Sensitivity to surface-level heuristics: A case from Turkish agreement attraction

Within-experiment statistics in agreement attraction

Utku Turk[a,*]

[a] *University of Maryland, College Park, Linguistics, Marie Mount Hall, College Park, 20742*

Abstract

Surface level does not affect it, but within-experiment statistics effect the findings.

Keywords: keyword1, keyword2

1. Introduction

Speakers often rely on additional sources of information when processing sentences, including plausability, frequency, distributional expectations about forms and tasks, as well as the overall composition of an experimental session (e.g., the ratio of fillers to critical items). Recent work has demonstrated that such task or item-specific factors can substantially modulate reading and judgment behavior (Laurinavichyute and von der Malsburg, 2024; Arehalli and Wittenberg, 2021; Hammerly et al., 2019; Logačev and Vasishth, 2016). One line of research has used the agreement-attraction phenomenon to probe the heuristics that influence sentence processing. Agreement attraction refers to cases in which a verb erroneously agrees with a nearby noun rather than the true subject, giving rise to so-called grammaticality illusions in both production and comprehension (Bock and Miller, 1991; Pearlmutter et al., 1999).

(1) * The key to the cabinets are rusty.

Agreement errors in sentences like (1) have been treated either as a failure of feature reconcilation or a failure of memory encoding. The former set of accounts explain these errors as a by-product of how number feature of a phrase is calculated in real-time (Bock and Miller, 1991; Eberhard et al., 2005; Hammerly et al., 2019). For example, Eberhard et al. (2005) argue that depending on

*Corresponding author
Email address: utkuturk@umd.edu (Utku Turk)

conceptual number, morphophonological number marking, or syntactic dependencies within a phrase, speakers assign a probabilistic number value to phrases. The errors arise when additional plurality features from different sources end up contributing to the final number representation of a phrase. On the other hand, the latter set of accounts claim that the initial representation is not erroneous, but speakers are sometimes unable to correctly retrieve the controller (Wagers et al., 2009; Dillon et al., 2013). For example, Wagers et al. (2009) argue that the parser normally check the agreement relation by retrieving the relevant chunk in memory using the retrieval cues provided by the agreement probe. In sentences like (1), speakers occasionally retrieve the incorrect element due to the fact that neither nouns fully match the relevant cues.

However, both group of accounts generally are underspecified in terms of how meta-linguistic information should be integrated to the inter-sentential dependency mechanisms. Recently, a growing literature have been testing how different types of additional sources that are independent of the linguistic information affects these errors. Recent experiments show that even small changes in task expectations can alter attraction patterns. For example, Laurinavichyute and von der Malsburg (2024) found that varying the practice structure and task demands (reading vs. judgment) affected reading times at the verb in sentences as in (2). In a series of high-powered self-paced reading tasks, they found that when participants answered a comprehension question after each trial, reading times at the verb 'admires' did not differ between (2a) and (2b). However, when participants were asked to judge grammaticality instead, they spent more time reading the verb 'admires' in (2b), suggesting that processing mechanisms can change depending on the expected task.

(2) a. The singer that the actor openly admires apparently received broad international recognition.

b. The singers that the actor openly admires apparently received broad international recognition.

A related set of findings came from Hammerly et al. (2019). They challenge long-standing assumption that the agreement errors only surfaced in ungrammatical sentences such as (1), but not in grammatical sentences as in (3). It has been repeatedly shown that a plural noun increased participants' likelihood to erroneously judge ungrammatical sentences as grammatical; however, participants rarely misidentified grammatical sentences as ungrammatical even when there is an attractor. Hammerly et al. (2019) showed that similar effect surfaced in grammatical sentences when participants' a priori expectations about the experiment is altered. They manipulated the instructions and the number of ungrammatical in an experiment so that participants expected to see more ungrammatical sentences than grammatical sentences. With reduced bias towards grammaticality, they found that the presence of a plural nearby noun affected how speakers completed ungrammatical sentences (1) and grammatical sentences (3) [see TURK2022 for acceptability].

(3) The key to the cabinets is rusty.

Another set of meta-linguistic information that comes from form-driven task strategies. In some languages, surface-form similarity to the agreement controller can exacerbate attraction [BUNCH OF PAPERS] or even drive illusions on its own [CHROMY]. Hartsuiker et al. (2003), for example, used the the accusative marked determiners' form similarity to the nominative marked articles in German, that can be the determiner of an agreement controller, and compared them to distinctively marked dative marked determiners. They showed that participants made agreement errors more often when the preambles contained two NPs that are not marked distinctively (4a) compared to cases where the attractor is distinguishable solely using the form (4b). More importantly, this effect only surfaced when the head and the distractors are both nouns with feminine gender, and not masculine or neuter gender, which do not show the surface level similarity in plurals.

(4) a. Die Stellungnahme gegen die
the.F.NOM.SG position against the.F.ACC.PL
Demonstration-en
demonstration-PL
'The position against the demonstrations'

b. Die Stellungnahme zu den Demonstration-en
the.F.NOM.SG position on the.F.DAT.PL demonstration-PL
'the position on the demonstrations'

The most striking evidence comes from work by Slioussar (2018), who showed that surface form can sometimes override abstract features in Russian. Exploiting the syncretism between singular genitives and nominative plurals, a pattern absent in plural genitives, she found that participants made more production and comprehension errors and showed faster reading times in sentences like (5a) than in (5b). This finding is fairly surprising given that a noun that is unambigously singular and only *seems* plural due to form similarity can induce more attraction errors, than grammatically plural nouns. She argued that, rather than accessing abstract case features, speakers and comprehenders sometimes relied on surface-level cues that were easier to retrieve.

(5) a. Korobka dlya kraski byla/*byli
box.F.SG.NOM for paint.GEN.SG$_{=\text{NOM.PL}}$ be.PST.F.SG/*be.PST.F.PL
…

'A box for the paint(s) was/*were …'

b. Korobka dlya krasok byla/*byli
box.F.SG.NOM for paint.GEN.PL$_{\neq\text{NOM.PL}}$ be.PST.F.SG/*be.PST.F.PL
…

'A box for the paint(s) was/*were …'

Together, these studies converge on the idea that speakers are sensitive not only to the immediate syntactic cues in a sentence, but also to broader distributional regularities that shape how those cues are interpreted. Some of these regularities reflect long-term experience with the language, such as recurring patterns of form syncretism and probability of being a controller, while others arise within the course of a single experiment, as participants adjust to the statistical composition of the materials, the frequency of ungrammatical items, or the mix of structure types they encounter. These findings raise the possibility that agreement attraction is not a fixed structural reflex, but rather a dynamic outcome of how the processing system weighs and re-weights cues based on both linguistic and contextual experience. If so, the strength and even the presence of attraction effects should depend on how strongly surface form and feature structure are correlated in the input, and on what participants learn about those correlations in real time.

Building on these observations, we utilize Turkish, a language that shows attraction effects robustly, as a testing ground to examine how surface-form overlap influences agreement processing and whether exposure to different kinds of distractors within an experiment modulates attraction. Previous research in Turkish has shown that participants do similar errors in agreement attraction comprehension (Lago et al., 2019; Türk and Logačev, 2020; Ulusoy, 2023). Lago et al. (2019) demonstrated agreement attraction effects using genitive-possessive contructions in a speeded acceptability judgment. Their sentences included a complex subjects 'milyonerlerin terzisi' (millionaires' tailor) similar to English saxon-genitives as in (6).

(6) * Milyoner-ler-in öğretmen-i mutfak-ta sürekli zıpla-dı-lar.
millionaire-PL-GEN fix-POSS kitchen-LOC non.stop jump-PST-PL
'Millionaires' fix jumped$_{pl}$ in the kitchen non-stop.'

Türk and Logačev (2020) showed that this effect survives even when the confounding form-ambiguity on the heads 'öğretmen-i' is resolved. The -i marking in Turkish is ambiguous between possessive and accusative case, and the accusative case cannot control agreement in Turkish. They conclude that Turkish speakers do not utilize the form-related features in processing. However, Ulusoy (2023) (Experiment 3) found that attraction effects disappeared when the case marking on the attractor, not the head as in Türk and Logačev (2020), is manipulated. In a series of experiments, they showed a nominative marked plural constituent (7a) can increase acceptability of ungrammatical sentences and decrease the reading difficulty compared to its singular constituent, however a similar effect does not arise with dative marked constituents (7b).

(7) a. * Kütüphaneci(-ler) çalışkan öğrenci-nin
librarian(-PL) hardworking student.SG-GEN
iste-dik-ler-i kitab-ı şimdi bul-du-lar.
want-NMLZ-PL-POSS book-ACC now find-PST-PL
'The instructors caught the child who the girl carelessly pushed.'

b. * Kütüphaneci(-ler)-e çalışkan öğrenci-nin
librarian-PL-DAT hardworking student.SG-GEN
iste-dik-ler-i kitap dün ver-il-di
want-NMLZ-PL-POSS book yesterday give-PASS-PST.
'The instructor caught the child who the girl carelessly pushed.'

Given that

- However, no work has directly tested whether verbal plural morphology can induce similar illusions, or how mixing different attractor types within an experiment affects the magnitude of attraction.

Turkish provides an especially informative case because both nominal and verbal plural markers are realized with the same morpheme, –lAr, yet only nominal plurals bear the syntactic features required for agreement. This allows us to ask whether participants rely on surface-form similarity or on abstract feature representations when evaluating agreement.

- Morphological properties
- Turkish marks number on both nouns and verbs using the identical plural morpheme –lAr.
- Only nominal plurals introduce number features that can agree with the verb; verbal –lAr expresses verbal agreement but is not a potential controller.
- Because of this homophony, Turkish allows form-overlap and feature-mismatch to be disentangled experimentally.

In our first experiment, we test whether plural marking on a verbal distractor—which is morphologically identical but syntactically irrelevant—can elicit attraction. In the second experiment, we combine these verbal distractor conditions with standard nominal attractor conditions to assess how their co-occurrence affects participants' judgments. If attraction effects reflect flexible, context-sensitive processing, the inclusion of verbal distractors should dilute or eliminate the illusion typically observed with nominal attractors.

Together, these experiments extend previous findings on agreement attraction and task sensitivity in two key ways. First, they show that surface-level overlap—even when morphologically identical—does not by itself produce agreement attraction, indicating that participants rely on abstract morphosyntactic features rather than phonological forms. Second, they reveal that participants are not only influenced by the global structure of an experiment (such as the proportion of fillers or grammatical items) but also by the presence of other

condition types within the same task. In other words, attraction effects are attenuated when competing, non-attracting conditions are included, suggesting that agreement processing is dynamically tuned to the statistical context of the experiment itself.

1.1. Experimental logic and predictions

- Goal 1: test whether purely form-based overlap (verbal –lAr) elicits attraction.
- Prediction: if attraction is driven by form, verbal plural distractors should yield higher "acceptable" rates for ungrammatical plurals.
- Alternative: if attraction depends on abstract features, no effect of verbal –lAr should appear.
- Goal 2: test whether the co-occurrence of different attractor types modulates attraction.
- Prediction: if participants adapt to the distribution of conditions, adding verbal distractors (which share the plural form but lack agreement features) should attenuate or eliminate the nominal-attractor illusion.
- Summary: These experiments jointly test whether agreement attraction in Turkish reflects shallow form matching or feature-based computation that is sensitive to the statistical context of the task.

2. Experiment 1: Testing Form-Driven Processing

2.1. Participants

We recruited 80 undergraduate students to participate in the experiment in exchange for course credit. All participants were native Turkish speakers, with an average age of 21 (range: 18 – 31). The experiment was carried out following the principles of the Declaration of Helsinki and the regulations concerning research ethics at Bogazici University. All participants provided informed consent before their participation and their identities were completely anonymised.

2.2. Materials

We used 40 sets of sentences like (8), in which we manipulated (i) the number of the attractor and (ii) the number agreement on the verb. Both plural markings were marked with the suffix -ler/-lar, while the singular number and singular agreement were marked by its absence.

(8) a. Tut-tuğ-u aşçı mutfak-ta sürekli zıpla-dı.
hire-NMLZ-POSS cook[NOM] kitchen-LOC non.stop jump-PST
'The cook they hired$_{sg}$ jumped$_{sg}$ in the kitchen non-stop.'

b. * Tut-tuğ-u aşçı mutfak-ta sürekli zıpla-dı-lar.
hire-NMLZ-POSS cook[NOM] kitchen-LOC non.stop jump-PST-PL
'The cook they hired$_{sg}$ jumped$_{pl}$ in the kitchen non-stop.'

c. Tut-tuk-lar-ı aşçı mutfak-ta sürekli zıpla-dı.
hire-NMLZ-PL-POSS cook[NOM] kitchen-LOC non.stop jump-PST
'The cook they hired$_{pl}$ jumped$_{sg}$ in the kitchen non-stop.'

d. * Tut-tuk-lar-ı aşçı mutfak-ta sürekli
hire-NMLZ-PL-POSS cook[NOM] kitchen-LOC non.stop
zıpla-dı-lar.
jump-PST-PL
'The cook they hired$_{pl}$ jumped$_{pl}$ in the kitchen non-stop.'

All sentences were adapted by previous studies in Turkish agreement attraction (Lago et al., 2019; Türk and Logačev, 2020). Sentences started with a complex subject NP like 'tuttukları aşçı' 'the cook they hired,' in which the nominalized relative clause functioned as the attractor, and the head noun were bare. Because the plural marking on nominals is not optional and the head noun was singular, absent of -lar, in all conditions, sentences with plural verb agreement were ungrammatical. To inhibit participants from forming a task-related strategy in which they deemed the sentence ungrammatical upon seeing a plural verb, half of our fillers included plural grammatical verbs, while the other half included singular ungrammatical verbs.

2.3. Procedures

The experiment was run online, using the web-based platform Ibex Farm (Drummond, 2013). Each experimental session took approximately 25 minutes to complete. Participants provided demographic information and gave informed consent to participate in the experiment. They then proceeded to read the instructions and were given nine practice trials before the experiment began.

Each trial began with a blank screen for 600 ms, followed by a word-by-word RSVP presentation of the sentence in the center of the screen, followed by a prompt to indicate their acceptability judgment. Sentences were presented word-by-word in the center of the screen in 30 pt font size, at a rate of 400 ms per word. Participants saw a blank screen for 100 ms between each word, and to see the next item, they needed to press the space key. Participants were asked to press the key P to indicate that a sentence is acceptable and Q to indicate that the sentence is unacceptable. They were instructed to provide judgments as quickly as possible. During the practice, but not during the experiment, a warning message in red font appeared if they did not respond within 5,000 ms.

Participants saw 40 experimental and 40 filler sentences. Experimental sentences were distributed among four different lists according to a Latin-square design. Every participant saw one version of the experiment with a specific list and one item per condition.

2.4. Analysis and Results

Participants showed high accuracy in both grammatical (M = 0.94, CI = [0.92,0.95]) and ungrammatical filler sentences (M = 0.92, CI = [0.9,0.93]), indicating that they understood the task and performed it reliably.

Figure 1 presents the overall means and credible intervals for 'yes' responses across experimental conditions. As shown, ungrammatical sentences with plural attractors were rated as acceptable as their counterparts with singular attractors (M = 0.06 and 0.05, CI = [0.04, 0.07] and [0.03, 0.07] for singular and plural attractors, respectively).

On the other hand, accuracy in grammatical conditions was modulated by the number of the attractor in an unexpected way. Participants rated grammatical sentences with singular attractors as grammatical less often (M = 0.92, CI = [0.9,0.94]) compared to their counterpars with plural attractors (M = 0.95, CI = [0.93,0.96]).

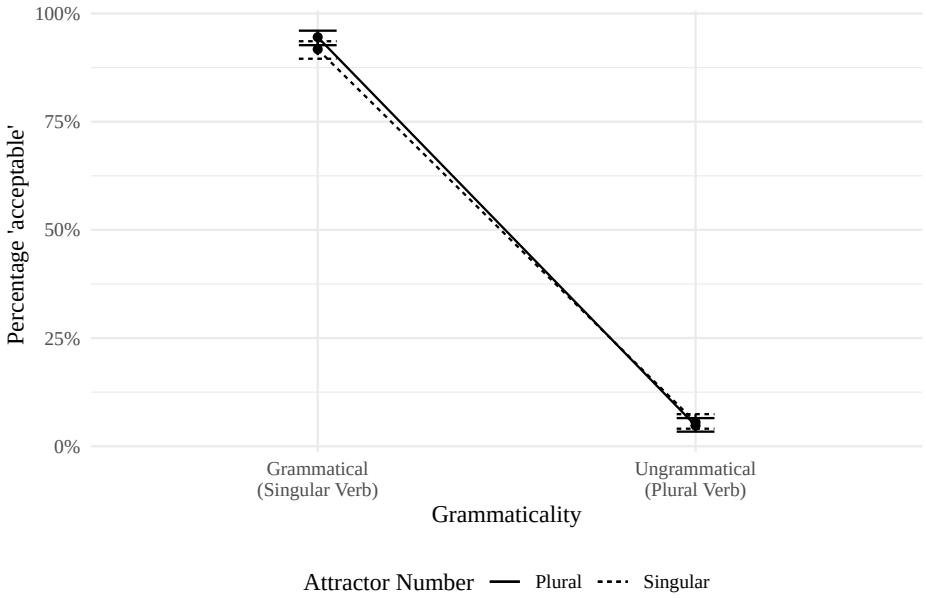

Figure 1: Mean proportion of 'acceptable' responses by grammaticality and attractor number. Error bars show 95% Clopper–Pearson confidence intervals.

These descriptive trends were confirmed by our Bayesian mixed-effects models implemented in brms, assuming a Bernoulli logit link. The model was fitted to the binary *yes/no* responses and included fixed effects for Grammaticality and Attractor Number and their interaction, and random intercepts and slopes for both subjects and items.

Posterior estimates are summarized in Figure 2. The model revealed a positive effect of grammaticality (β = 5.92 [5.41, 6.46], P($\beta > 1.00$)), but no reliable

main effect of attractor number (β = 0.15 [-0.19, 0.51], P($\beta > 0.81$)). On the other hand, there was a small but positive interaction (β = 0.66 [-0.02, 1.38], P($\beta > 0.97$)). To clarify the effects' presence in grammaticals only, we fitted two more models that is fitted to the subset of the data. While the model fitted to grammatical conditions only showed an effect of attractor number (β = 0.51 [0.06, 1.00], P($\beta > 0.99$)), the model fitted to ungrammatical conditions did not provide evidence for the effect of number manipulation (β = -0.05 [-0.45, 0.37], P($\beta > 0.99$)). These results suggest that the presence of a plural attractor did not increase the acceptability of ungrammatical sentences, nor was this relationship modulated by grammaticality.

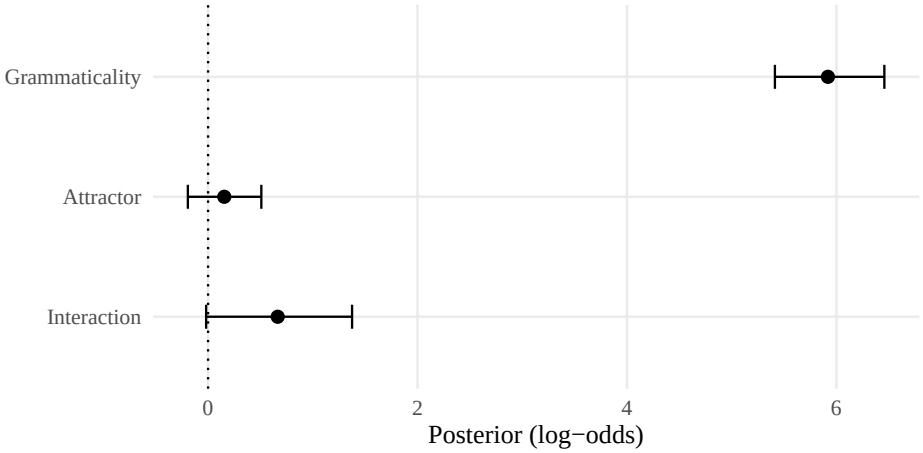

Figure 2: Posterior means and 95% credible intervals for fixed effects in the two Bayesian models. The x-axis shows the posterior mean (log-odds scale). The blue intervals correspond to the model in which a positive interaction was assumed, and the orange intervals to the model in which it was not.

2.5. Discussion

- No attraction effect
- There is an unexpected effect, which is might be due to interaction between the plausability and the availability of a referent. While the plural morpheme can give a general reading, the singular RC probably requires an overt referent. It is outside of the scope of this paper.

3. Experiment 2: Testing Within-Experiment Statistical Sensitivity

	response_yes	
exp_condition	FALSE	TRUE
condition_gen_a	411	54
condition_gen_b	61	404
condition_gen_c	444	22
condition_gen_d	45	424

```
  condition_rc_a    450   22
  condition_rc_b     29  440
  condition_rc_c    450   22
  condition_rc_d     29  433
  filler_g           96 1723
  filler_ung       1750  113
  practice          369  435
[1] "number of bad subjects: 3.000000"
```

3.1. Participants

We recruited 95 undergraduate students to participate in the experiment in exchange for course credit. All participants were native Turkish speakers, with an average age of 21 (range: 18 – 30). The experiment was carried out following the principles of the Declaration of Helsinki and the regulations concerning research ethics at Bogazici University. All participants provided informed consent before their participation and their identities were completely anonymised.

3.2. Materials

The same materials were used with Exp1. We added items from **?** as an additional condition for nominal cases.

3.3. Procedures

The same procedure with Experiment 1 was used.

3.4. Analysis and Results

Participants showed high accuracy in both grammatical (M = 0.95, CI = [0.94,0.96]) and ungrammatical filler sentences (M = 0.94, CI = [0.93,0.95]), indicating that they understood the task and performed it reliably.

Figure 3 presents the overall means and credible intervals for 'yes' responses across experimental conditions, as well as the previous data from Türk and Logačev (2020), which is quite similar to the magnitude of Lago et al. (2019). As shown, in our study, participant gave more 'yes' responses to ungrammatical sentences with plural genitive-marked nominal attractors (M = 0.12, CI = [0.09,0.15]) compared to their singular counterparts (M = 0.12, CI = [0.09,0.15]).

However, similar increase in acceptability was not found with relative clause attractors (M = 0.05 and 0.05, CI = [0.03, 0.07] and [0.03, 0.07] for singular and plural attractors, respectively). Participants rated grammatical sentences similarly independent of the attractor number or attractor type.

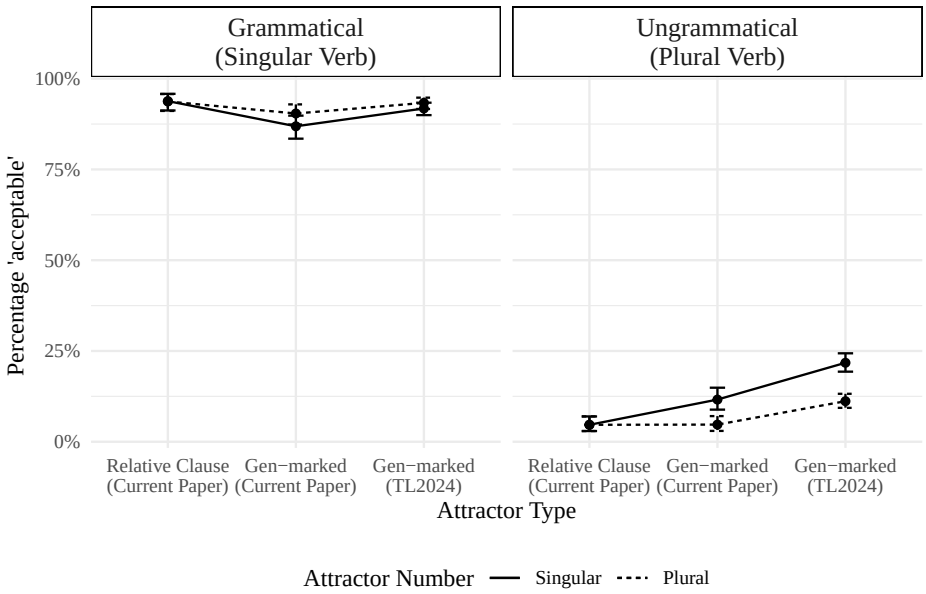

Figure 3: Mean proportion of 'acceptable' responses by grammaticality, attractor number and attractor type. Error bars show 95% Clopper–Pearson confidence intervals.

Our models also showed similar results, assuming a Bernoulli logit link. The model was fitted to the binary *yes/no* responses and included fixed effects for Grammaticality, Attractor Number, and Attractor Type and their interaction, along with random intercepts and slopes for both subjects and items. Since our main question was whether within-experiment statistics affect the grammaticality magnitudes, we fitted another model with genitive marked nominals from data from our experiment and Türk and Logačev (2020).

Talk about the important points. not all of them. attraction effect existed. and it also manipulated as a three way which tells us that participant only did in a single type.

as for our second model, we present the illusion estimate as a function of experiment. Attraction:Current, Attraction:TL24

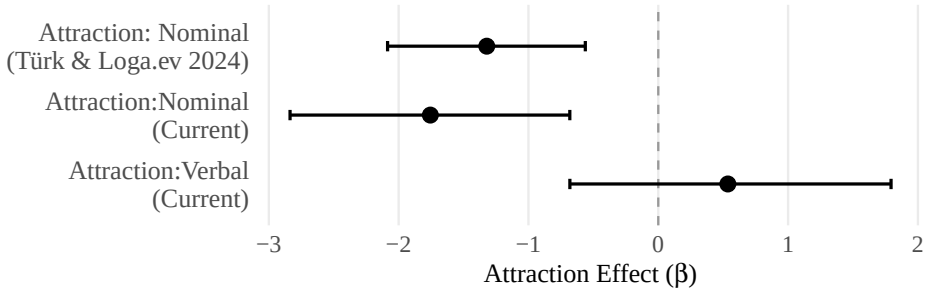

Figure 4: Posterior means and 95% credible intervals for fixed effects in the two Bayesian models. The x-axis shows the posterior mean (log-odds scale). The blue intervals correspond to the model in which a positive interaction was assumed, and the orange intervals to the model in which it was not.

3.5. Discussion

- Goal: test whether attraction changes when both attractor types occur in one experiment.
- Participants: 95 Turkish speakers.
- Design: 2 × 2 × 2 (Grammaticality × Attractor Number × Attractor Type [nominal vs verbal]).
- Procedure & analysis: same as Experiment 1.
- Results:
  - Attraction replicated for nominal attractors (Δ 0.07).
  - Verbal attractors again showed null effect.
  - Global decline in yes-responses relative to earlier studies → participants became more conservative.
- Discussion:
  - Exposure to verbal conditions reduced attraction magnitude overall.
  - Indicates participants adapt to statistical properties of the task.
  - Aligns with learning-based cue-weighting accounts (Haskell et al. 2010).

4. General Discussion

- Synthesis:
  - No evidence for surface-form matching; effects are feature-based.
  - Attraction magnitude changes with condition distribution → adaptive tuning.
- Interpretation:
  - Supports an adaptive parser sensitive to within-experiment statistics.
  - Challenges "shallow" or "good-enough" accounts that attribute attraction to phonological overlap.
- Broader implication:
  - Agreement processing is flexible and probabilistic; illusions arise from learned cue validity.

- Limitations:
  - Syntactic depth asymmetry (verbal attractors more embedded).
  - Need future designs equating structure (e.g., embedded-object attractors).
- Conclusion:
  - Turkish attraction effects arise from abstract feature retrieval not surface level shallow form-matching.
  - The evaluation of abstract features are modulated by distributional learning within the experiment.

References

Arehalli, S., Wittenberg, E., 2021. Experimental filler design influences error correction rates in a word restoration paradigm. Linguistics Vanguard 7, 20200052. URL: https://doi.org/10.1515/lingvan-2020-0052, doi:http://dx.doi.org/doi:10.1515/lingvan-2020-0052.

Bock, K., Miller, C.A., 1991. Broken agreement. Cognitive Psychology 23, 45–93. doi:http://dx.doi.org/10.1016/0010-0285(91)90003-7.

Dillon, B., Mishler, A., Sloggett, S., Phillips, C., 2013. Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. Journal of Memory and Language 69, 85–103. doi:http://dx.doi.org/10.1016/j.jml.2013.04.003.

Drummond, A., 2013. *Ibex farm*. Https://spellout.net/ibexfarm.

Eberhard, K.M., Cutting, J.C., Bock, K., 2005. Making syntax of sense: Number agreement in sentence production. Psychological review 112, 531–559. doi:http://dx.doi.org/10.1037/0033-295X.112.3.531.

Hammerly, C., Staub, A., Dillon, B., 2019. The grammaticality asymmetry in agreement attraction reflects response bias: Experimental and modeling evidence. Cognitive Psychology 110, 70–104. doi:http://dx.doi.org/10.1016/j.cogpsych.2019.01.001.

Hartsuiker, R.J., Schriefers, H.J., Bock, K., Kikstra, G.M., 2003. Morphophonological influences on the construction of subject–verb agreement. Memory & Cognition 31, 1316–1326. doi:http://dx.doi.org/10.3758/bf03195814.

Lago, S., Gračanin-Yuksek, M., Şafak, D.F., Demir, O., Kırkıcı, B., Felser, C., 2019. Straight from the horse's mouth: Agreement attraction effects with Turkish possessors. Linguistic Approaches to Bilingualism 9, 398–426. doi:http://dx.doi.org/10.1075/lab.17019.lag.

Laurinavichyute, A., von der Malsburg, T., 2024. Agreement attraction in grammatical sentences and the role of the task. Journal of Memory and

Language 137, 104525. URL: https://www.sciencedirect.com/science/article/pii/S0749596X24000287, doi:http://dx.doi.org/https://doi.org/10.1016/j.jml.2024.104525.

Logačev, P., Vasishth, S., 2016. A multiple-channel model of task-dependent ambiguity resolution in sentence comprehension. Cognitive Science 40, 266–298.

Pearlmutter, N.J., Garnsey, S.M., Bock, K., 1999. Agreement processes in sentence comprehension. Journal of Memory and Language 41, 427–456. doi:http://dx.doi.org/10.1006/jmla.1999.2653.

Slioussar, N., 2018. Forms and features: The role of syncretism in number agreement attraction. Journal of Memory and Language 101, 51–63. doi:http://dx.doi.org/10.1016/j.jml.2018.03.006.

Türk, U., Logačev, P., 2020. The role of shallow processing in agreement attraction. doi:http://dx.doi.org/10.17605/osf.io/X9PV7.

Ulusoy, E., 2023. Connectivity and case effects in agreement attraction: The case of turkish.

Wagers, M.W., Lau, E.F., Phillips, C., 2009. Agreement attraction in comprehension: Representations and processes. Journal of Memory and Language 61, 206–237. doi:http://dx.doi.org/10.1016/j.jml.2009.04.002.