

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

Utku Turk^{a,*}

^aUniversity 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: form-sensitivity, memory, agreement attraction

1. Introduction

Sentence processing is shaped not only by grammatical constraints but also by plausibility, frequency, task-specific factors, and phonological processes. Recent work shows that such influences 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). Form-based overlap between elements in a sentence can also influence how sentences are processed. A substantial body of work has shown that the parser and the production system are sensitive not only to syntactic or semantic relations but also to the surface form of words. These effects have been taken to suggest that, under certain circumstances, speakers and comprehenders rely on shallow or heuristic cues to complete dependencies. Acheson and MacDonald (2011), for example, found that participants showed slower reading times when the subjects of the two embedded clauses share phonological similarity (*baker-banker* in 1 vs. *runner-banker* in 2). Moreover, participants were less accurate in answering comprehension questions with phonological overlap present. Related work in short-term memory and word recognition shows similar effects—items that overlap phonologically or morphologically are more confusable and more easily retrieved (Copeland & Radvansky, 2001; Rastle & Davis, 2008).

(1) The baker that the banker sought bought the house.

(2) The baker that the banker sought bought the house..

One domain in which these influences are observed is the research on agreement attraction as in (3), a phenomenon in which a verb erroneously agrees with a nearby noun rather than its grammatical subject, producing so-called grammaticality illusions (Bock and Miller, 1991; Pearlmutter et al., 1999). This effect have been robustly attested in many languages with various methodologies [to name a few]. Bock and Eberhard (1993) tested whether attractors that only sound plural, pseudoplurals such as *course 4a*, increase agreement errors compared to true plural nouns (4c). They reasoned that if participants rely on phonological cues rather than abstract features, words ending with plural-like sounds (/s/ or /z/) should behave like true plurals. Participants completed sentence preambles such as (4), where the head noun (*player*) was singular but the attractor varied in form. They found that pseudoplural attractors did not increase plural agreement rates.

*Corresponding author

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

(3) * The player on the courts are tired from a long-game.

- (4) a. Pseudoplural Attractor
The player on the course ...
b. Singular Attractor
The player on the court ...
c. Plural Attractor
The player on the courts ...

Even though modulation from a pure phonological similarity was not found, several experiments have manipulated morphological case similarity between controllers and attractors, reasoning that syncretism or surface ambiguity could enhance competition during retrieval or interfere in production [PAPERS]. For example, [Hartsuiker et al. \(2003\)](#) used the overlap between accusative and nominative forms of feminine determiners in German and compared these ambiguous forms to distinctively marked dative forms. Participants produced more agreement errors when the preambles contained two noun phrases whose determiners were not distinctively marked, as in (5a), compared to cases where the attractor could be distinguished by form alone, as in (5b). Crucially, this additive effect was limited to feminine nouns, the only gender showing nominative-accusative syncretism in plural forms while other nouns showed the base effect of plural.

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

Similar effects of surface similarity are also found in comprehension studies. [Slioussar \(2018\)](#), for example, showed that phonological overlap affects the reading pattern and accuracy of participants in Russian agreement. A group of accusative marked nouns in Russian surfaces ambiguously with their nominative counterparts when they are plural (6a-6b). Meanwhile, it is possible to assign a different case to the attractors using a different preposition as in (6c-6d). Crucially, in her experiment the genitive marked plural nouns were not ambiguous with their nominative counterparts. [Slioussar \(2018\)](#) showed that participants not only exhibited faster reading times at the verb in (6b) compared to (6a), but also judged sentences with a plural attractor as grammatical more often. These effects of plural attractor were only present in cases with ambiguous case marking.

- (6) a. ssylka na sajto byli dany.
link[NOM.SG] to website[ACC.SG(≠NOM.PL)] were given
b. ssylka na sayty byli dany.
link[NOM.SG] to website[ACC.PL(=NOM.PL)] were given
'The link to the website(s) were given.'
c. material dlja kryši byli brakovannymi.
material[NOM.SG] for roof[GEN.SG(=NOM.PL)] were defective
d. material dlja kryš byli brakovannymi.
material[NOM.SG] for roof[GEN.PL(≠NOM.PL)] were defective
'The material for the roof(s) were defective.'

However, a more intriguing aspect of the study by [Slioussar \(2018\)](#) is her results with respect to attractors marked with genitive singular. Another interesting characteristic of Russian is such that a subset of *singular* genitive nouns share the same form with their plural nominal counterpart. In addition to plural nouns

not increasing grammatical judgments to ungrammatical sentences and not creating a reading advantage, the verbs of singular attractors were read faster and resulted in more ‘yes’ responses to grammaticality judgments.

In this aspect, the findings of Slioussar (2018) targets the initial question raised by Bock and Eberhard (1993): whether the pure phonological similarity can drive the agreement attraction effects. Given the contention of initial findings of Bock and Eberhard (1993) with Russian data and the theoretical importance of the empirical generalization, we tested whether we could find attraction effects with another type morphologically rich language, Turkish. Turkish, an almost-strict agglutinative language, presents another typological aspects of morphological marking. English, a predominantly analytic language that uses separate words, such as prepositions, particles, and auxiliary verbs, to express grammatical meaning rather than relying on inflections or affixes attached to words did not show an effect of pure phonological overlap. Meanwhile, Russian, a fusional language in which a single affixal morpheme can express multiple grammatical meanings, exhibited the effect of phonological overlap. Turkish represents another group of languages in which there is close to 1-to-1 mapping between grammatical meanings and affixal morphemes.

In this paper, we test whether pure phonological overlap can derive agreement attraction effects in two high-powered speeded acceptability judgment experiments. To this end we use Turkish, a language where verbal and nominal plural marking share the same surface form, the suffix –lar. We use reduced relative clause (RRC) structures, in which the verb with the plural marking alone can appear as the attractor (7). Importantly, Turkish –lar syncretism here is not feature-ambiguous (as in cases of syncretism); it is a form-only overlap that does not share possible argument status with the subject. Even when the RRC can surface without its head as the subject, they cannot control the agreement (8).

- (7) Gör-dük-ler-i çocuk koş-tu-(*lar).
 go-NMLZ-PL-POSS kid[NOM] run-PST-(*PL)
 ‘The kid that (they) saw ran.’
- (8) Gör-dük-ler-i koş-tu-(*lar).
 go-NMLZ-PL-POSS run-PST-(*PL)
 ‘(The kid) that (they) saw ran.’

In Experiment 1, we tested the form hypothesis directly by comparing ungrammatical sentences with verbal-plural vs. verbal-singular attractors. Experiment 2 replicated this design but included additional nominal-attractor items from a previous Turkish attraction study (Türk and Logačev, 2024), allowing us to test whether the distribution of item types and the presence of genuine attraction-inducing elements modulates the outcome. Across both experiments, we found no evidence that verbal –lar induces attraction, even when canonical nominal attractors are present in the same session. This pattern aligns with prior findings in general attraction literature and Turkish agreement attraction, namely surface-form overlap alone does not drive agreement illusions. Rather, attraction appears to depend on abstract feature overlap between potential controllers and agreement probes. In this lights, findings of Slioussar (2018) becomes even more surprising given that singular attractors that are homophonous with plurals were able to induce attraction effects. One way to reconcile these findings is to refer to types of morphological encoding or the functional utility of morphemes in specific languages following Dillon and Keshev (2024).

1.1. Primer on Attraction Accounts and Phonological Modulation

Agreement errors in sentences like (3) have been treated either as a failure of feature reconciliation 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), in their Marking and Morphing account, argue that speakers and comprehenders does not necessarily create binary singular or plural representation, instead they consider various number

related information in a phrase and create a continuous value. This related information include (i) the inherent conceptual number of the head of phrase, namely collectiveness or distributiveness of a noun, (ii) grammatical number markings on all nouns within a phrase weighted by their syntactic distance, and (iii) idiosyncretic (scissors) or grammatical (books) presence of a plural marking. The errors probabilistically arise when additional plurality features from different sources contribute to the final number representation of a phrase. Because this account ties attraction to various sources including the presence of a plural morpheme, it could in principle accommodate effects of surface-form similarity. However, this surface form similarity effects should be limited to the cases where it can be associated with a plural number marking, and should not arise with pure phonological similarities.

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; Ryskin et al., 2021). For example, Wagers et al. (2009), in their memory account, assumes a cue-based retrieval model in which the verb cues a search for a memory chunk matching relevant cues for subjecthood and number. In the ungrammatical attraction sentences as in (3), each of chunks for ‘courts’ and ‘player’ matches a subset of relevant cues. In these partial match scenarios, erroneously retrieval of the attractor may lead comprehenders to judge the sentence grammatical. In this framework, surface-form overlap affects processing only if it contributes to cue overlap. For example, plural morphology or phonological endings could influence retrieval if the system treats them as diagnostic of plural number. Because the model allows cues to be weighted differently depending on their reliability, it trivially accounts for cross-linguistic variability in the role of case marking and other morphosyntactic features.

The two main set of explanations, therefore, make different predictions for the influence of surface similarity. Feature reconciliation accounts predict form-based attraction if overlapping phonology also results in overlapping morphological representations, whereas cue-based retrieval predicts form effects only when they are encoded as retrieval cues, i.e. relevant for predictions. Indeed, Lago et al. (2019) previously argued that Turkish genitive case marking is associated with subjecthood in Turkish, due to they marking subjects in embedded clause, and thus derives attraction effects in Turkish. A similar account from Dillon and colleagues pushed for sensitivity to looking like a controller (Bhatia and Dillon, 2022; Bleotu and Dillon, 2024). However, both of these accounts argue for distributional or within-item possibility of being a controller, instead of surface level similarity. Turkish verbal -lAr presents a clear dissociation between form-association with subjecthood and being a controller. As we previously noted, they can surface as a subject, they do look like a controller, but they cannot control the agreement and they are not possible controllers. In this aspect, Turkish verbal -lAr is similar to Russian genitive cases. If verbal -lAr, like Russian genitive case, induce attraction errors, current models of attraction need to be revised to include phonological similarity beyond being a possible controller.

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 (9), 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.

- 160 (9) a. Tut-tuğ-u aşçı mutfak-ta sürekli zıpla-dı.
 hire-NMLZ-POSS cook[NOM] kitchen-LOC non.stop jump-PST
 161 ‘The cook they hired_{sg} jumped_{sg} in the kitchen non-stop.’
 162 b. *Tut-tuğ-u aşçı mutfak-ta sürekli zıpla-dı-lar.
 hire-NMLZ-POSS cook[NOM] kitchen-LOC non.stop jump-PST-PL
 163 ‘The cook they hired_{sg} jumped_{pl} in the kitchen non-stop.’
 164 c. Tut-tuk-lar-ı aşçı mutfak-ta sürekli zıpla-dı.
 hire-NMLZ-PL-POSS cook[NOM] kitchen-LOC non.stop jump-PST
 165 ‘The cook they hired_{pl} jumped_{sg} in the kitchen non-stop.’
 166 d. *Tut-tuk-lar-ı aşçı mutfak-ta sürekli zıpla-dı-lar.
 hire-NMLZ-PL-POSS cook[NOM] kitchen-LOC non.stop jump-PST-PL
 167 ‘The cook they hired_{pl} jumped_{pl} in the kitchen non-stop.’

168 All sentences were adapted by previous studies in Turkish agreement attraction (Lago et al., 2019; Türk
 169 and Logačev, 2024). Sentences started with a complex subject NP like ‘tuttukları aşçı’ ‘the cook they
 170 hired,’ in which the nominalized relative clause functioned as the attractor, and the head noun were bare.
 171 Because the plural marking on nominals is not optional and the head noun was singular, absent of -lar, in all
 172 conditions, sentences with plural verb agreement were ungrammatical. To inhibit participants from forming
 173 a task-related strategy in which they deemed the sentence ungrammatical upon seeing a plural verb, half of
 174 our fillers included plural grammatical verbs, while the other half included singular ungrammatical verbs.

175 2.3. Procedures

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

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

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

191 2.4. Analysis and Results

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

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

199 On the other hand, accuracy in grammatical conditions was modulated by the number of the attractor in
 200 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 counterparts with plural attractors ($M = 0.95$, $CI = [0.93, 0.96]$).

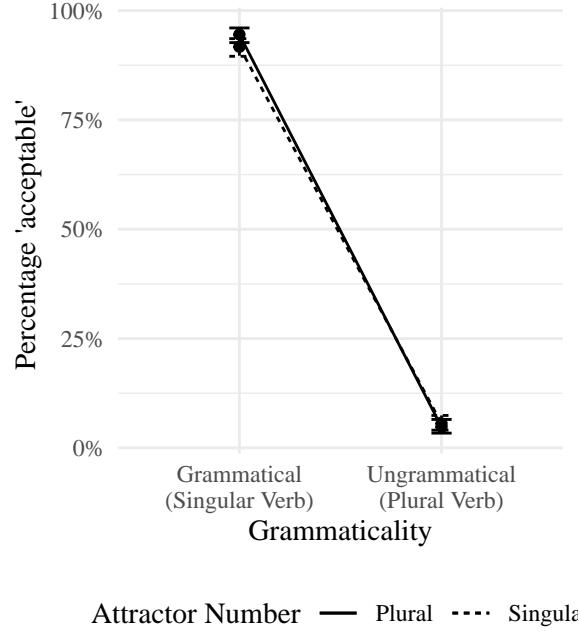


Figure 1: Mean proportion of ‘acceptable’ responses by grammaticality and attractor number. Error bars show 95% Clopper–Pearson confidence intervals.

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 ($\beta = 5.92$ [5.41, 6.46], $P(\beta > 1.00)$), but no reliable main effect of attractor number ($\beta = 0.15$ [-0.19, 0.51], $P(\beta > 0.81)$). On the other hand, there was a small but positive interaction ($\beta = 0.66$ [-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 ($\beta = 0.51$ [0.06, 1.00], $P(\beta > 0.99)$), the model fitted to ungrammatical conditions did not provide evidence for the effect of number manipulation ($\beta = -0.05$ [-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.

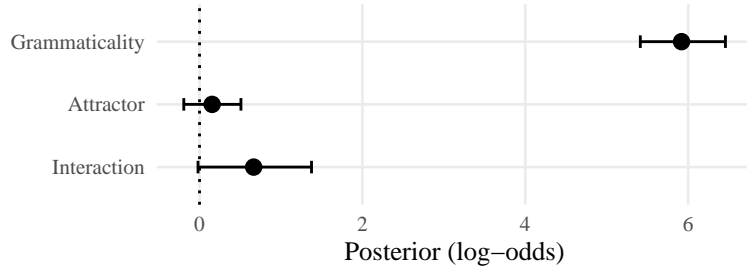


Figure 2: Posterior means and 95% credible intervals for fixed effects in the two Bayesian models. The x-axis shows the posterior mean (log-odds scale). The blue intervals correspond to the model in which a positive interaction was assumed, and the orange intervals to the model in which it was not.

2.5. Discussion

3. Discussion

Experiment 1 tested whether phonological overlap between nominal and verbal plural morphemes in Turkish induces agreement attraction. The results provided no evidence for attraction driven by surface-form similarity. Ungrammatical sentences with plural-marked verbs were not judged more acceptable when the relative clause verb contained a plural morpheme. Instead, participants reliably rejected such sentences regardless of attractor number. This indicates that the verbal plural marker *-lAr* does not create the same type of interference observed with nominal plural attractors in previous studies.

Unexpectedly, grammatical sentences with singular attractors were judged less acceptable than those with plural attractors. This effect is unlikely to reflect agreement attraction, since it arises in the opposite direction. One possibility is that it results from an interaction between plausibility and referential availability. The plural morpheme can license a more general interpretation by allowing an arbitrary or unspecific reference, whereas the singular reduced relative clause more strongly invites a specific referent, which may be less accessible in the context of the task. In other words, plural morphology may facilitate an *arbitrary PRO* interpretation of the embedded clause, in which the understood subject of the relative clause is not controlled by any overt antecedent and has a generic or impersonal reference. A similar effect can be seen in English sentences like ‘Just to sit there should be forbidden.’ Here, the subject of the infinitival clause has arbitrary reference. We do not pursue this explanation further, as it falls outside the scope of the present paper.

One possible reason for the absence of attraction may lie in the within-experiment statistics. Previous work has shown that participants’ global expectations about the frequency of grammatical and ungrammatical sentences can alter attraction patterns. Hammerly et al. (2019) and Türk (2022) demonstrated that reducing the proportion of grammatical trials led to attraction effects even in otherwise grammatical sentences. Similarly, Arehalli and Wittenberg (2021) reported that filler distribution affects error correction rates. It is possible that the current experiment’s distribution discouraged attraction: if participants rarely encountered conditions that supported attraction, they may have maintained a strong bias against plural-marked verbs, reinforcing this bias throughout the session.

To test this possibility, Experiment 2 introduced additional conditions that have previously been shown to elicit attraction in Turkish (Türk and Logačev, 2024; Lago et al., 2019). This allowed us to assess whether the inclusion of genuine nominal attractors modulates the likelihood of errors and whether participants adapt to the statistical environment of the task.

4. Experiment 2: Testing Within-Experiment Statistical Sensitivity

4.1. 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.

4.2. Materials

The same materials were used with Exp1. We added items from Türk and Logačev (2024) as an additional condition for nominal cases.

4.3. Procedures

The same procedure with Experiment 1 was used.

4.4. 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 (2024), 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.

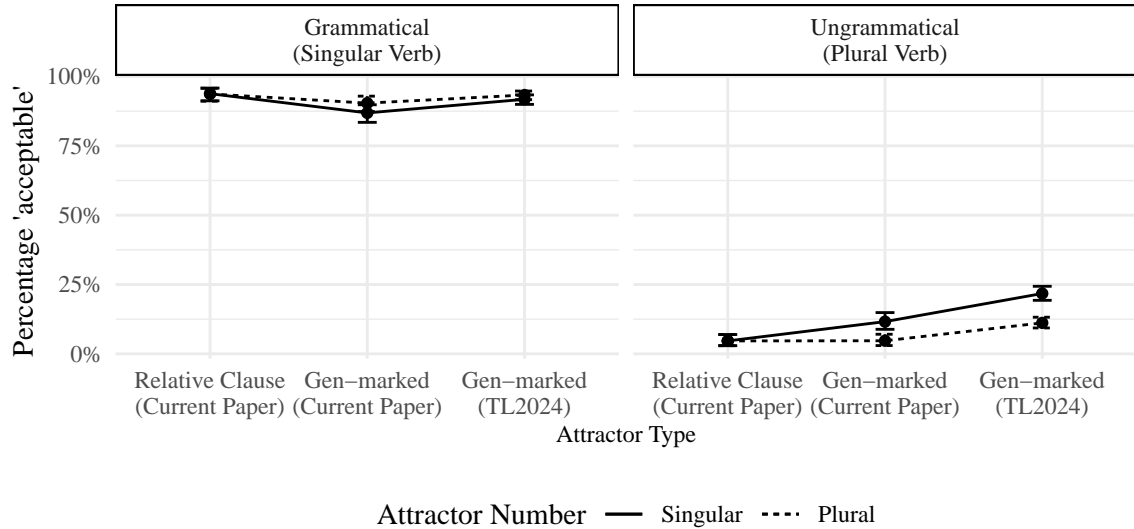


Figure 3: Mean proportion of ‘acceptable’ responses by grammaticality, attractor number and attractor type. Error bars show 95% Clopper–Pearson confidence intervals.

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 (2024).

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

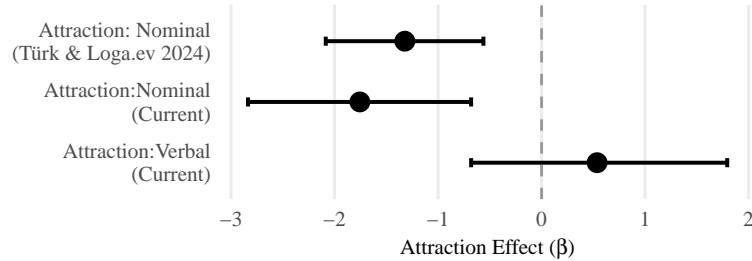


Figure 4: Posterior means and 95% credible intervals for fixed effects in the two Bayesian models. The x-axis shows the posterior mean (log-odds scale). The blue intervals correspond to the model in which a positive interaction was assumed, and the orange intervals to the model in which it was not.

4.5. Discussion

- Goal: test whether attraction changes when both attractor types occur in one experiment.
- Participants: 95 Turkish speakers.
- Design: $2 \times 2 \times 2$ (Grammaticality \times Attractor Number \times Attractor Type [nominal vs verbal]).
- Procedure & analysis: same as Experiment 1.
- Results:
 - Attraction replicated for nominal attractors ($\Delta = 0.07$).
 - Verbal attractors again showed null effect.
 - Global decline in yes-responses relative to earlier studies \rightarrow 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).

5. General Discussion

- Synthesis:
 - No evidence for surface-form matching; effects are feature-based.
 - Attraction magnitude changes with condition distribution \rightarrow 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

- Acheson, D.J., MacDonald, M.C., 2011. The rhymes that the reader perused confused the meaning: Phonological effects during on-line sentence comprehension. *Journal of memory and language* 65, 193–207.
- 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>.
- Bhatia, S., Dillon, B., 2022. Processing agreement in hindi: When agreement feeds attraction. *Journal of Memory and Language* 125, 104322.
- Bleotu, A.C., Dillon, B., 2024. Romanian (subject-like) dps attract more than bare nouns: Evidence from speeded continuations. *Journal of Memory and Language* 134, 104445.
- Bock, K., Eberhard, K.M., 1993. Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes* 8, 57–99. doi:<http://dx.doi.org/10.1080/01690969308406949>.
- 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](http://dx.doi.org/10.1016/0010-0285(91)90003-7).
- Dillon, B., Keshev, M., 2024. Syntactic dependency formation in sentence processing: A comparative perspective, in: Barbiers, S., Corver, N., Polinsky, M. (Eds.), *The Cambridge Handbook of Comparative Syntax*. Cambridge University Press, Cambridge, pp. 1031–1069.
- 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-Yukse, M., Şafak, D.F., Demir, O., Kırkıcı, B., Felsler, 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>.
- Ryskin, R., Bergen, L., Gibson, E., 2021. Agreement errors are predicted by rational inference in sentence processing. Preprint available at OSF: <https://osf.io/uaxsq>.
- 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., 2022. Agreement attraction in Turkish. Ph.D. thesis. Bogaziçi University. İstanbul, Turkey.
- Türk, U., Logačev, P., 2024. Agreement attraction in turkish: The case of genitive attractors. *Language, Cognition and Neuroscience* 39, 448–454.
- 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>.