

# Quantifying the cross-linguistics effects of syncretism on agreement attraction

Utku Turk (utkuturk@umd.edu)

Department of Linguistics, University of Maryland, College Park

## Abstract

Speakers' attention to distinctive case markers varies depending on the language they speak. Speakers of languages like German, Russian, or English less often deem ungrammatical sentences grammatical when the case marking is distinctive. However, other languages like Turkish, French, or Armenian do not show this sensitivity, namely additional morphological information provided does not help speakers to create a better representation of the sentence. One explanation for this dichotomy is related to functional utility of morphology in a language. In this study, I first review the psycholinguistic literature on speakers' sensitivity to morphological cues during agreement processing in Turkish, Russian, and English. I then aim to quantify the functional utility of these morphological cues using large language models and their attention patterns.

**Keywords:** syncretism; agreement attraction; attention; large language models

## 1 Introduction

Consider the following sentences in (1).

- (1) a. \* The key to the cabinet are rusty.  
b. \* The key to the cabinets are rusty.

Even though both sentences are ungrammatical due to agreement related reasons, i.e. using 'are' instead of 'is', psycholinguistic research have shown that speakers are more likely to deem sentences like (1b) as grammatical or produce them when the initial preamble includes a plural noun compared to a singular noun (Wagers, Lau, & Phillips, 2009).

One theory that explains these systematic errors known as agreement attraction is the cue-based retrieval account (Wagers et al., 2009). Following the ACT-R model of sentence processing, this model assumes that words are stored in a content-addressable memory as a chunk that can be later retrieved using specific cues (Lewis & Vasishth, 2005). Recall the example in (1b). According to this theory, the noun key would be stored as a following chunk in which the number, the overt case marking, and the dominating syntactic position are encoded: [SG, NOM, TP].<sup>1</sup> Similarly, the word cabinets would be stored as [PL, NOM, PP].

In the account put forward by Wagers et al. (2009), readers would form a prediction based on the chunks they have processed first. In this case, their prediction would be that the agreement carrying element should be singular, given that the noun resides in the syntactic position of an agreement controller [TP] is singular. In cases where participants see 'is' or

other elements that are singular, no disruption in reading or judgment is seen. On the other hand, if the element they see is 'are' or other elements that do not match with the predicted number, a repair process is initiated. In this process, participants would use the relevant cues provided by 'are,' such as [PL, NOM, TP], to search for an agreement controller. In (1b), there is no single match, thus most of the time participants do not find these sentences grammatical. However, systematically but occasionally, participants may erroneously retrieve *cabinets* [PL, NOM, PP] as the controller of the agreement due to a partial match between the cues provided by 'are.'

On the other hand, this partial match would be a lot less likely to give rise to erroneous retrieval of the word *cabinet* in (1a), since the features shared between the cues of 'are' [PL, NOM, TP] and the chunk of 'cabinet' [SG, NOM, PP] are even fewer.

### 1.1 The role of syncretism

One crucial assumption shared among the cue-based retrieval models is the role of case marking and morphology (Lewis & Vasishth, 2005). In every iteration of the cue-based retrieval theory, the string-based morphological cues are argued to be encoded. For example, in sentences like 'I heard you,' even though the syntactic case assigned to 'you' is the accusative case, due to the morphological realization the chunk information would include nominative and accusative at the same time. This ambiguity in the morphological output is known to be syncretism. In non-syncretic elements like 'he/him,' the case is marked distinctively. Thus, in the sentence 'I heard him,' the chunk for the word 'him' would contain accusative, but not nominative.

The role of distinctive case marking is not specific to agreement processes. For example, Fedorenko, Babyonyshev, and Gibson (2004) found that participants were slower to read sentences in sentential embedding configurations like [NP-case NP-case VP VP] when the output of cases are syncretic. The presence of syncretic cases also decreased the overall accuracy of participants in comprehension questions.

### 1.2 The effect of syncretism in English Attraction

Given the role of attractor morphology, one prediction of the cue-based retrieval is that participants should not make similar errors when a possible attractor is marked distinctively. Nicol, Barss, and Barker (2016) tested this prediction with

<sup>1</sup>The selection of which features will be encoded is not exactly clear. The features we present here are not the exclusive list, but rather an exemplary set.

English possessive marked nouns. Consider the preambles in (2).

- (2) a. The statue in the elf's garden ...
- b. The statue in the elf's gardens ...
- c. The statue in the elves' garden ...
- d. The statue in the elves' gardens ...

They found that participants produced fewer agreement attraction errors with preambles like (2c) compared to (2b) or (2d). Moreover, there were no significant difference between (2b) and (2d). They took their results as an evidence for a confirmation of the use of overt morphological marking in the computation of agreement attraction.

In another experiment, Nicol and Antón-Méndez (2009) found that using pronouns as attractors as in '*The bill from them ...*' diminished the attraction effects compared to sentences with nouns as attractors whose case marking is syncretic between accusative and nominative. Findings of both Nicol and Antón-Méndez (2009) and (Nicol et al., 2016) follows from the predictions of the cue based retrieval models.

### 1.3 Cross-linguistic picture of syncretism

However, the effect of syncretism is not clear when one looks at various languages. While certain languages like Turkish or Armenian do not exhibit changes in attraction as a function of case syncretism, other languages like Russian aligns with English and shows increased attraction effects, i.e. more errors in agreement marking, when the forms of the nouns involved in agreement processing is ambiguous.

**Russian.** Slioussar (2018) tested syncretic attractors in Russian, which exhibits an interesting syncretism that is only present in nouns with feminine gender, as in '*statja*' (paper). The singular form of paper when it is marked with a genitive case '*stati*' is syncretic with the plural nominative marked paper '*stati*'. On the other hand, the plural genitive marked paper '*statej*' is distinctively marked and not ambiguous with the nominative plural '*stati*' or singular '*statja*'. She found that when participants were provided with preambles like '*material for paper<sub>GEN.SG</sub> (=NOM.PL)*', they made more errors compared to '*material for papers<sub>GEN.PL</sub> ( $\neq$ NOM.PL/NOM.SG)*'. This effect was robust in both comprehension and production. Her results suggest the string syncretism of case marking can induce the attraction even in cases where the number feature is not plural, which is a stronger version of the tests provided in German and English.

**Turkish.** Turkish along with Armenian and French differs from previous languages. Lago et al. (2019) tested agreement attraction in Turkish by using sentences as in (3). However, all the sentences used in their experiment included head nouns '*eğitmen*' (instructor) that ends with a consonant. The possessive marking in Turkish is syncretic with the accusative case when the noun ends with a consonant. Moreover, the genitive case marked attractor '*teknisyen*' (technician) can be interpreted as a subject of an embedded clause at the time of reading since Turkish embedded subjects are generally marked with a genitive case. Türk and Logačev (2024) tested

the role of syncretism in Turkish by minimally changing the the experimental items in (3). Instead of using consonant ending head nouns, they used vowel ending head nouns '*hoca*' (teacher) which disambiguated between the accusative case and the possessive marking as in (4). They expected that given how syncretism matter in other languages and additionally there is a structural ambiguity in Turkish, using distinctively marked nouns would diminish the attraction effects significantly. However, they found that using distinctive case marking did not affect the magnitude of attraction. Both sentences containing syncretic possessive marking *-i* and distinctive marking *-si* had comparable attraction effects.

- (3) Teknisyen-(ler)-in eğitmen-i ... koştı-(lar).  
technician-(PL)-GEN instructor-POSS ran-(PL)  
'The instructor of the technician(s) ... ran<sub>SG/PL</sub>.'
- (4) Teknisyen-(ler)-in hoca-sı ... koştı-(lar).  
technician-(PL)-GEN teacher-POSS ran-(PL)  
'The teacher of the technician(s) ... ran<sub>SG/PL</sub>.'

### 1.4 Quantifying Retrieval

Is there a way to quantify this retrieval processes and check what happens at each language? The ACT-R model provided by Lewis and Vasishth (2005) gives the mathematical formulation of such quantification, called *fan* effect. The fan effect is dependent on how many chunks are there and the number of shared features. As the features shared between the chunks increase, the fan effect also increases, and the activation of a chunk decreases. However, for this calculation to go through, one needs to know the which features are tracked. This creates the paradoxical situation, since our main question in this paper is to understand what is different in these languages. One can create a plethora of feature sets and try to explain differences via specific feature engineering.

Another possible way is to use an approximation. One symptom of the fan effect is the reading times. As the fan effect is increased, it is expected that participants should spend more time to resolve a cue-feature dependency. Recent findings in Deep Neural Network (DNN) suggest that these DNN models can provide quantified numbers, called surprisal values, that can predict humans' reading times in linguistic contexts from 11 different languages (Wilcox, Pimentel, Meister, Cotterell, & Levy, 2023). These 11 languages include English, Russian, and Turkish, which are the languages we test in this paper. If one assumes that the underlying model of human sentence comprehension includes a fan-effect-like mechanism as in cue-based retrieval, it is possible that we can use this approximate number from language models.

Ryu and Lewis (2021) recently tested this approximation in English agreement attraction scenarios. They showed that both surprisal values, i.e. how likely a word to appear in a context (Levy, 2008), and attention values, i.e. how much importance a specific word in the context carries (Clark, Khandelwal, Levy, & Manning, 2019), predicted the attraction effects in English. In this paper, we are hoping that these values

Table 1: Materials used in our simulations. A bold font is used for heads and the attractors are provided underlined.

Language	Source	Condition	Sentence
English	Nicol et al. (2016), Exp3	Non-Syncretic	The <b>statue</b> in the <u>elf's/elves'</u> garden is/*are ...
English	Wagers et al. (2009), Exp4-7	Syncretic	The <b>key</b> to the <u>cabinet/cabinets</u> is/*are ...
Russian	Slioussar (2018), Exp3	Non-Syncretic	<b>Korobka</b> dlya <u>kraski/krasok</u> byla/*byli ... a box for the paint(s) was/*were ...
Russian	Slioussar (2018), Exp3	Syncretic	<b>Tropinka</b> cherez lug/luga byla/*byli ... the path through the meadow(s) was/*were ...
Turkish	Türk and Logačev (2024)	Non-Syncretic	Teknisyennin/Teknisyenlerin <b>hocası</b> durmadan koştu/*koştular. The teacher of the technician(s) ran <sub>sg</sub> /*ran <sub>pl</sub> non-stop.
Turkish	Lago et al. (2019)	Syncretic	Teknisyenin/Teknisyenlerin <b>eğitmeni</b> durmadan koştu/*koştular. The instructor of the technician(s) ran <sub>sg</sub> /*ran <sub>pl</sub> non-stop.

will pattern with presence of absence of attraction effects as a function of case syncretism.

### 1.5 Reconciling cross-linguistic facts

To the best of our knowledge, there is no principled way to explain cross-linguistic facts. One language-specific explanation provided by Avetisyan, Lago, and Vasishth (2020) is that speakers only use case related features for thematic assignment, and they are not used to decide number marking. However, this explanation would fail to explain languages like English, German, or Russian. Another explanation is provided by Dillon and Keshev (2024). They propose that one way to reconcile these opposite cross-linguistic facts is to attribute these differences to language-specific cue dynamics. They raise the possibility that the language-specific utility of morphological cues may underlie the observed effects of syncretism. However, this remains a theoretical proposal, as there is currently no empirical evidence supporting this view. In this work, I am aiming to find evidence for this proposal. If the reason for cross-linguistic differences are due to encoding utility differences, these encoding utility differences might also reveal themselves in surprisal and attention values in LLMs. One would hope that with the amount of data these models are fed, they would be able to pick up distributional facts with respect to case, case syncretism, and the probability of how important morphological information is in controlling dependencies.

## 2 Experimental Setup

We used the following pretrained models: pre-trained GPT2-small (Radford et al., 2019) and BERT (Devlin, Chang, Lee, & Toutanova, 2018) for English, Turkish-GPT2-small (Kesgin et al., 2024) and BERT (<https://huggingface.co/dbmdz/bert-base-turkish-cased>) for Turkish, and BERT for Russian (<https://huggingface.co/deepvk/bert-base-uncased>). All models were accessed through the transformers or minicons library, and used with the Huggingface API.

### 2.1 Materials

We used data from Türk and Logačev (2024) and Lago et al. (2019) for Turkish, Nicol et al. (2016) and Wagers et al. (2009) for English, and Slioussar (2018) for Russian. Table 1 shows the set of experimental sentences as well as their translations. These three sets of sentences have a 2 x 2 x 2 experimental manipulation with the factors grammaticality (grammatical/ungrammatical), number mismatch with the head and the attractor (match/non-match), and case marking (distinctive, non-syncretic/non-distinctive, syncretic) While syncretism (non-distinct case marking) is shown as a different line, other manipulations such as the attractor number and the verb number is shown within the example with slashes.

### 2.2 Metrics

In this paper, we have used two measures from LLMs: attention to possible agreement controllers and surprisal following (Ryu & Lewis, 2021).

**Attention.** We extracted the soft attention weights for each relevant noun that could either correctly (grammatical head) or erroneously (attractor) control agreement. We only checked the layer-head combination that has the most likelihood of predicting the subject head.

**Surprisal.** We estimated surprisal for the agreement bearing words (auxiliary or main verbs) using two family of models. For GPT-2, we used its left-to-right generation to obtain token-wise surprisal directly. For BERT, which does not predict sequentially, we used masked language modeling: we masked the auxiliary and computed its negative log probability given the surrounding context.

### 2.3 Probing heads

In order to decide which heads to track for agreement related dependencies, we used Universal Dependencies treebanks for English (35,581 total sentences from EWT, GUM, and LinES Treebanks), Turkish (36,760 total sentences from BOUN, Atis, GB, and Kenet Treebanks), and Russian (209,304 total sentences from Taiga and SynTagRus Treebanks). After

extracting `nsubj` dependencies that control dependencies for each sentence, we obtained which attention heads attributed most attention to the element that was tagged with the `nsubj`. We only used sentences where the subject precedes the verb. Following is the list of layers and attention heads as well as the percentage of these heads paying the highest attention to the subject: English: GPT2: layer4-head11 (36% accuracy), BERT: layer8-head10 (12% accuracy), Russian: BERT: layer1-head5 (32% accuracy), and Turkish: GPT2: layer3-head0 (17% accuracy), BERT: layer4-head0 (26% accuracy).

## 2.4 Behavioral Results and Predictions

Our main question is whether an LLM-based approximation for the effects of case-syncretism in agreement attraction is possible. These attraction effects are mainly discussed as differences between plural attractor and singular attractor conditions in ungrammatical sentences (*the key to the cabinets vs. cabinet are*). For this reason, we are going to focus on ungrammatical sentences in our results sections while plots will include values for grammatical sentences too as a control.

Firstly, we should see a difference between the measures as a function of attractor number in Turkish and English only in ungrammatical sentences. In our plots, we will plot the difference ( $\Delta$ ) between Match vs. Mismatch. Since all head nouns are singular, match means a singular attractor, and mismatch means a plural attractor. Assuming that attention to noun phrases in a large language models may reflect the probability of that noun being retrieved as the controller,  $\Delta\text{Head-attention}$  for these languages should be positive and  $\Delta\text{Attractor-attention}$  should be negative, indicating a reduced attention to head but increased attention to attractor in mismatch condition. This differences should get closer to 0 in English, but not change in Turkish as a function of syncretism.

On the other hand, non-syncretic Russian examples should pattern like English, yet, syncretic examples should have a different picture. In syncretic conditions, both match and mismatch condition should have the same sign for the difference: negative for  $\Delta\text{Head}$  and positive for  $\Delta\text{Attractor}$ . These differences should be more amplified for match conditions.

Secondly, assuming surprisal values reflect reading times as a function of how much features are overlapping between candidates, we expect to see lower surprisal values for mismatch conditions in English and Turkish ungrammatical sentences. This effect should only be visible in non-syretic conditions in Russian. While there is no consistent findings with respect to reading times, the cue-based retrieval account would expect non-syncretic cases to be processed slower, thus increased surprisal values.

## 2.5 Simulation Results

In the next section, we refer to our plotted descriptive results for each metric and report posteriors' probability of being lower than zero from our maximal

Bayesian models fitted to these metrics assuming a log-normal distribution. Our code and data can be found in: [https://github.com/utkuturk/syncretism\\_crossling\\_llm](https://github.com/utkuturk/syncretism_crossling_llm)

**Attention to head.** Figure 1 shows the attention difference as a function of number mismatch to head metric across 2 models and 3 languages for all 4 conditions: verb number and syncretism.

In singular verb conditions, the attention to the head does not change as a function of head-attractor number relation ( $P(\beta_{Eng} < 0) = 0.53$ ,  $P(\beta_{Ru} < 0) = 0.44$ ,  $P(\beta_{Tr} < 0) = 0.59$ ), independent of the model type, language, and syncretism. No interactions involving match were significant (all  $P(\beta < 0) > 0.11$ ). This is expected given the general attraction results cross-linguistically. In grammatical sentences (singular verb-singular head), attraction results are rarely observed.

In plural verb conditions, the picture is slightly different. First of all, our expectation was such that there should be a difference between the match and mismatch conditions across languages. However, our models revealed a significant main effect of mismatch only in English ( $P(\beta < 0) > 0.99$ ), but not in Russian ( $P(\beta < 0) = 0.71$ ) or Turkish ( $P(\beta < 0) = 0.40$ ). Recall that null main effect is expected in Russian given that both match and mismatch condition induce attraction.

As for our main question, we predicted that mismatch effect should be visible in English syncretic cases. Our model verified this negative interaction: a reliable mismatch effect was found independent of model type ( $P(\beta < 0) = 0.09$ ); however, this interaction was amplified in the BERT model ( $P(\beta < 0) = 0.07$ ). Turkish data showed no evidence for an interaction between match and syncretism ( $P(\beta < 0) = 0.44$ ), which is expected. On the other hand, Russian showed a strong positive evidence for an interaction between match and syncretism ( $P(\beta < 0) = 0.98$ ). This is expected given that syncretic match conditions are found to cause more attraction.

Overall, we were able to predict English and Russian attraction results and syncretism data. However, we failed to replicate number mismatch difference in Turkish. Even though there were no difference due to syncretism, there were also no baseline attraction-pattern.

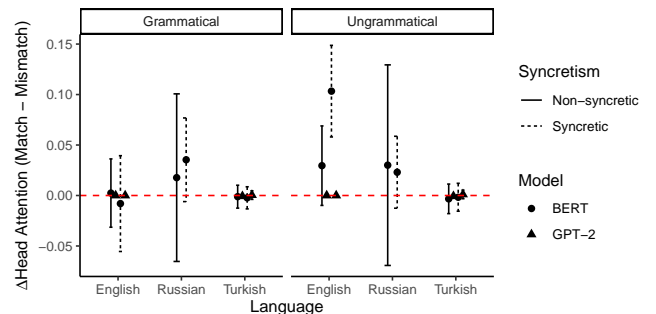


Figure 1: Head Attention Difference for each language from GPT-2small and BERT models. Error bars signal SE.

**Attention to Attractor.** Figure 2 shows the attention difference to attractor metric across the same conditions. One

possibility is that even though attention to the head did not change, overall attention to the attractor was funneled from other elements in the sentence and thus was increased. For a cue-based model this would indicate that parser consider non-nominal chunks too.

As predicted, singular verb (grammatical) conditions do not show any effect as a function of number mismatch ( $P(\beta_{Eng} < 0) = 0.54$ ,  $P(\beta_{Ru} < 0) = 0.79$ ,  $P(\beta_{Tr} < 0) = 0.29$ ). One outlier is Turkish with a BERT language model. There seems to be a more attention allocated to the attractor when both of them are singular ( $P(\beta < 0) < 0.001$ ), as opposed to the mismatch condition where the head is singular and the attractor is plural.

In the plural verb conditions, BERT models for English and Russian results show a pattern similar to agreement attraction, that is, there is an increased attention to attractors in mismatch conditions (*The key to the cabinets are ...*). Our model verified this effect ( $P(\beta_{Eng} < 0) > 0.001$ ,  $P(\beta_{Ru} > 0) = 0.01$ ). In our plot, this is visible as a negative  $\Delta$ . Turkish on the other hand showed comparable results to grammatical conditions.

For our syncretism question, syncretism overall increases the attention to attractor in English ( $P(\beta < 0) < 0.99$ ). More importantly, this attention increase was amplified with number mismatch interaction ( $P(\beta < 0) < 0.99$ ). On the other hand, Russian BERT model did not show such an interaction ( $P(\beta < 0) = 0.34$ ). This is unexpected mainly because we expected reversed sign for the difference between Match and Mismatch conditions as a function of syncretism.

Overall, we were able to predict English and Russian attraction data. However, the syncretism related findings were only replicated in English. As for Turkish, even though basic attraction patterns were not found, the insensitivity to morphological form was verified.

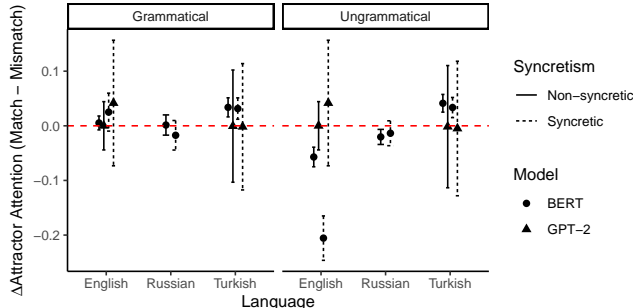


Figure 2: Attractor attention difference for each language from GPT-2small and BERT models. Error bars signal SE.

**Surprisal** Figure 3 shows the surprisal difference computed at the auxiliary verbs as a function of number mismatch. It is important to note that even though we have reading data for the Russian data and the syncretic English data, we do not have the reading time data for Turkish or non-syncretic English dataset. However, we will compare our simulation findings to our expectations with minimal additional assumptions.

Masking-based BERT results shows no effect of number mismatch between the head and the attractor ( $P(\beta < 0) =$

0.50) and no interaction with the grammaticality ( $P(\beta < 0) = 0.63$ ), an important pattern observed in attraction results. Even though there seems to be a difference in surprisal values as a function of syncretism ( $P(\beta < 0) < 0.001$ ), there is also no interaction between the syncretism and the number mismatch ( $P(\beta < 0) = 0.48$ ).

On the other hand, GPT-2small results reveals an interesting pattern. Firstly, there is no difference as a function of number-mismatch when the verb is singular ( $P(\beta < 0) = 0.55$ ), which is expected given the expectation cue-based model laid on in Wagers et al. (2009), meanwhile the vanilla cue-based model would expect increased reading times and thus surprisal with match conditions (singular head-singular attractor-singular verb) due to competition of the cues compared to mismatch conditions.

For the English dataset, we see that surprisal values decreased for the mismatch conditions in non-syncretic nouns when the verb was plural ( $P(\beta < 0) > 0.99$ ). For syncretic nouns, surprisal increased slightly in mismatch conditions, signalled with a significant interaction between mismatch and syncretism ( $P(\beta < 0) < 0.001$ ). This aligns with our expectations, since lower surprisal was previously taken to signal decreased reading times. Moreover, previous behavioral results also suggest that reading times decrease in sentences like *the key to the cabinets are ...* compared to *the key to the cabinet are ....* More importantly, we found that the magnitude of the difference is greater in syncretic conditions as opposed to non-syncretic conditions. This difference follows from the predictions of cue-based retrieval model. Considering that reading times increase with more features shared between the competitors, it is expected to see lower surprisal in cases with less feature overlap, i.e. no-syncretism cases. Moreover, the fact that there is an interaction between the syncretism and number match is a good sign for the predictability of human behavior with surprisal approximation.

As for Turkish, we again see the effect of number mismatch in the mean surprisal values. As predicted, we see lower surprisal values in mismatch conditions (singular head-plural attractor-plural verb) compared to match conditions (singular head-singular attractor-plural verb) ( $P(\beta < 0) < 0.001$ ). More importantly, the syncretism does not have an effect on the difference between match and mismatch conditions ( $P(\beta < 0) = 0.51$ ). This is in line with the behavioral findings that shows syncretism does not effect the attraction results. However, there is no direct evidence on reading times in Turkish.

### 3 Conclusion

This paper aimed to quantify, not explain, cross-linguistic facts of morphological form of distinctive case marking using measures from large language models. We specifically tested the agreement attraction sentences. Dillon and Keshev (2024) suggest that cross-linguistic differences might be due to the functional informativity of morphological marking in a specific language. Following the link proposed by Ryu and



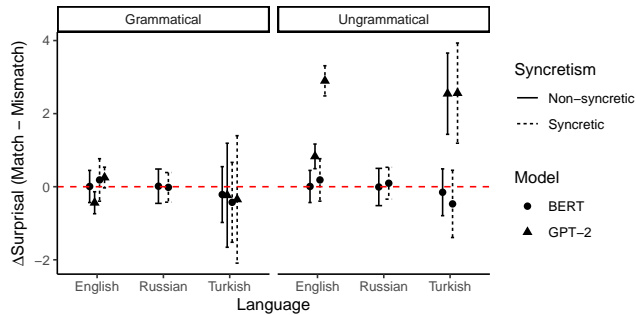


Figure 3: Surprisal values for each model computed at the auxiliary verb. Error bars signal SE.

Lewis (2021), we argued that attention values are linked to likelihood of memory retrieval which is based on the cues utilized, while surprisal values are related to reading times which again depends on cue dynamics.

We were able to predict the behavioral findings in English in both surprisal values and attention values. We were also able to predict morphological insensitivity in Turkish, even though the attraction patterns in attention measure were not conclusive. Finally, we were able to predict Russian attraction findings, but not the effect of morphological manipulation with none of the models.

One of the reasons for this might be due to the very low attention head accuracy. We only used hand-annotated data for head-probing. A larger data set would prove more useful in increasing the accuracy.

Moreover, we assume that there is one head at play. However, many-to-many relation is also possible given that it is unlikely case is related to syntactic position, choice of preposition, and other semantic features. Moreover, case and number representations are also independent.

Lastly, we leave exploration of this many-to-many relation and explanation of why such a cross-linguistic difference exist in the first place to future studies. We believe that a Counterfactual Intervention method proposed by Hao and Linzen (2023) would prove useful.

## References

- Avetisyan, S., Lago, S., & Vasishth, S. (2020). Does case marking affect agreement attraction in comprehension? *Journal of Memory and Language*, 112, 104087.
- Clark, K., Khandelwal, U., Levy, O., & Manning, C. D. (2019). What does bert look at? an analysis of bert's attention. *arXiv preprint arXiv:1906.04341*.
- Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2018). BERT: pre-training of deep bidirectional transformers for language understanding. *CoRR, abs/1810.04805*. Retrieved from <http://arxiv.org/abs/1810.04805>
- Dillon, B., & Keshev, M. (2024). Syntactic dependency formation in sentence processing: A comparative perspective. *Cambridge Handbook of Comparative Syntax*.
- Fedorenko, E., Babyonyshev, M., & Gibson, E. (2004). The nature of case interference in online sentence processing in Russian. In M. Wolf & K. Moulton (Eds.), *Proceedings of north east linguistics society* 34 (Vol. 34, pp. 215–226). Amherst, MA: GLSA Publisher.
- Hao, S., & Linzen, T. (2023). Verb conjugation in transformers is determined by linear encodings of subject number. *arXiv preprint arXiv:2310.15151*.
- Kesgin, H. T., Yuce, M. K., Dogan, E., Uzun, M. E., Uz, A., Seyrek, H. E., ... Amasyali, M. F. (2024). Introducing cosmogpt: Monolingual training for turkish language models. *arXiv preprint arXiv:2404.17336*.
- Lago, S., Gračanin-Yukse, M., Şafak, D. F., Demir, O., Kırkıcı, B., & Felser, C. (2019). Straight from the horse's mouth: Agreement attraction effects with Turkish possessors [Journal Article]. *Linguistic Approaches to Bilingualism*, 9(3), 398–426.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126–1177.
- Lewis, R. L., & Vasishth, S. (2005). An Activation-Based Model of Sentence Processing as Skilled Memory Retrieval. *Cognitive Science*, 29(3), 375–419.
- Nicol, J., & Antón-Méndez, I. (2009). The effect of case marking on subject-verb agreement errors in English. In W. D. Lewis, S. Karimi, H. Harley, & S. O. Farrar (Eds.), *Time and again: Theoretical perspectives on formal linguistics in honor of d. terence langendoen* (pp. 135–150). Amsterdam, The Netherlands: John Benjamins Publishing.
- Nicol, J., Barss, A., & Barker, J. E. (2016). Minimal interference from possessor phrases in the production of subject-verb agreement. *Frontiers in psychology*, 7, 548.
- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I., et al. (2019). Language models are unsupervised multitask learners. *OpenAI blog*, 1(8), 9.
- Ryu, S. H., & Lewis, R. L. (2021). Accounting for agreement phenomena in sentence comprehension with transformer language models: Effects of similarity-based interference on surprisal and attention. *arXiv preprint arXiv:2104.12874*.
- Slioussar, N. (2018). Forms and features: The role of syncretism in number agreement attraction. *Journal of Memory and Language*, 101, 51–63.
- Türk, U., & Logachev, P. (2024). Agreement attraction in turkish: The case of genitive attractors. *Language, Cognition and Neuroscience*, 39(4), 448–454.
- Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, 61(2), 206–237.
- Wilcox, E. G., Pimentel, T., Meister, C., Cotterell, R., & Levy, R. P. (2023). Testing the predictions of surprisal theory in 11 languages. *Transactions of the Association for Computational Linguistics*, 11, 1451–1470.