

Rational inference does not predict agreement errors:

Gender versus number attraction in Hebrew comprehension

Mandy Cartner ♦ Brian Dillon ♦
Aya Meltzer-Asscher ♦ Maayan Keshev

A noisy channel approach to Agreement Attraction

Number attraction: More agreement errors when a verb-matching plural distractor appears in the sentence [1-3].

- A) *No match*: *The key to the cabinet were rusty ☹️
B) *Dist. match*: *The key to the cabinets were rusty 😊

Markedness based accounts: The *marked* feature of a distractor interferes with the encoding [4-5] or retrieval [2] of the head noun.

Noisy-channel based accounts: Prior knowledge facilitates rational misinterpretation of the head noun as plural given a distractor [6-7]. A comprehender is more likely to infer that S_p is a corruption of S_i the larger the ratio in (i) [8]. This ratio is larger in *Dist. match* than *No match*, due to prior probability of each relevant S_p .

$$(i) \frac{P(S_i|S_p)}{P(S_p|S_p)} = \frac{P_L(S_i)P_N(S_i \rightarrow S_p)}{P_L(S_p)P_N(S_p \rightarrow S_p)}$$

S_p : The key to the cabinets
↓
 S_i : The keys to the cabinets

Number versus Gender attraction

Gender attraction: More agreement errors are observed with a marked feminine distractor [9-11].

- C) *No match*: *The tailor.M of the designer.M improvises.F ☹️
D) *Dist. match*: *The tailor.M of the designer.F improvises.F 😊

Markedness based accounts do not predict a contrast in the rates of number and gender attraction.

Noisy-channel based accounts predict attraction rates to vary based on the degree to which the marked distractor lowers $P(S_p|S_p)$, and this may vary across number and gender.

$$S_p: \text{The tailor.M of the designer.F} \rightarrow S_i: \text{The tailor.F of the designer.F}$$

References: [1] Bock & Miller (1991); [2] Wagers et al. (2009); [3] Pearlmutter et al. (1999); [4] Bock et al. (2001); [5] Keshev et al. (2024); [6] Ryskin et al. (2021); [7] Brehm et al. (2021); [8] Levy (2008); [9] Deutsch & Dank (2009); [10] Tucker et al. (2021); [11] Antón-Méndez et al. (2002).

This work was supported by NSF-BSF grant #2146798 to AM & BD.



Hebrew corpus analysis

- heTenTen21 corpus
- 10k tokens of [NP₁ of NP₂]; all animate

Noisy-channel does not predict attraction errors since S_p is more probable than S_i . Still, attraction is attributed to a difference in $P(S_i|S_p)$ for *No match* compared to *Dist. match* [6].

No match S_p : Sg of Sg is 4.590 more likely than S_i : Pl of Sg
Dist. match S_p : Sg of Pl is 1.425 more likely than S_i : Pl of Pl

In *gender*, the degree to which *Dist. match* misinterpretation is more likely than *No match* misinterpretation is **smaller** than *number*, **predicting less attraction in gender than number**.

No match S_p : M of M is 4.249 more likely than S_i : F of M
Dist. match S_p : M of F is 3.366 more likely than S_i : F of F

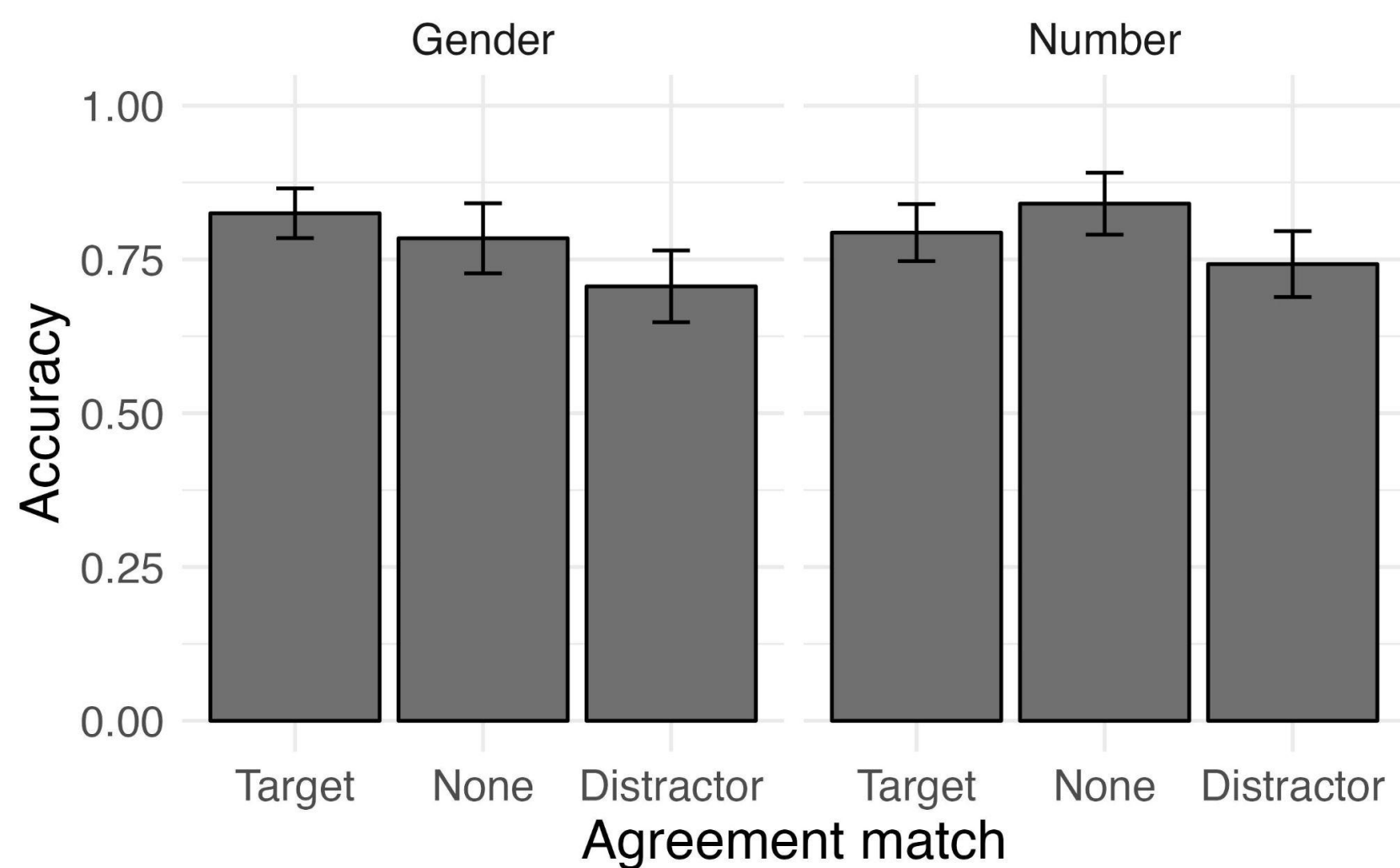
Note: In Hebrew, gender errors involve omission of 1 letter; number involves 2 → Noisy channel predicts more gender errors overall (i.e., in *No match* too).

Hebrew speeded acceptability task: Methods

- 94 Hebrew speakers
- RSVP + speeded acceptability
- 36 item sets, crossing: **Feature** (*Gender, Number*) and **Agreement match** (*Target, None, Distractor*).

Feature: Gender	
Target	...ha-xayetet šel ha-me'acev tamid me'alteret... ...the-tailor.F of the-designer.M always improvises.F...
None	ha-xayat šel ha-me'acev tamid me'alteret the-tailor.M of the-designer.M always improvises.F
Distractor	ha-xayat šel ha-me'acevet tamid me'alteret the-tailor.M of the-designer.F always improvises.F
Feature: Number	
Target	...ha-xayatim šel ha-me'acev tamid me'alterim... ...the-tailor.Pl of the-designer.Sg always improvise.Pl...
None	ha-xayat šel ha-me'acev tamid me'alterim the-tailor.Sg of the-designer.Sg always improvise.Pl
Distractor	ha-xayat šel ha-me'acvim tamid me'alterim the-tailor.Sg of the-designer.Pl always improvise.Pl

Hebrew speeded acceptability task: Results



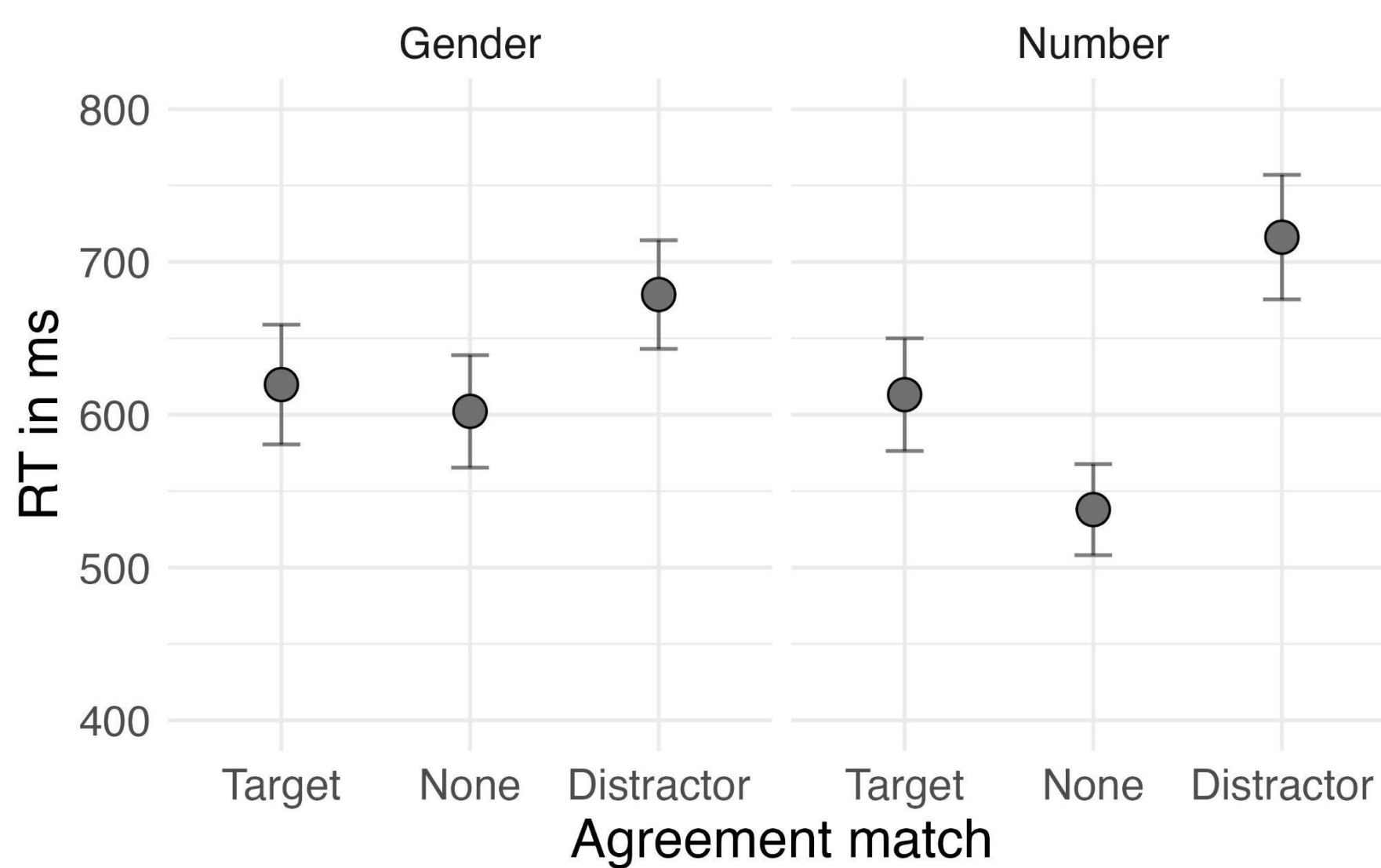
Attraction: Lower accuracy in *Distractor* compared to *No match*.

Gender (0.31, [0.11, 0.52] 95% CRI)

Number (0.46, [0.24, 0.7] 95% CRI).

Null interaction: no evidence for contrast in attraction between *Number* and *Gender* (Est. 0.21, SE 0.17, 95% CRI [-0.13, 0.54]).

BF <1/100: Overwhelming evidence against an interaction.
Priors: Intercept ~ normal(0,3); effects ~ normal(0,1)



Slowdown in Distractor compared to *No match* (-0.07, [-0.11, -0.03] 95% CRI).

Unreliable interaction with *Feature* (-0.02, [-0.06, 0.01] 95% CRI).

BF > 100: Overwhelming evidence for model with an interaction.
Priors: Intercept ~ normal(6.6,1); effects ~ normal(0,1)

Discussion

- The frequency of different [NP₁ of NP₂] types in a Hebrew corpus predicts more attraction errors to arise in *number* compared to *gender*, if attraction were due to rational misinterpretation.
- Gender and number attraction rates are equivalent** in a speeded acceptability judgment task [cf. 9-11], although RT data shows more uncertainty in *number*.
- Attraction is unlikely to result from rational misinterpretation.** The similarity between gender and number suggests attraction is sensitive to the way abstract features are represented.
- Moreover, our results do not support the claim that number features are more cognitively salient than gender [10].