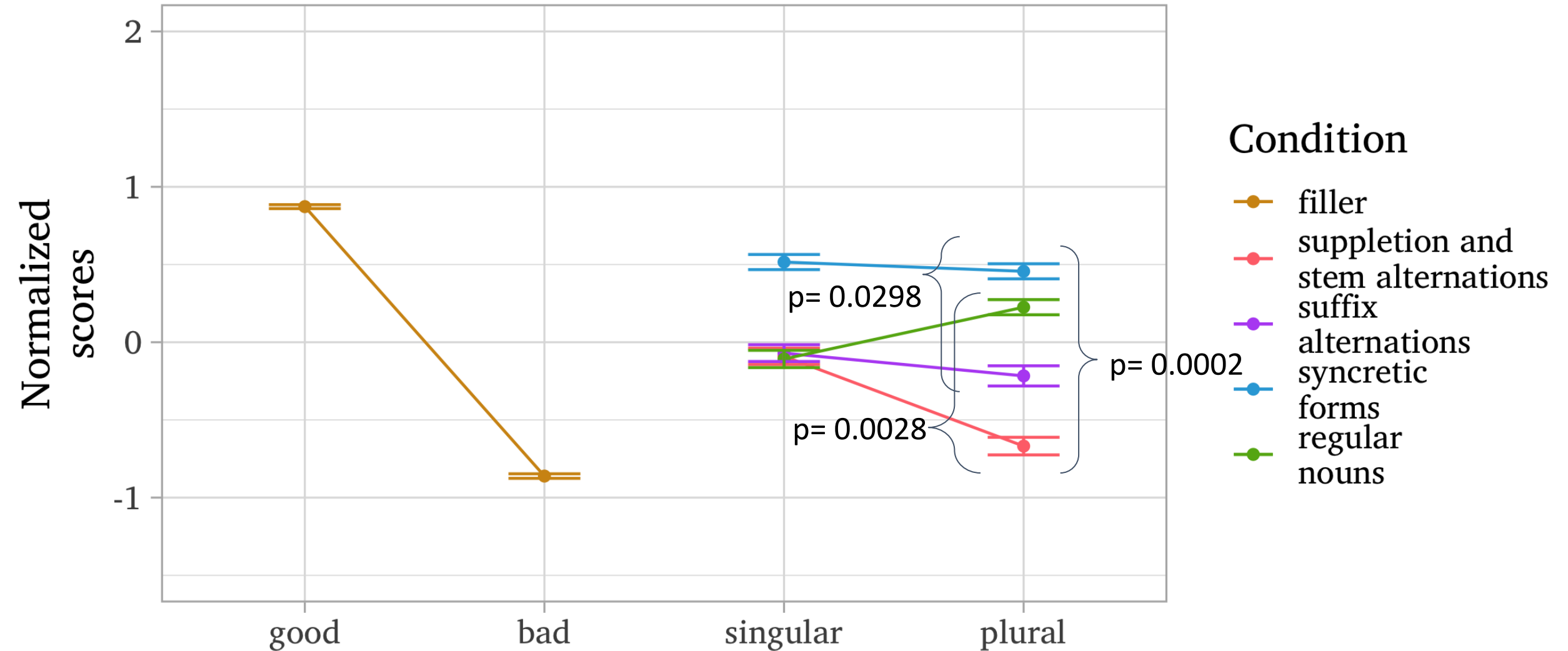
1. The effect of noun morphology: language modeling



1. The effect of noun morphology: acceptability judgements

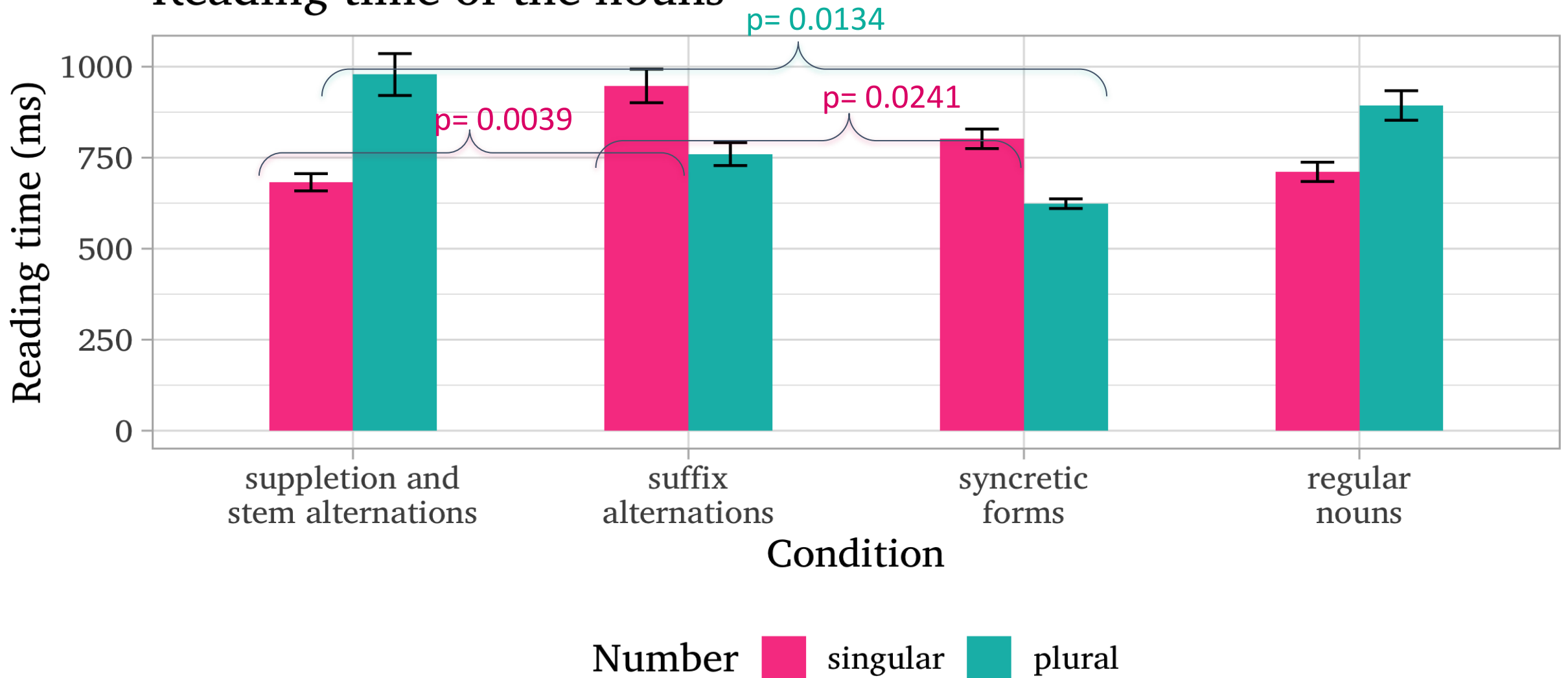


1. The effect of noun morphology: acceptability judgements



1. The effect of noun morphology: reading time

Reading time of the nouns



2. The effect of premodifier agreement: reading time

