Neural network learning of the Russian genitive of negation: optionality and structure sensitivity

Natalia Talmina
Department of Cognitive Science
Johns Hopkins University
talmina@jhu.edu

Tal Linzen
Department of Cognitive Science
Johns Hopkins University
tal.linzen@jhu.edu

Abstract

A number of recent studies have investigated the ability of language models (specifically, neural network language models without syntactic supervision) to capture syntactic dependencies. In this paper, we contribute to this line of work and investigate the neural network learning of the Russian genitive of negation. The genitive case can optionally mark direct objects of negated verbs, but it is obligatory in the existential copula construction under negation. We find that the recurrent neural network language model we tested can learn this grammaticality pattern, although it is not clear whether it learns the locality constraint on the genitive objects. Our results further provide evidence that RNN models can distinguish between optionality and obligatoriness.

1 Introduction

Statistical language models are probability distributions over sequences of words, which they learn from large corpora during training. For any given context, these models assign a probability to all of its possible continuations: for a example, given the context "he was eating soup with a. . . ", language models can predict that the word "spoon" is much more likely to occur next than "shoe".

A class of language models – Recurrent Neural Network (RNN) models – have been particularly successful on various applied language tasks (Mikolov et al., 2010; Vinyals et al., 2015; Kiperwasser and Goldberg, 2016; Bahdanau et al., 2014). But what kind of linguistic knowledge do these models capture? Arguably, human language knowledge is comprised of more than word co-occurrence statistics – it encompasses abstract rules and generalizations that concern hierarchical structure. According to the argument from the poverty of the stimulus (Chomsky, 1980), the kind of structural knowledge that underlies human linguistic performance is impossible to derive purely from the input language learners receive, since many structure-dependent linguistic phenomena are too infrequent in the type of input humans encounter during language acquisition. Therefore, according to the argument, human sensitivity to the structure in language must be innate.

Since neural networks do not possess this innate bias – but perform applied natural language tasks with high accuracy – they can provide a rich source of information about the mechanisms underlying hierarchical structure rule learning. A number of questions need to be asked. How much grammar can language models learn just from a corpus? What are the limitations on the generalizations they can make about hierarchical structures? Recently, several studies have addressed these questions by testing RNNs' performance on structure-sensitive grammatical tasks. The results of these studies showed that RNNs can learn subject-verb agreement (Linzen et al., 2016; Gulordava et al., 2018; Ravfogel et al., 2018), filler-gap dependencies (Wilcox et al., 2018), hierarchical rules of question formation (McCoy et al., 2018), and the contexts that license negative polarity items (Jumelet and Hupkes, 2018).

In this paper, we contribute to this line of research by extending it to issues in Russian syntax. What makes Russian compelling is that it has rich morphology, which allows us to expand the range of tasks that have been used in previous work to explore RNN learning of structural dependencies. In particular, Russian has case-marking alternations involving the genitive case: along with the accusative case (which is typical cross-linguistically), the genitive can mark direct objects of transitive verbs. However, it is only licensed under negation, and is **optional** – the accusative case can be used in both affirmative and negative clauses. The genitive also alternates with the nominative case to mark the subjects of existential copula constructions, where it is **obligatory** under negation. Nominative subjects are only allowed with affirmative sentences. We spell out these properties in more detail in the next section.

2 Background: Russian genitive-of-negation

In Russian, direct objects are usually marked by the accusative case, as is common in languages with overt case marking:

(1) Uchitel proveril domasniye zadaniya
Teacher graded homeworksACC
"The teacher graded the homeworks."

However, non-oblique arguments can receive genitive case in the scope of sentential negation – a phenomenon known as the genitive of negation (Bailyn, 1997; Pesetsky, 1982; Paducheva, 2004; Harves, 2002; Timberlake, 1975; Babby, 1980):

(2) Uchitel **ne** proveril domasniye zadaniya
Teacher **neg** graded homeworks.ACC
"The teacher did not grade the homeworks."

(3) Uchitel **ne** proveril domasnih zadaniyj
Teacher **neg** graded homeworks.GEN
"The teacher did not grade the homeworks."

If the sentence is affirmative, only the accusative case can be used to mark the direct object:

(4) * Uchitel proveril domasnih zadaniyj
Teacher graded homeworks.GEN
"The teacher graded the homeworks."

Further, the genitive is only licensed when the negation term is local: in sentences like (5), the relative clause negation cannot license genitive case-marking on the main verb object *domasnih zadaniyj*. We will refer to this licensing pattern as the LOCALITY CONSTRAINT.

(5) * Uchitel, kotoryj ne lyubil studentov,
Teacher who **neg** like students
proveril domasnih zadaniyj
graded homeworks.GEN
"The teacher, who didn't like the students, graded the homeworks."

The genitive of negation is considered to be **optional** in sentences like (3) (Kagan 2010, although see Bailyn 1997; Harves 2002 for discussion), but it is **obligatory** in the existential copula construction, where the genitive alternates with the nominative case:

(6) (Bailyn, 1997)

a. Na stole **net** knig
on table **neg** books.GEN
"There are no books on the table."

b. * Na stole **net** knigi
on table **neg** books.NOM
"There are no books on the table."

3 Overview of experiments

Motivated by the observations in the previous section, we explored how well language models can capture the properties of the genitive of negation. We ran a series of experiments to study the behavior of an RNN language model trained by Gulordava et al. (2018). In Experiment 1, we tested the language model on simple sentences with case-marking alternation on direct objects, finding that the model learned the grammaticality pattern in (3–4). In Experiments 2–4, we tested whether the model was sensitive to the structurally defined scope of negation. We found that the model correctly predicted the genitive-accusative alternation even when there was no overt marking of sentential scope. In Experiment 5, we tested the model on the existential copula construction in which the genitive case is obligatory under negation. Our results suggest that the model could differentiate between the syntactic structures where the genitive case is obligatory from those where it is optional.

4 Methodology

To explore whether RNN language models can capture the constraints on genitive-marked direct objects, we studied the performance of the model presented in Gulordava et al. (2018). The model was trained on a 90-million-word corpus extracted from the Russian Wikipedia and had two layers of 650 hidden LSTM units. Additionally, we trained a 3-gram model on the same corpus to provide a baseline for our experiment. The 3-gram model which backs off to smaller n-grams using linear interpolation.

Following previous work (Linzen et al., 2016; Gulordava et al., 2018; Marvin and Linzen, 2018), we assessed the model's performance by examining the probabilities it assigned to grammatical sentences from our dataset, compared to ungrammatical ones. We used surprisal (Hale, 2001):

$$\text{surprisal}(w_i) = -\log \mathrm{P}(w_i \mid w_1 \ldots w_{i-1})$$

The higher the surprisal, the more unexpected a word is under the model's probability distribution. Since the sentences in (3) and (4) are minimally different from each other (the only difference being that the verb in (3) is negated), we can directly compare the surprisal the model assigned to the genitive-marked objects in these sentences. Assuming the probability distribution defined by the model reflects the grammar of the genitive of negation construction, we expected that the genitive-marked object would be assigned higher surprisal in (4), where it is not licensed by negation. Since accusative objects are grammatical independently of polarity, we did not expect the same difference between (1) and (2).

5 Experiments

5.1 Experiment 1: Simple sentences

5.1.1 Materials

We constructed a dataset of 64 sentences, each consisting of a subject, a verb, and an object. For each sentence, we included four versions which varied in main verb polarity (positive or negative) and the case marking of the direct object (accusative or genitive), yielding a total of 256 experimental items. Examples (7a–7d) represent all four conditions for one item in our dataset. Only the sentence in (7b) is ungrammatical: both (7a) and (7c) are grammatical because accusative objects are always licensed, and in (7d), the genitive of negation is grammatical because it is within the scope of a negated verb. In (7b), however, the genitive-marked object is not licensed by negation, which makes the whole sentence ungrammatical.

(7) a. **positive-accusative**

Vystavka artista poterpela proval
Exhibition of-artist suffered failure.ACC

"The artist's exhibition was a failure."

b. **positive-genitive**

* Vystavka artista poterpela
Exhibition of-artist suffered
provala
failure.GEN

"The artist's exhibition was a failure."

c. **negative-accusative**

Vystavka artista **ne** poterpela
Exhibition of-artist **neg** suffered
proval
failure.ACC

"The artist's exhibition wasn't a failure."

d. **negative-genitive**

Vystavka artista **ne** poterpela
Exhibition of-artist **neg** suffered
provala
failure.GEN

"The artist's exhibition wasn't a failure."

Given this pattern, we expected that the model would assign higher surprisal to the word *provala* 'failure.GEN' in (7b) than in (7d), but there would be no such difference for the word *proval* 'failure.ACC' in (7a) and (7c).

5.1.2 Results

LSTM Consistent with our predictions, the genitive-marked direct objects were less surprising when the verb was negated (see Figure 2a). Figure 3a shows that the difference between the positive and negative conditions is much bigger for genitive-marked objects than for the accusative-marked ones. This suggests the model learned that the negative-polarity constraint only applies to objects marked by the genitive case.

We further tested this by running a linear mixed effects model (Baayen et al., 2008) with the model-assigned surprisal as the dependent variable, and case, polarity, their interaction, and item frequency as predictors. We found a main effect of case ($p = 0.004$), as well as an interaction between case and polarity ($p < 0.0001$). Surprisal was significantly affected by polarity for genitive-marked objects ($p < 0.0001$), but not for accusative objects ($p = 0.09$).

Although we did not find a main effect of frequency, we performed a follow-up analysis aimed to rule out the possibility that unigram frequency could be a confound for these results. Figure 1 shows that accusative-marked objects in our dataset had much higher unigram frequency in the training corpus than the genitive-marked objects. To test for the presence of the frequency effects, we re-ran the linear mixed effects analysis on surprisal scores that we normalized by subtracting the target word's log frequency from its surprisal score. The pattern remained the same: we found main effects of frequency ($p = 0.006$) and, as before, of case ($p = 0.004$), as well as an interaction between case and polarity ($p < 0.0001$).

N-gram We found a main effect of case ($p < 0.0001$) and frequency ($p = 0.001$), but not of polarity ($p = 0.7$). There was no interaction between case and polarity ($p = 0.8$). Figure 4b shows there was no difference between the positive and negative conditions for either case. We observed this pattern in all experiments we ran, unless otherwise stated.

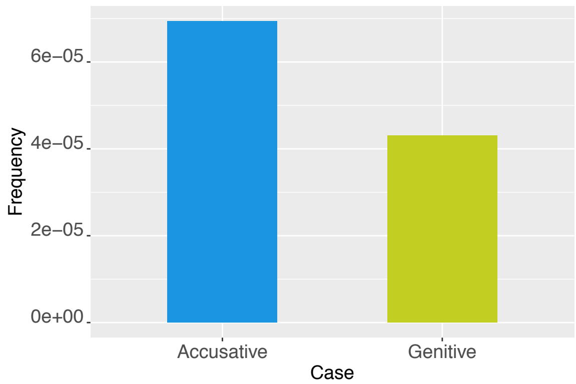


Figure 1. Average unigram frequency (word count divided by the size of the training corpus) of accusative and genitive objects from our dataset.

5.1.3 Discussion

Our results suggest that the model at least learned to encode case: to predict the grammaticality pattern in (7a–7d), the model needed to infer that the grammaticality of the genitive case – but not the accusative – is constrained by the presence of negation.

However, these results alone are not sufficient to conclude that the model was able to infer the syntactic structure that licenses the genitive of negation. Since our experimental items had SVO word order, it could have instead learned a linear rule where the genitive-marked object is allowed whenever it follows negation. Instead, the locality constraint would predict that the object in the genitive case is licensed only when it is in the scope of negation.

To test whether the model has learned the locality constraint, we ran a series of experiments in which we modified our experimental sentences to include the following distractors: (1) a negated relative clause, while the genitive-marked object was licensed by the negated main clause verb, (2) a complement clause, whose polarity varied between positive and negative, and whose main clause was always negative, and (3) a negated participial construction. We give a detailed description of these constructions in the following sections.

5.2 Experiment 2: Relative clauses

5.2.1 Materials

To test whether the model learned that the genitive of negation is only licensed under the scope of sentential negation, we modified the simple sentences from our dataset to include a relative clause with a negated verb. It is crucial for the model to infer the syntactic structure of these sentences: the model needs to be able to represent local scope in order to correctly predict that (8b) is ungrammatical – since the genitive-marked object in this case is outside the scope of negation.

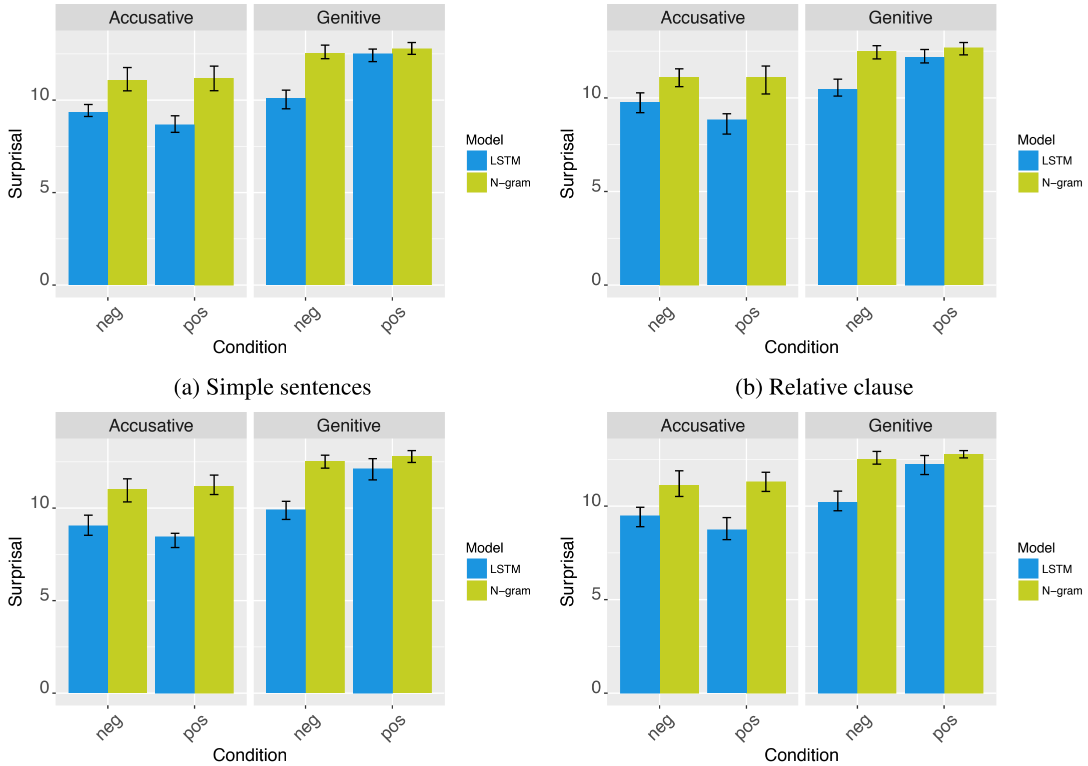
(8) a. * Vystavka artista, kotoryj **ne** lyubil
Exhibition of-artist who **neg** loved
vnimaniya publiki, poterpela provala
attention public suffered failure.GEN
"The exhibition of the artist, who didn't like public attention, was a failure."

b. Vystavka artista, kotoryj **ne** lyubil
Exhibition of-artist who **neg** loved
vnimaniya publiki, **ne** poterpela
attention public **neg** suffered
proval
failure.GEN
"The exhibition of the artist, who didn't like public attention, was not a failure."

5.2.2 Results

LSTM The model's surprisal was highest in the positive-genitive condition (Figure 2b), suggesting that genitive-marked direct objects were more expected when they were licensed by the negated main clause verb. We found main effects of case ($p = 0.01$) and polarity ($p = 0.04$), and the two terms interacted ($p < 0.0001$). Polarity significantly affected both genitive-marked ($p = 0.0001$) and accusative-marked ($p = 0.04$) objects. Figure 3b shows that for the accusative-marked objects, the difference between positive and negative conditions was the inverse of the genitive case: an accusative-marked object was more surprising when the main clause verb was negated.

The analysis of frequency effects revealed that normalized surprisal scores were significantly af-



(a) Simple sentences

(b) Relative clause

(c) Complement clause

(d) Participial construction

Figure 2. Surprisal averaged by condition (Experiments 1–4). Error bars indicate 95% confidence intervals.

fected by case ($p = 0.01$), frequency ($p = 0.001$), and the interaction of case and polarity ($p < 0.0001$).

N-gram The trigram model's performance was the same as in Experiment 1.

5.2.3 Discussion

Our results suggest the model learned the genitive-marked object was licensed only when it appeared in the scope of negation – which in turn required the representation of syntactic structure. If the model had learned only the linear rule, it would have assigned the same surprisal in both positive-genitive and negative-genitive conditions, since both linearly followed the negation in the scope of the relative clause.

The main effect of polarity suggests that the model possibly learned an interaction between case and polarity, preferring accusative objects with affirmative sentences and genitive objects under negation.

5.3 Experiment 3: Complement clauses

5.3.1 Materials

In the previous experiment, the distractor (i.e. the negation term that needed to be ignored) was always in the relative clause. This implies that there are two possible interpretations of the results: 1) the model could represent the scope of negation and apply it to the genitive licensing rule, or 2) the model learned to ignore negation if it immediately followed the word *kotoryj* 'that/who', which marked the beginning of an embedded clause. To rule out the second possibility, we tested the model's performance on sentences with complement clauses. In this set of sentences, the distractor was in the main clause, while the target word (the accusative- or genitive-marked direct object) was in an embedded clause. The embedded clause varied between positive and negative polarity – and only the latter licensed the genitive object:

(9) a. * Zhurnalist **ne** znal chto vystavka
Journalist **neg** knew that exhibition
artista poterpela provala
of-artist suffered failure.GEN
"The journalist didn't know that the artist's exhibition was a failure."

b. Zhurnalist **ne** znal chto vystavka
Journalist **neg** knew that exhibition
artista **ne** poterpela provala
of-artist **neg** suffered failure.GEN

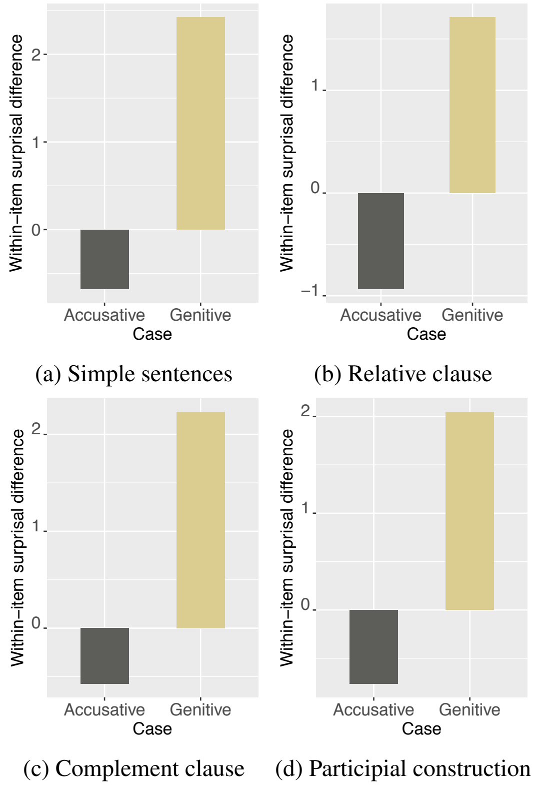


Figure 3. Within-item difference between positive and negative conditions, averaged by case (Experiments 1–4).

"The journalist didn't know that the artist's exhibition was not a failure."

5.3.2 Results

LSTM Average surprisal was lower for genitive-marked objects when the embedded clause contained a negated verb (Figure 2c), suggesting the model learned to represent sentential scope and did not mistake main clause negation for a licensor. Average within-item difference between positive and negative conditions was also greater for the genitive case (Figure 3c).

As before, we ran a linear mixed effects model to test the significance of these findings. We found a main effect of case ($p = 0.0006$), as well as an interaction between case and polarity ($p < 0.0001$). The surprisal the language model assigned to genitive-marked objects was significantly affected by the embedded clause's polarity ($p < 0.0001$), while there was no such effect for the accusative case ($p = 0.17$).

Our analyses of surprisal scores normalized by frequency revealed main effects of case ($p = 0.0004$) and frequency (0.002), as well as an interaction between case and polarity ($p < 0.0001$).

N-gram The model's performance was the same as in Experiment 1.

5.3.3 Discussion

These results provide further evidence that the model learned the locality constraint on genitive licensing: although the main clause verb was negated in all four conditions, the surprisal the model assigned to the genitive-marked object was reduced when the verb in the embedded clause was negated as well.

5.4 Experiment 4: Participial constructions

5.4.1 Materials

Experiments 2 and 3 provide some evidence that the model learned the scope constraint on the genitive of negation. However, the sentences we tested in these experiments contained overt cues that indicated the scope of negation that the model needed to ignore: in Experiment 1, the relative pronoun *kotoryj* indicates the beginning of the relative clause, and in Experiment 2, the pronoun *chto* indicates the beginning of the complement clause. Would the model be able to identify the scope of negation without these cues? We investigated this by testing the model's performance on the Russian participial construction, which has no overt function words marking the scope of negation. We constructed an experimental set of sentences which consisted of simple sentences such as those in (7a-7d) with an active present or past participle modifying the subject.

(10) a. * **Ne** poluchivshaya vnimaniya pressy
neg received.PTCP attention of-press
vystavka artista poterpela provala
exhibition of-artist suffered failure.GEN
"The artist's exhibition, which did not receive attention from press, was a failure."

b. **Ne** poluchivshaya vnimaniya pressy
neg received.PTCP attention of-press
vystavka artista **ne** poterpela
exhibition of-artist **neg** suffered
provala
failure.GEN
"The artist's exhibition, which did not receive attention from press, was not a failure."

In (10a), the genitive-marked object *provala* 'failure' is outside of the scope of negation, so we expected that it would be more surprising than in

(10b), where the genitive is licensed by sentential scope.

5.4.2 Results

LSTM Figure (2d) shows the model assigned higher probability to genitive-marked objects when they were licensed by a negated verb. A linear mixed effects analysis confirmed surprisal was affected by case ($p = 0.01$), as well as the interaction between case and polarity ($p < 0.0001$). Polarity was significant for genitive-marked objects ($p < 0.0001$), but not for accusative-marked ones ($p = 0.098$).

Surprisal scores normalized by frequency were significantly affected by case ($p = 0.01$), frequency ($p = 0.003$), and the interaction between polarity and case ($p < 0.0001$).

N-gram The model's performance was the same as in Experiment 1.

5.4.3 Discussion

The model was able to capture the grammaticality pattern in (10a–10b) despite the lack of overt scope marking cues – suggesting that the model in fact represents the scope of negation instead of relying on cues such as function words introducing embedded clauses.

5.5 Experiment 5: Existential copula construction

5.5.1 Materials

In the experiments we have presented so far, the genitive case was always optional: genitive-marked direct objects were only grammatical in the scope of sentential negation, while the accusative case was licensed whether the sentence had positive or negative polarity. We expected to see higher surprisal for genitive-marked objects when they were outside of the scope of negation, but we did not expect any polarity-related difference for the accusative case.

The situation is different in the Russian existential copula construction. First, in this construction the case alternation concerns the subject, which can be assigned the nominative or the genitive case. Second, the genitive case is always obligatory under negation. Finally, the nominative case marking is also constrained (unlike the accusative with direct objects): subjects can only receive nominative case when the sentence is affirmative. In other words, although in previous examples only the positive genitive condition was ungrammatical, in the case of the existential construction the negative nominative condition is ungrammatical as well:

(11) a. U vystavki byl proval
At exhibition was failure.NOM
"The exhibition was a failure."

b. * U vystavki byl provala
At exhibition was failure.GEN
"The exhibition was a failure."

c. * U vystavki **ne** bylo proval
At exhibition **neg** was failure.NOM
"The exhibition was not a failure."

d. U vystavki **ne** bylo provala
At exhibition **neg** was failure.GEN
"The exhibition was not a failure."

5.5.2 Results

LSTM A linear mixed-effects analysis revealed main effects of polarity ($p < 0.0001$), case ($p < 0.0001$), and frequency ($p = 0.0003$). The interaction between case and polarity was significant as well ($p < 0.0001$).

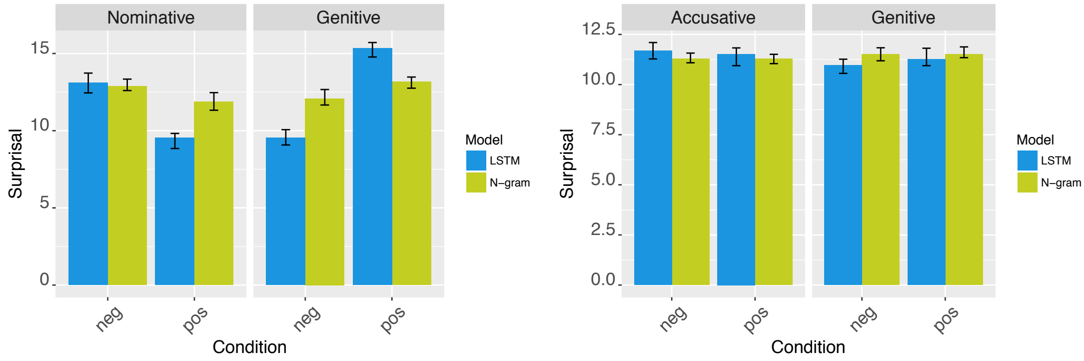
N-gram We found main effects of polarity ($p = 0.001$), case ($p = 0.0007$), and frequency ($p < 0.0001$). There was also a significant interaction of case and polarity ($p < 0.0001$).

5.5.3 Discussion

The main effect of polarity shows that the model learned constraints on both the nominative and the genitive case: the genitive is licensed under negation and ungrammatical in affirmative sentences, while the opposite is true for the nominative.

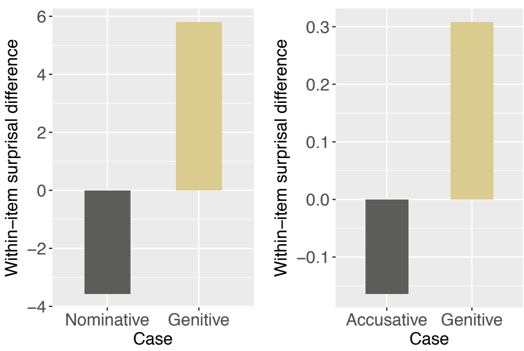
Further, within-item difference for both the nominative and the genitive is much bigger than in other experiments (Figure 5a) – which suggests that the model distinguished between optionality and obligatoriness. I.e., the magnitude of surprisal was reduced in the positive-genitive condition when it was optional under negation. However, when it was required under negation, genitive-marking with positive polarity was more surprising.

Compared to previous experiments, there was a stark difference in surprisal scores between positive and negative conditions. This could be due to the fact that the the verb *byt'* 'to be' always appears in 3rd person singular under negation, which could have provided the model with an additional cue that the genitive case is required.



(a) Existential construction (b) Parenthetical

Figure 4. Surprisal averaged by condition (Experiments 5–6). Error bars indicate 95% confidence intervals.



(a) Existential construction (b) Parenthetical

Figure 5. Within-item difference between positive and negative conditions, averaged by case (Experiments 5–6).

5.6 Experiment 6

5.6.1 Materials

In the grammatical sentences used in Experiments 1–5, the genitive objects were directly preceded by the **neg** + main verb bigram, which left open the possibility that the LSTM model relied on this linear structure as a cue that the genitive case was licensed. We constructed a new dataset where the main verb was separated from the direct object by a parenthetical (e.g. "to the surprise of the press" in 12a-12b). If the model is learning the locality rule correctly, this parenthetical should not intervene with inferring the grammaticality pattern in 12a-12b.

(12) a. * Vystavka artista poterpela, k
Exhibition of-artist suffered to
udivleniju pressy, provala
surprise of-press failure.GEN
"The artist's exhibition was a failure, to the surprise of the press."

b. Vystavka artista **ne** poterpela, k
Exhibition of-artist **neg** suffered to
udivleniju pressy, provala
surprise of-press failure.GEN
"The artist's exhibition wasn't a failure, to the surprise of the press."

5.6.2 Results

LSTM We found a main effect of case ($p < 0.0004$) and frequency ($p = 0.01$), but not of polarity ($p = 0.6$); there was no interaction between case and polarity ($p = 0.1$). Figure 4b shows there was almost no difference in surprisal the model assigned to the genitive objects licensed by negation compared to those that were ungrammatical.

N-gram There was a main effect of frequency ($p < 0.0001$), but not of case ($p = 0.34$) or polarity ($p = 0.96$). There was no interaction between case and polarity (0.97).

5.6.3 Discussion

In (12b), the negation term was local to the target genitive object, but linearly separated from it. If the model was correctly learning the locality constraint, it would be able to predict that the genitive object *provala* is grammatical in (12a), but not (12b). However, the model could not identify the negation term as the licensor in these types of sentences, assigning similar surprisal to the genitive objects in (12a) and (12b). This result, however, may be due to the rarity of the parenthetical sentences in the training corpus, and does not necessarily imply the model was not learning the constraint in Experiments 1–5.

6 General discussion and future work

In this paper, we have examined the ability of an RNN language model to learn several properties of the Russian genitive of negation. The genitive of negation can **optionally** mark direct objects of transitive verbs when the latter are negated, and is **obligatory** with subjects of existential copula constructions under negation.

To be able to learn the polarity constraint on the genitive case, the model needed to represent the scope of negation. In Experiments 2 and 3, we tested this by introducing distractors to our experimental items: negated relative clauses and complement clauses that were not licensed by sentential negation. We found that the model's performance matched our predictions, assigning higher surprisal to those genitive-marked objects that were outside of the scope of negation. The results from Experiment 4 further suggest that the model could represent the scope of negation without relying on such cues as function words explicitly marking clause boundaries.

Our results from Experiment 5 provide some evidence that the model could differentiate between optionality and obligatoriness. First, we found that both the nominative and the genitive case were significantly impacted by polarity (while only the genitive was affected in other types of sentences we tested). Second, for both the nominative and the genitive case the average within-item difference between positive and negative conditions was much bigger than in other experiments. Taken together, these results suggest that the model learned that the genitive of negation was obligatory in existential sentences.

The results of Experiment 6 reveal that the model could not learn the locality constraint on the genitive of negation when the linear distance between the main verb and the direct object was increased. We tested sentences where a parenthetical intervened before the main verb and its object, and the model did not differentiate between the sentences in which the genitive object was licensed by a local negation term from those where it was not. However, this finding does not necessarily imply that the model did not learn the locality constraint in Experiments 1–5. One possible explanation for the model's behavior on the task in Experiment 6 is that constructions where a parenthetical intervenes between the main verbs and its object are not frequent in a natural corpus.

Further, more evidence is needed to asses whether the model could differentiate between syntactic structures which optionally licensed the genitive case from those where it was obligatory. One limitation of our approach is that we used the same metric for both optional and obligatory uses of the genitive of negation: we compared the surprisal the model assigned to grammatical and ungrammatical sentences, and the negated sentences with the genitive case were grammatical whether the genitive was obligatory or optional. A possible direction for future work could involve a comparison of our results to human processing data (e.g. as in Futrell and Levy 2018). Since surprisal scores tend to correlate with reaction times (Smith and Levy, 2013), we would expect our results to match human performance.

Finally, our study only addressed some properties of the genitive of negation and only a subset of the syntactic structures in which it can appear. We haven't looked, for instance, into the genitive case marking of unaccusative subjects (13) and derived subjects of passives (14) (Bailyn, 1997):

(13) (Babby, 1980)

Zdes'	ne	rastet	gribov
here	**neg**	grows	mushrooms.GEN

"No mushrooms grow here."

(14) (Bailyn, 1997)

Ne	bylo	polucheno	gazet
neg	was	received	newspapers.GEN

"No newspapers were received."

There is also a slight difference in meaning between the genitive and accusative direct objects that we haven't addressed: while accusative direct objects usually receive a definite interpretation, the genitive ones have an existential or indefinite interpretation (Bailyn, 1997; Harves, 2002).

While future investigation into these issues is needed to gain a full picture of neural network learning of the genitive of negation, our study adds to the growing body of evidence that RNN language models do not need syntactic supervision or a hierarchical bias to capture syntactic dependencies. Whether the same is true for human language learners remains to be seen.

References

R Harald Baayen, Douglas J Davidson, and Douglas M Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, 59(4):390–412.

Leonard H Babby. 1980. *Existential Sentences and Negation in Russian*. Karoma Publishers: Ann Arbor, MI.

Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2014. Neural machine translation by jointly learning to align and translate. In *International Conference for Learning Representations*.

John F Bailyn. 1997. Genitive of negation is obligatory. In *Formal approaches to slavic linguistics*, volume 4, pages 84–114. University of Michigan Press: Ann Arbor, MI.

Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2014. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Noam Chomsky. 1980. Rules and representations. *Behavioral and brain sciences*, 3(1):1–15.

Richard Futrell and Roger P Levy. 2018. Do RNNs learn human-like abstract word order preferences? *arXiv preprint arXiv:1811.01866*.

Kristina Gulordava, Piotr Bojanowski, Edouard Grave, Tal Linzen, and Marco Baroni. 2018. Colorless green recurrent networks dream hierarchically. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1195–1205. Association for Computational Linguistics.

John Hale. 2001. A probabilistic earley parser as a psycholinguistic model. In *Proceedings of the second meeting of the North American Chapter of the Association for Computational Linguistics on Language technologies*, pages 1–8. Association for Computational Linguistics.

Stephanie Harves. 2002. Genitive of negation and the syntax of scope. In *Proceedings of ConSOLE*, volume 9, pages 96–110.

Jaap Jumelet and Dieuwke Hupkes. 2018. Do language models understand anything? On the ability of LSTMs to understand negative polarity items. In *Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, pages 222–231, Brussels, Belgium. Association for Computational Linguistics.

Olga Kagan. 2010. Genitive objects, existence and individuation. *Russian linguistics*, 34(1):17–39.

Nina Kazanina. 2017. Predicting complex syntactic structure in real time: Processing of negative sentences in Russian. *The Quarterly Journal of Experimental Psychology*, 70(11):2200–2218.

Eliyahu Kiperwasser and Yoav Goldberg. 2016. Simple and accurate dependency parsing using bidirectional lstm feature representations. *Transactions of the Association for Computational Linguistics*, 4:313–327.

Tal Linzen, Emmanuel Dupoux, and Yoav Goldberg. 2016. Assessing the ability of LSTMs to learn syntax-sensitive dependencies. *Transactions of the Association for Computational Linguistics*, 4:521–535.

Rebecca Marvin and Tal Linzen. 2018. Targeted syntactic evaluation of language models. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP 2018*, pages 1192–1202.

R. Thomas McCoy, Robert Frank, and Tal Linzen. 2018. Revisiting the poverty of the stimulus: Hierarchical generalization without a hierarchical bias in recurrent neural networks. In *Proceedings of the 40th Annual Conference of the Cognitive Science Society*, pages 2093—2098, Austin, TX.

Tomáš Mikolov, Martin Karafiát, Lukáš Burget, Jan Černockỳ, and Sanjeev Khudanpur. 2010. Recurrent neural network based language model. In *Eleventh annual conference of the international speech communication association*.

Elena V Paducheva. 2004. The genitive subject of the verb byt' (to be). *Studies in Polish linguistics*, 1:47–59.

David Pesetsky. 1982. *Paths and categories*. Ph.D. thesis, MIT.

Shauli Ravfogel, Yoav Goldberg, and Francis Tyers. 2018. Can LSTM learn to capture agreement? The case of Basque. In *Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, pages 98–107, Brussels, Belgium. Association for Computational Linguistics.

Nathaniel J Smith and Roger Levy. 2013. The effect of word predictability on reading time is logarithmic. *Cognition*, 128(3):302–319.

Alan Timberlake. 1975. Hierarchies in the genitive of negation. *The Slavic and East European Journal*, 19(2):123–138.

Oriol Vinyals, Łukasz Kaiser, Terry Koo, Slav Petrov, Ilya Sutskever, and Geoffrey Hinton. 2015. Grammar as a foreign language. In *Advances in neural information processing systems*, pages 2773–2781.

Ethan Wilcox, Roger Levy, Takashi Morita, and Richard Futrell. 2018. What do RNN language models learn about filler-gap dependencies? In *Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, pages 211–221, Brussels, Belgium. Association for Computational Linguistics.