Sensitivity to within-experiment statistics: 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 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-specific factors can substantially modulate reading and judgment behavior (e.g., Malsburg, Logacev; Hammerly et al.; Arehalli). 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.

1. *The key to the cabinets are rusty.

Recent experiments show that even small changes in task expectations can alter attraction patterns. For example, Malsburg and colleagues found that varying the practice structure and task demands (reading vs. judgment) affected reading times at the verb in sentences such as the following:

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

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

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

In a self-paced reading task, when participants answered a comprehension question after each trial, reading times at the verb did not differ between (2) and (3). However, when participants were asked to judge grammaticality instead, they spent more time reading the verb admires in (3), suggesting that processing heuristics depend on the expected output of the task.

A related set of findings comes from work by Slioussar, 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 errors and showed longer reading times in sentences like (4) than in (5). She argued that, rather than accessing abstract case features, readers relied on surface-level cues that were easier to retrieve.

4. [Example sentence with syncretic form …]
5. [Example sentence with non-syncretic form …]

Building on these observations, we utilize Turkish as a testing ground to examine how surface-form overlap influences agreement processing and whether exposure to different kinds of distractors modulates 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.

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. Agreement attraction and its modulation

- Core phenomenon and theoretical accounts. Briefly summarize leading explanations:
- Cue-based retrieval / similarity-based interference (Wagers, Lau & Phillips 2009; Dillon et al. 2013): attraction arises because plural attractors partially match number cues used in retrieval.
- Feature-percolation accounts (Bock & Miller 1991; Franck et al. 2002): features spread upward within noun phrases, confusing the parser.
- Context and task modulation
- Show that attraction strength varies across tasks and experimental context:
- Task type: self-paced reading vs. acceptability judgment (Malsburg).
- Response expectations and bias: Hammerly et al. 2019 — participants' error patterns change when they expect more ungrammatical items.
- Item composition: Laurinavichyute & von der Malsburg 2023 — manipulating filler ratios or condition mixes alters accuracy patterns.
- Conclude that participants adapt their behavior to the statistical environment of the experiment; attraction is not fixed but context-sensitive.
- Form-based influences. in some languages, surface-form similarity can exacerbate attraction or even drive illusions on its own.
- Cite Slioussar (2018) and related Russian work: case syncretism between nominative plurals and genitive singulars increased attraction or slowed reading, suggesting reliance on form-based heuristics rather than abstract case features.
- Cite Chromy and checz data. tell it is not very consistent.

open question — when morphology overlaps across categories, does attraction arise from form or feature?

1.2. Background on Turkish

- 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.
- Previous attraction findings in Turkish
- Prior work has reported typical attraction effects with genitive-marked nominal attractors, showing higher acceptance of ungrammatical plural-verb sentences.
- 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.

1.3. 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

- Goal: determine if surface plural forms (verbal -lAr) elicit illusionary agreement.
- Participants: 80 Turkish speakers (Boğaziçi undergraduates).
- Design: 2 × 2 (Grammaticality × Attractor Number).
- Materials: relative-clause verbs as attractors; same surface morphology as nominal plurals.
- Procedure: speeded acceptability judgments, 1500 ms deadline.
- Analysis: Bayesian probit GLM (brms); random intercepts/slopes by subject/item.
- Results:

  - High filler accuracy ($> .9$).
  - No difference in ungrammatical sentences between plural vs singular attractors.
  - Posterior coefficients near 0; 95 % CIs within ROPE.
- Discussion:
  - No evidence for form-driven guessing.
  - Participants rely on abstract number features, not phonological similarity.

3. Experiment 2: Testing Within-Experiment Statistical Sensitivity

- 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 ($\Delta \quad 0.07$).
  - 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).

4. 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