Novel analysis of response bias challenges representational accounts in attraction

Utku Türk & Pavel Logačev

Department of Linguistics
University of Maryland
www.utkuturk.com
utkuturk@umd.edu

Department of Linguistics
Boğaziçi University

Agreement Attraction

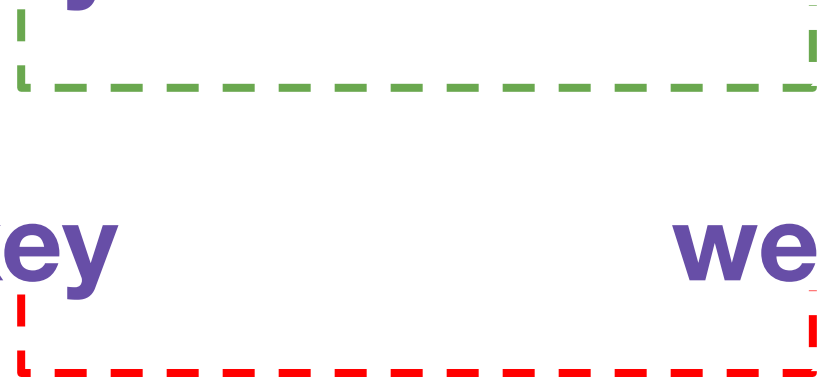
(1) The **key** **was** rusty.

(Wagers et al., 2009)

Agreement Attraction

(1) The **key** **was** rusty.

(2) *The **key** **were** rusty.



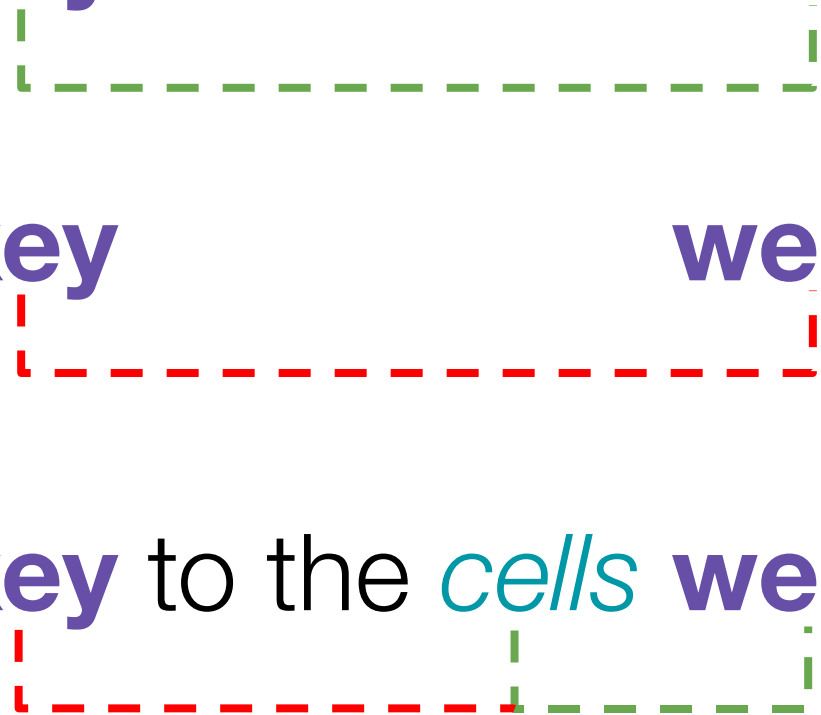
(Wagers et al., 2009)

Agreement Attraction

(1) The **key** **was** rusty.

(2) *The **key** **were** rusty.

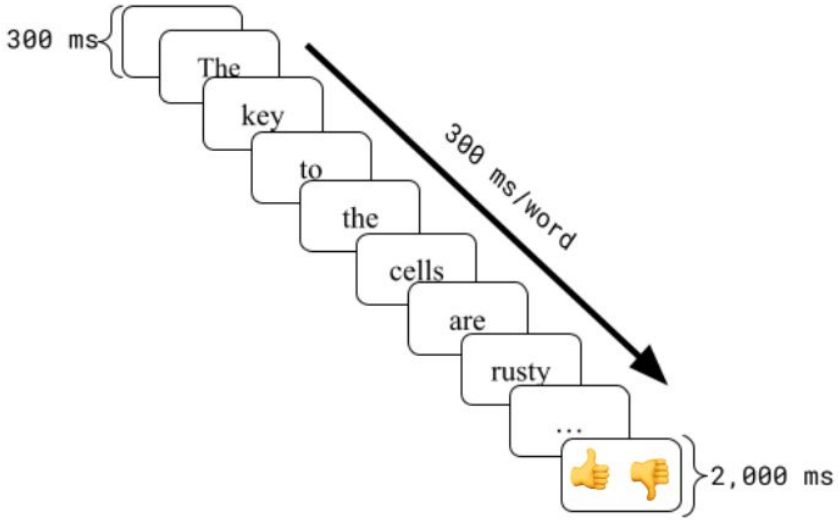
(3) *The **key** to the *cells* **were** rusty.



(Wagers et al., 2009)

Agreement Attraction: **Comprehension**

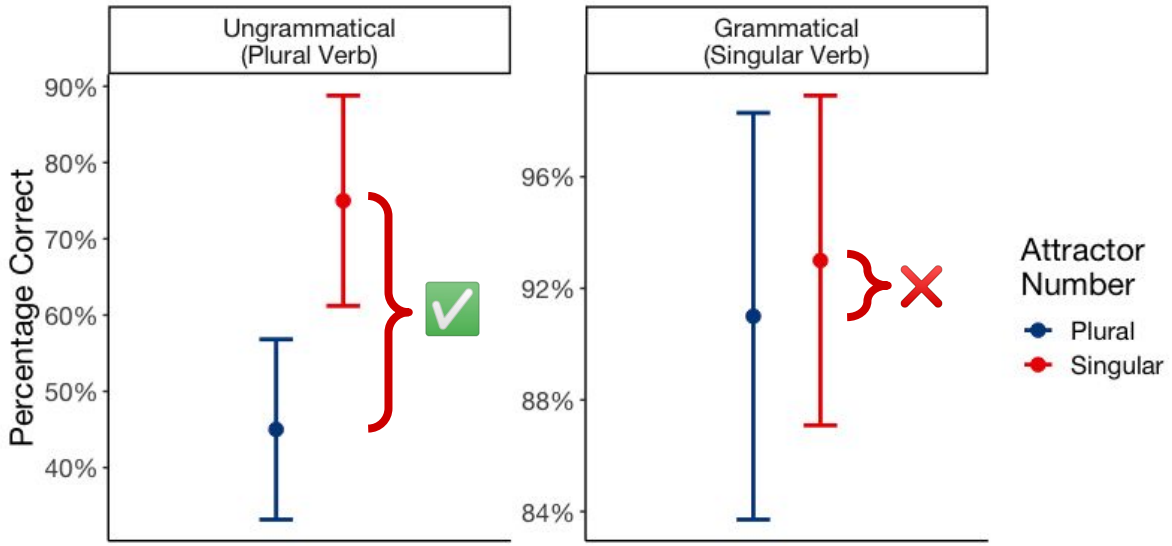
- Word-by-word, speeded acceptability judgment task



(Wagers et al., 2009)

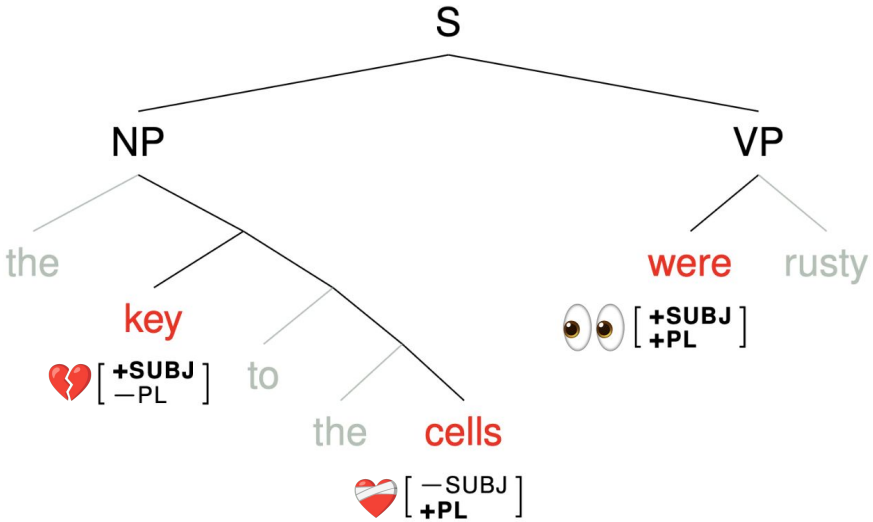
Agreement Attraction: **Grammaticality Asymmetry**

- Attractor number mattered only in ungrammatical sentences

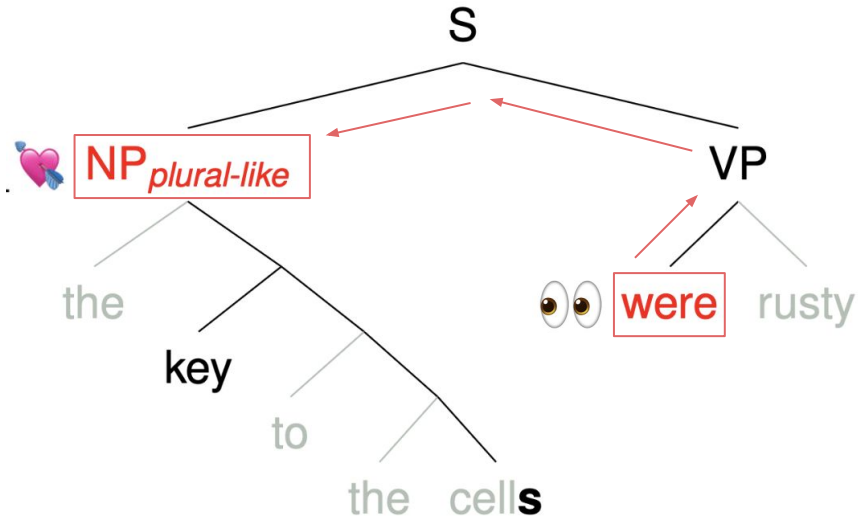


(Wagers et al., 2009)

Agreement Attraction: **Accounts**



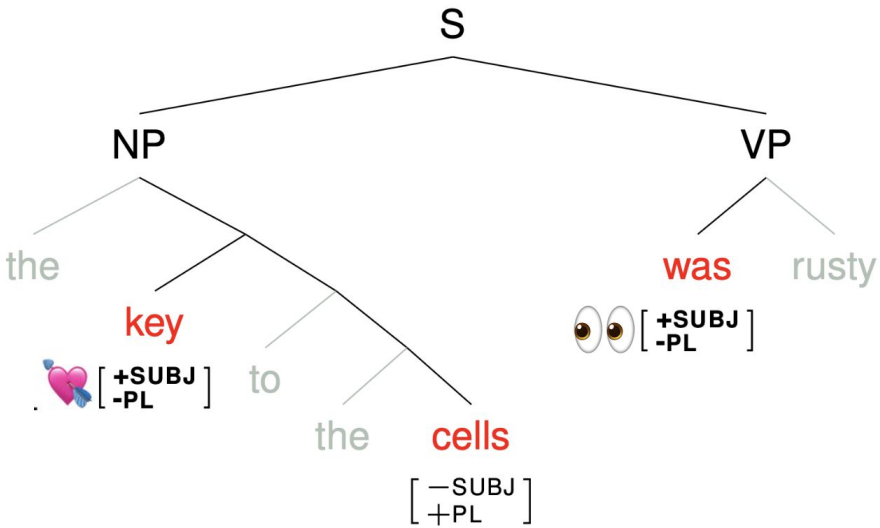
- **Retrieval:** Partial match may occasionally save the retrieval.



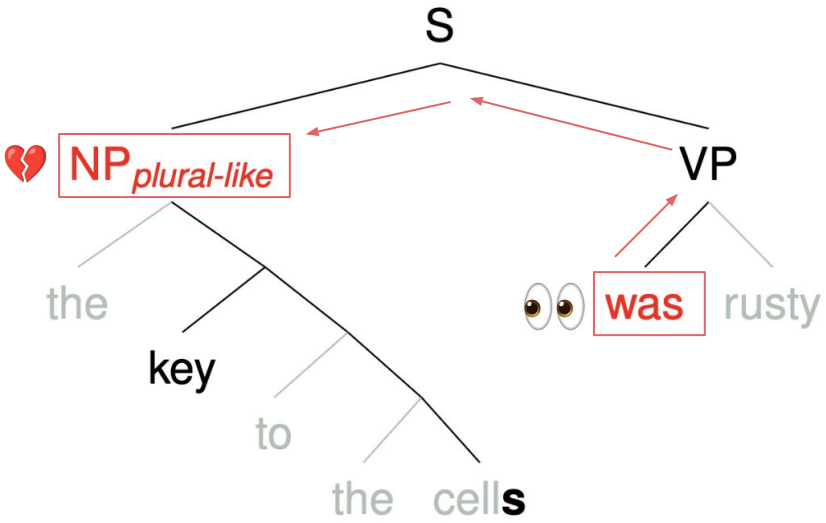
- **Representational:** probing acceptability in ungrammatical sentences

(Eberhard et al., 2005; Wagers et al., 2009; Engelmann, Jäger, & Vasishth, 2019)

Agreement Attraction: **Accounts**



- **Retrieval:** Less interference when the true subject is a *perfect match*.



- **Representational:** probing *unacceptability* even in grammatical sentences

(Eberhard et al., 2005; Wagers et al., 2009; Engelmann, Jäger, & Vasishth, 2019)

Roadmap

- 1. Hammerly et al. (2019) and Bias Calculation Problem
- 2. Turkish Experiment
- 3. Re-analysis of Hammerly et al. (2019)
- 4. Meta-analysis

Grammaticality Asymmetry: **Response Bias**

Hammerly et al. (2019)

- Grammaticality asymmetry does not have to favor Retrieval

(Hammerly et al., 2019)

Grammaticality Asymmetry: **Response Bias**

Hammerly et al. (2019)

- Grammaticality asymmetry does not have to favor Retrieval
- People have **a priori 'yes' bias**

(Ratcliff, 1978; Hammerly et al., 2019)

Grammaticality Asymmetry: **Response Bias**

Hammerly et al. (2019)

- Grammaticality asymmetry does not have to favor Retrieval
- People have **a priori 'yes' bias**
- When bias is reduced, both grammatical and ungrammatical sentences will show attraction effects

(Ratcliff, 1978; Hammerly et al., 2019)

Grammaticality Asymmetry: **Response Bias**

Hammerly et al. (2019)

- Manipulated bias through
  - instructions
  - ungrammatical to grammatical filler ratios

(Ratcliff, 1978; Hammerly et al., 2019)

Grammaticality Asymmetry: **Response Bias**

Hammerly et al. (2019)

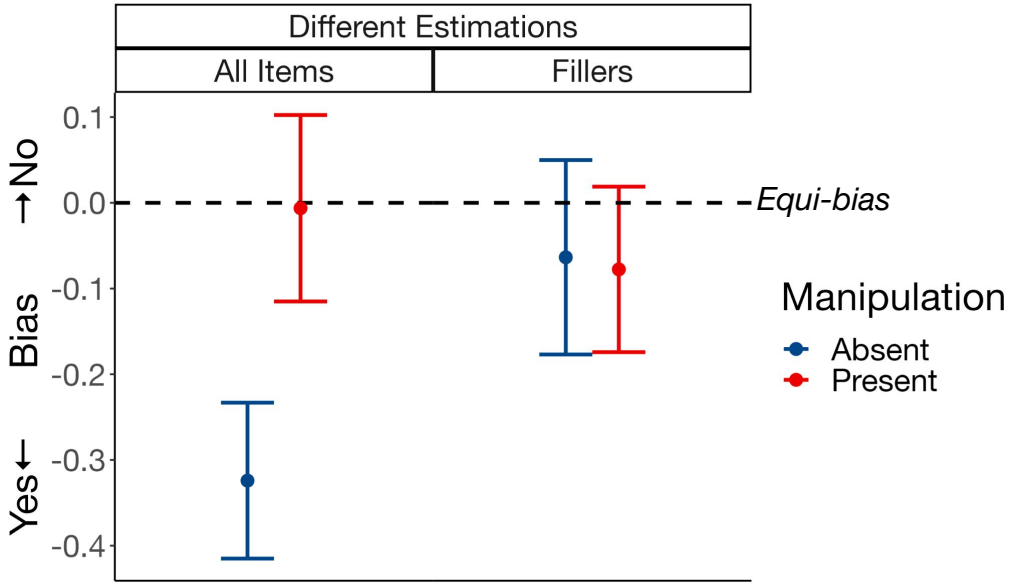
- Manipulated bias through
  - instructions
  - ungrammatical to grammatical filler ratios

- Results: Symmetrical effects independent of well-formedness.
  Thus, asymmetry is a residue of a response bias.

(Ratcliff, 1978; Hammerly et al., 2019)

Grammaticality Asymmetry: **Response Bias**

- <u>Problem:</u> They used all items in bias estimation.



(Macmillan & Creelman, 2005)

Grammaticality Asymmetry: **Why Fillers?**

- Experimental items can inflate the bias estimate

Grammaticality Asymmetry: **Why Fillers?**

- Experimental items can inflate the bias estimate

- Bias in experimental items may tap into different mechanisms

Grammaticality Asymmetry: **Why Fillers?**

- Experimental items can inflate the bias estimate
- Bias in experimental items may tap into different mechanisms
- Fillers are constant between participants

Grammaticality Asymmetry: **Why Fillers?**

- Experimental items can inflate the bias estimate
- Bias in experimental items may tap into different mechanisms
- Fillers are constant between participants
- More conservative test of the hypothesis

Grammaticality Asymmetry: **Why Fillers?**

- Experimental items can inflate the bias estimate
- Bias in experimental items may tap into different mechanisms
- Fillers are constant between participants
- More conservative test of the hypothesis
- Indifferent to experimental manipulations

Roadmap

- ✔ 1. Hammerly et al. (2019) and Bias Calculation Problem
- ❑ 2. Turkish Experiment
- ❑ 3. Re-analysis of Hammerly et al. (2019)
- ❑ 4. Meta-analysis

Our Study ($N_{subject}$ = 114)

Goal: to test Hammerly et al. theoretical findings and verify bias results.

Our Study: **Experimental Items**

- Used Genitive modifier DPs as attractors (Lago et al., 2019)

(4) [$_{DP}$ [$_{DP}$ *Milyoner-**ler**-in*] terzi-si]

millionaire-**PL**-GEN tailor-POSS

"the tailor of *the millionaires*"

Our Study: **Experimental Items**

- Ungrammaticality due to singular head and plural verb.

(5) * [$_{DP}$ [$_{DP}$ *Milyoner-**ler**-in*] terzi-si] kov-ul-du-**lar.**

millionaire-**PL**-GEN tailor-POSS fire-PASS-PST-**PL**

"the tailor of *the millionaires* were fired."

Our Study: **Experimental Items**

- Within-subject factors: *Verb x Attractor number*

(6) a. * [$_{DP}$ [$_{DP}$ *Milyoner-**ler**-in* terzi-si] tamamen gereksizce kov-ul-du-**lar**.
millionaire-PL-GEN tailor-POSS completely without_reason fire-PASS-PST-PL
"The tailor of *the millionaires* were fired for no reason at all."

b. * *Milyonerin* terzisi tamamen gereksizce kovuldu**lar**.

c. *Milyoner**ler**in* terzisi tamamen gereksizce kovuldu.

d. *Milyonerin* terzisi tamamen gereksizce kovuldu.

(Lago et al., 2019)

Our Study: **Bias**

- Between subjects factor: Bias.

- Manipulated bias through
  - instructions
  - ungrammatical to grammatical filler ratios

(Hammerly et al., 2019)

Our Study: **Bias**

- Between subjects factor: Bias.

- Manipulated bias through
  - instructions
  - ungrammatical to grammatical filler ratios

- No bias difference between groups

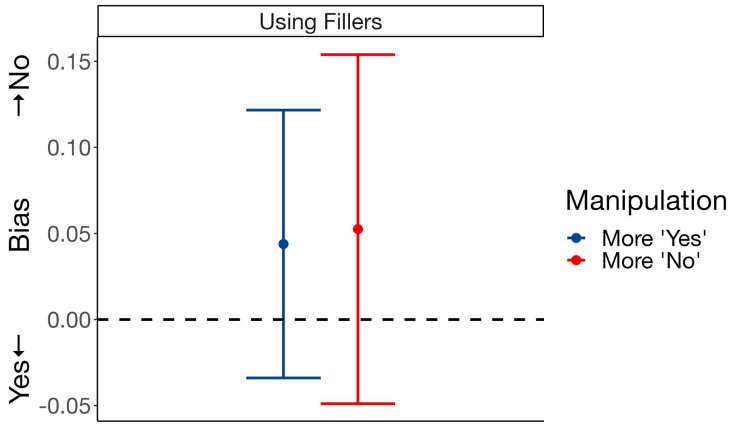
(Hammerly et al., 2019)

Our Study: **Bias**

- Between subjects factor: Bias.

- Manipulated bias through
  - instructions
  - ungrammatical to grammatical filler ratios

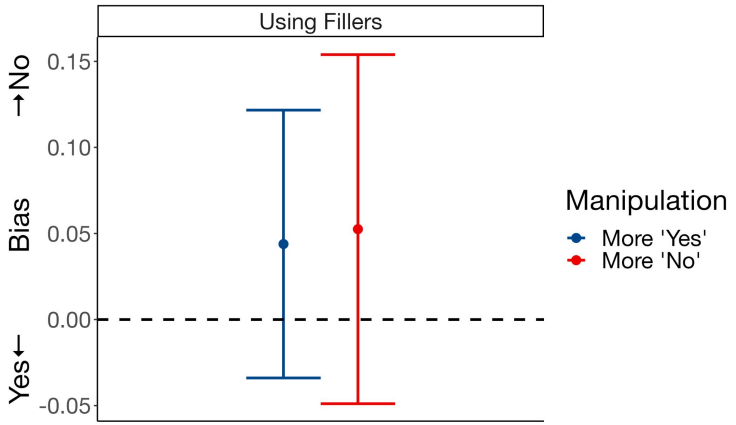
- No bias difference between groups



(Macmillan & Creelman, 2005; Hammerly et al., 2019)

Our Study: **Bias**

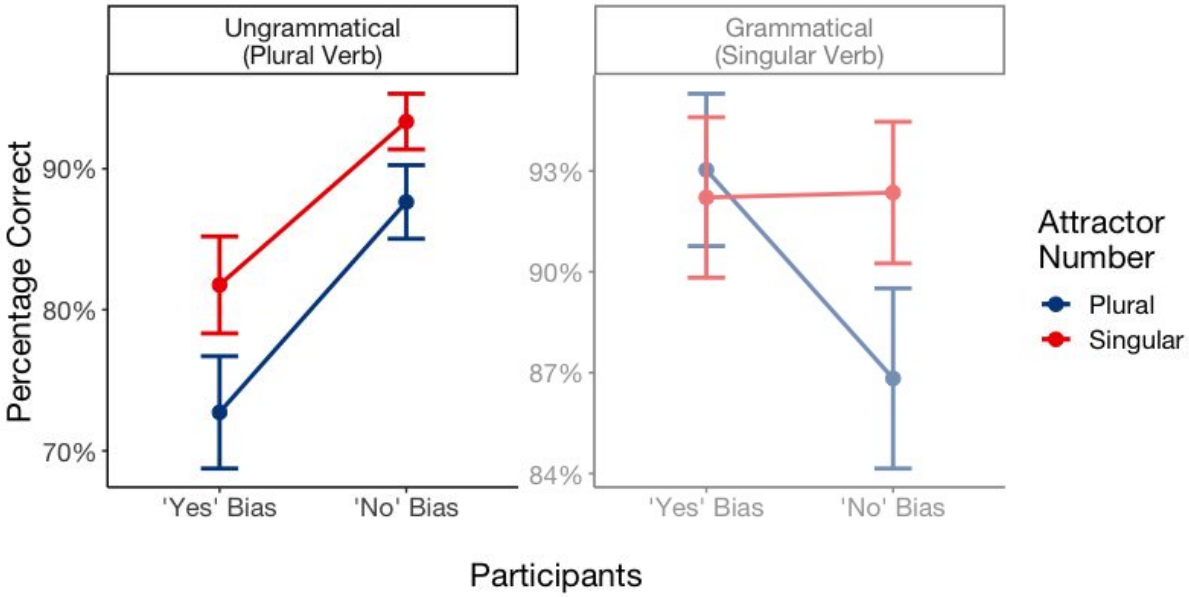
- Between subjects factor: Bias.
- Manipulated bias through
  - instructions
  - ungrammatical to grammatical filler ratios
- No bias difference between groups
- Exploited the individual bias differences



(Macmillan & Creelman, 2005; Hammerly et al., 2019)

Our Study: **Results**

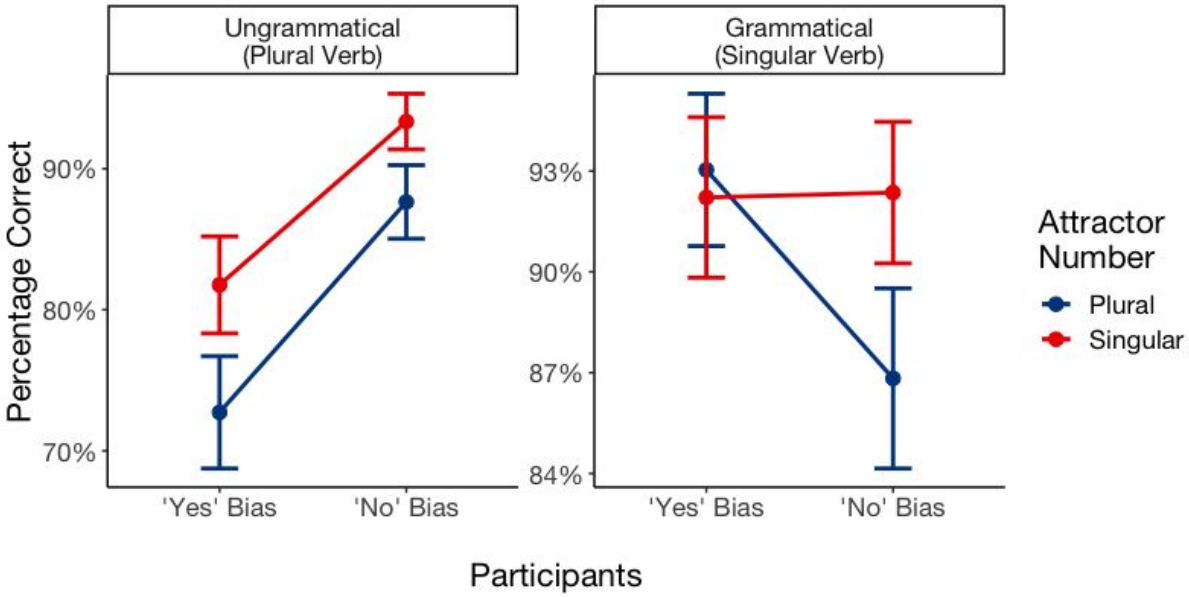
- Attraction in ungrammatical sentences independent of response bias



(Cousineau, 2005; Morey, 2008)

Our Study: **Results**

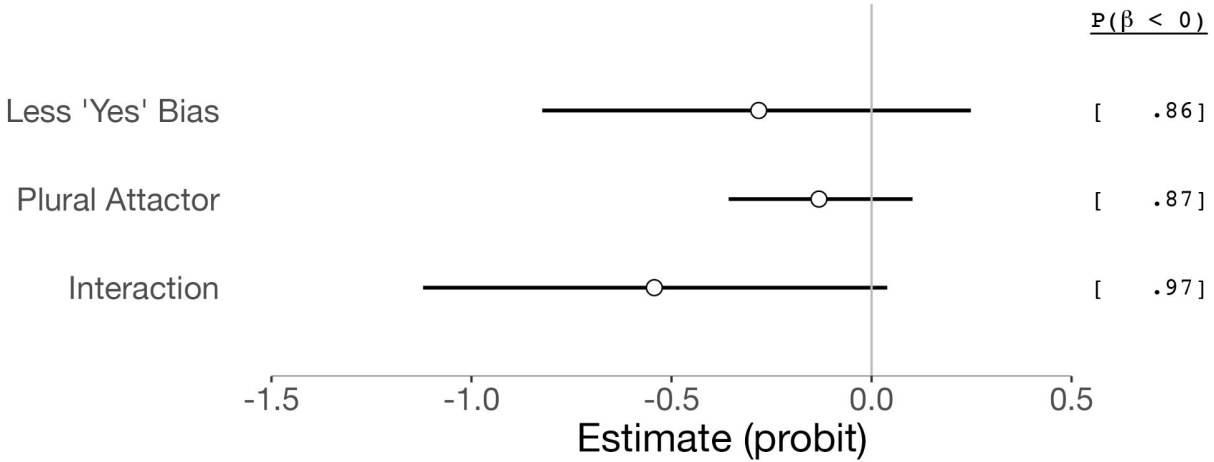
- Attraction in **grammatical** sentences as a function of bias



(Cousineau, 2005; Morey, 2008)

Our Study: **Bayesian Model**

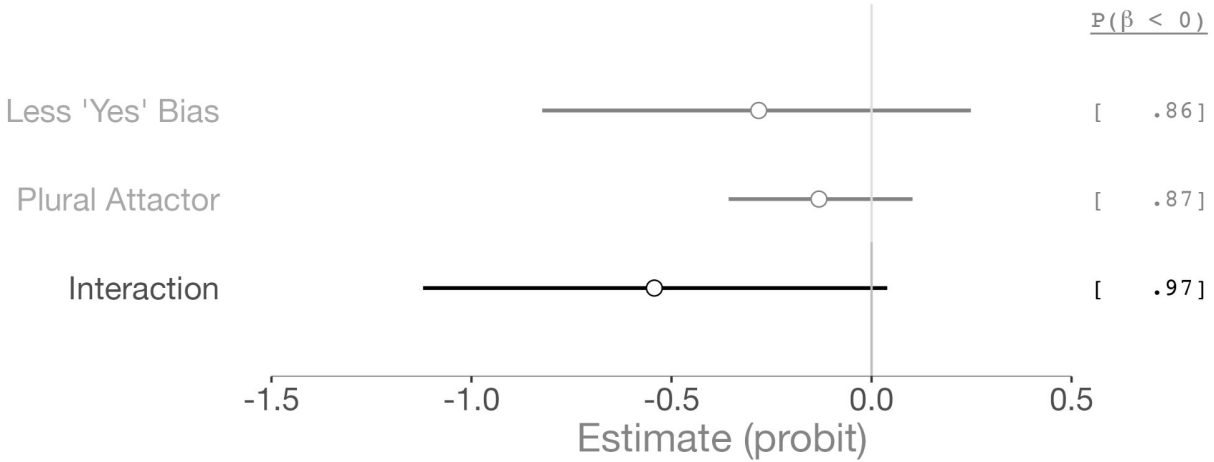
- Verified our results with a maximal Bayesian GLM.
  - Fitted to **grammatical** sentences
  - No main effect of PLURAL ATTRACTOR
  - P(INTERACTION < 0) = 0.97



(Gelman & Hill, 2007; Barr et al., 2013; Nicenboim & Vasishth, 2016; Kruschke, 2018)

Our Study: **Bayesian Model**

➔ The effect of plural attractor is more pronounced in people with less “yes” bias in grammatical sentences



(Gelman & Hill, 2007; Barr et al., 2013; Nicenboim & Vasishth, 2016; Kruschke, 2018)

Our Study: **Findings**

- Replicated theoretically significant findings of Hammerly et al.

Our Study: **Findings**

- Replicated theoretically significant findings of Hammerly et al.
- Grammaticality asymmetry can be explained via response bias

Our Study: **Findings**

- Replicated theoretically significant findings of Hammerly et al.
- Grammaticality asymmetry can be explained via response bias
- No need for a strong preference of retrieval accounts

Roadmap

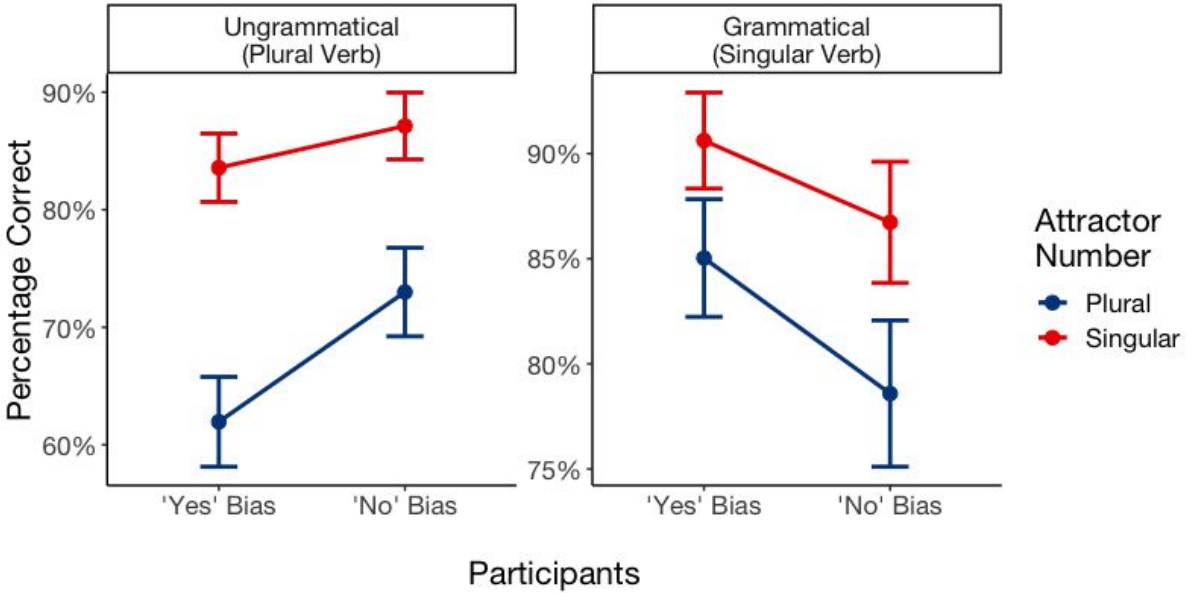
- ✔ 1. Hammerly et al. (2019) and Bias Calculation Problem
- ✔ 2. Turkish Experiment
- ❑ 3. Re-analysis of Hammerly et al. (2019)
- ❑ 4. Meta-analysis

Reanalysis of Hammerly et al. (2019)

- Grouped participants according to their bias in fillers

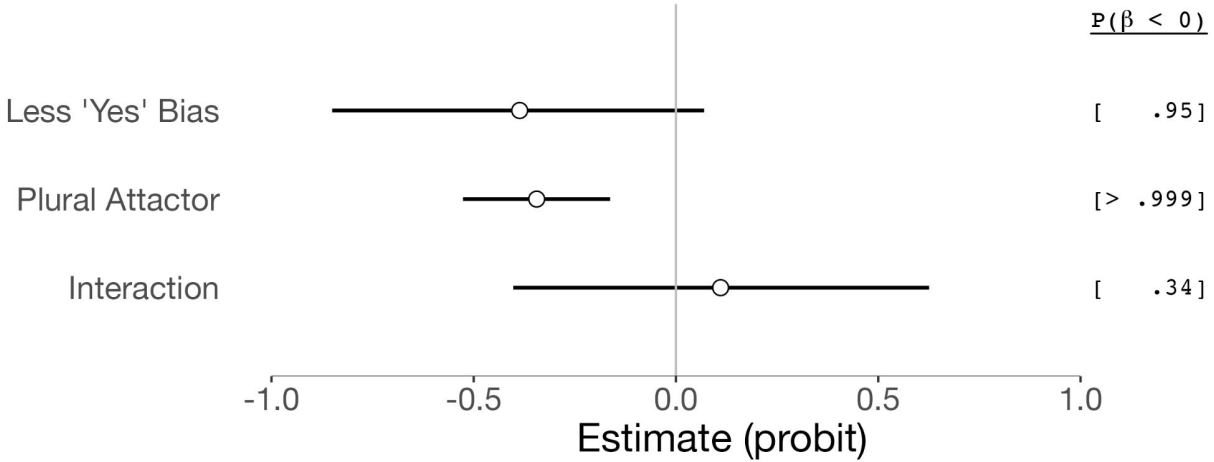
Reanalysis of Hammerly et al. (2019): **Results**

- Grouped participants according to their bias in fillers
- Attraction in both grammatical and ungrammatical sentences **independent of response bias**



Reanalysis of Hammerly et al. (2019): **Bayesian Model**

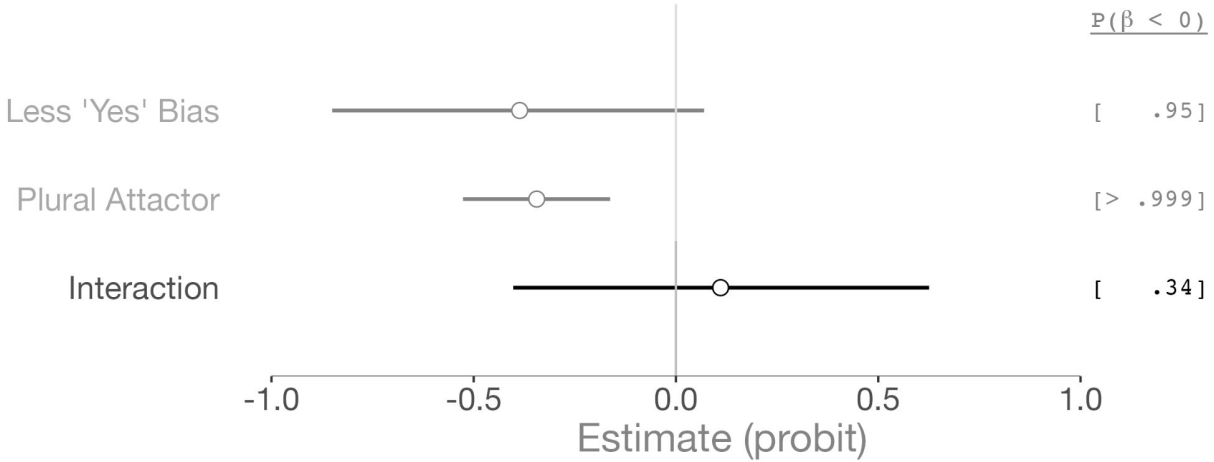
- Verified lack of bias effect with a maximal Bayesian GLM.
  - Fitted to **grammatical** sentences
  - Clear main effect of PLURAL ATTRACTOR, $P(\beta < 0) > 0.999$
  - No interaction, $P(\text{INTERACTION} < 0) = 0.34$



(Gelman & Hill, 2007; Barr et al., 2013; Nicenboim & Vasishth, 2016; Kruschke, 2018)

Reanalysis of Hammerly et al. (2019): **Bayesian Model**

➔ Having weaker “yes” bias did not affect the contribution of the plural attractor



(Gelman & Hill, 2007; Barr et al., 2013; Nicenboim & Vasishth, 2016; Kruschke, 2018)

Reanalysis of Hammerly et al. (2019): **Findings**

- Attraction in grammatical sentences surfaces even with “yes” bias
- Bias in filler was not reflected in experimental items
- Original findings does not reflect participants’ a priori bias

Roadmap

✔ 1. Hammerly et al. (2019) and Bias Calculation Problem

✔ 2. Turkish Experiment

✔ 3. Re-analysis of Hammerly et al. (2019)

❑ 4. Meta-analysis

Meta-analysis

- What about other experiments
  that does not have bias manipulation?

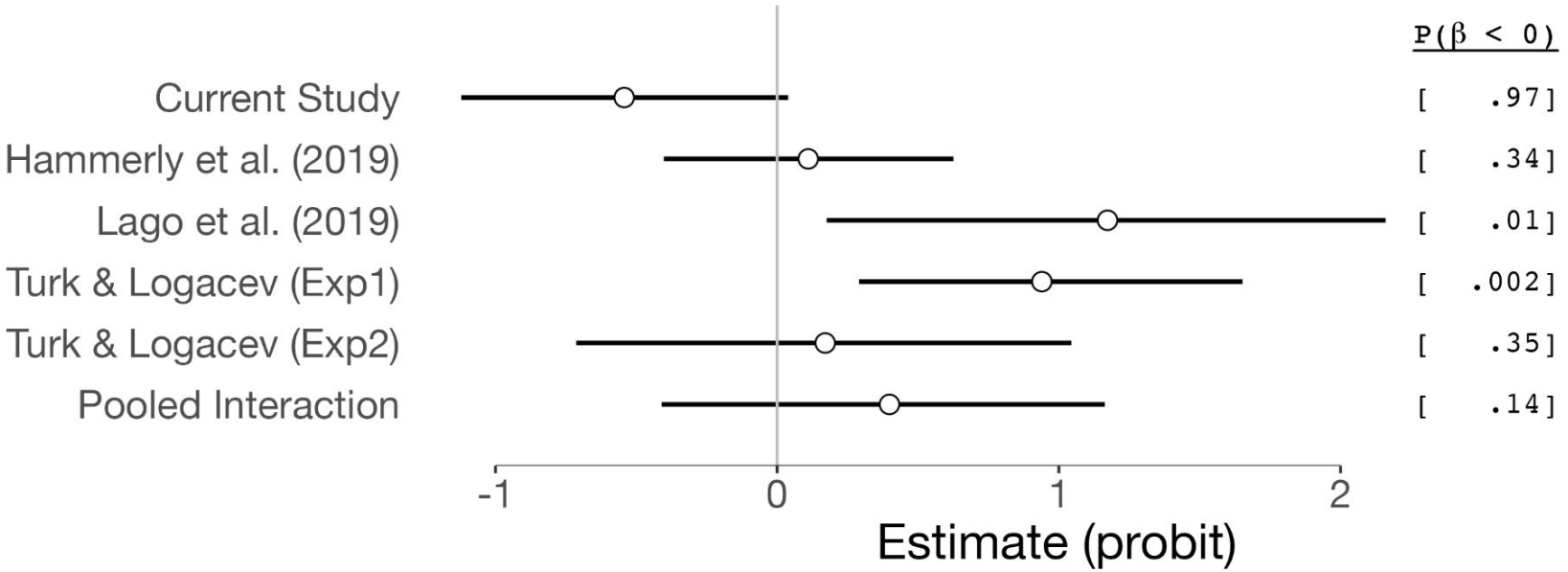
Meta-analysis: **Bayesian Model Details**

- What about other experiments without bias manipulation?

  - Conducted a multilevel Bayesian meta-analysis
  - Fitted to correct responses to **grammatical** sentences
  - Predictors:
    - Experiments, subjects, and items as random effects
    - Bias Value (calculated using fillers)
    - Attractor Number
    - The interaction
    - Trial number

(Berkey et al., 1998; Gelman & Hill, 2007; Gleser & Olkin, 2009; Konstantopoulos, 2011)

Meta-analysis: **By-experiment Interaction Posteriors**



➔ Cannot say grammaticality asymmetry reflects response bias, it sometimes does.

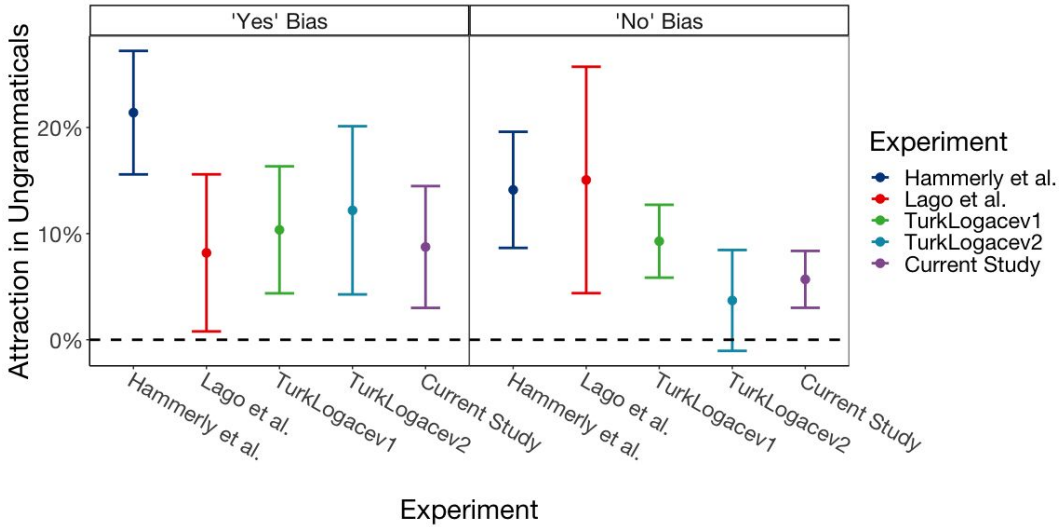
(Gelman & Hill, 2007; Barr et al., 2013; Nicenboim & Vasishth, 2016; Kruschke, 2018)

Roadmap

- ✔ 1. Hammerly et al. (2019) and Bias Calculation Problem
- ✔ 2. Turkish Experiment
- ✔ 3. Re-analysis of Hammerly et al. (2019)
- ✔ 4. Meta-analysis

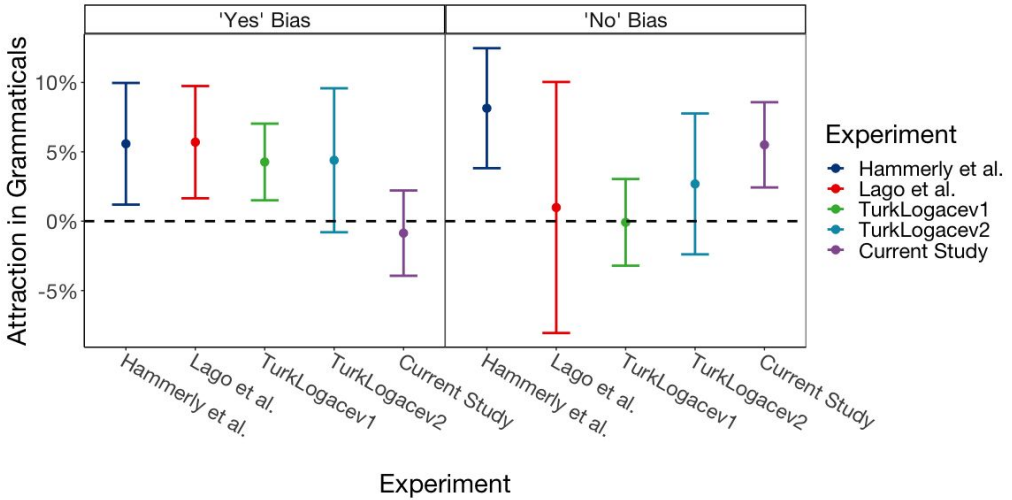
Take-Home Messages

- The effect in ungrammatical sentences: **Persistent**



Take-Home Messages

- The effect in ungrammatical sentences: **Persistent**
- The effect in grammatical sentences? **Finicky**



Take-Home Messages

- The effect in ungrammatical sentences: **Persistent**
- The effect in grammatical sentences? **Finicky**
- Asymmetry is still important.

Take-Home Messages

- The effect in ungrammatical sentences: **Persistent**
- The effect in grammatical sentences? **Finicky**
- Asymmetry is still important.
- Grammaticality impacts attraction effects' variability.

Selected References

Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, *23*(1), 45–93.

Eberhard, K. M., Cutting, J. C., & Bock, K. (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review*, *112*(3), 531–559.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

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.

Jäger, L. A., Mertzen, D., Van Dyke, J. A., & Vasishth, S. (2020). Interference patterns in subject-verb agreement and reflexives revisited: A large-sample study. *Journal of Memory and Language*, *111*, 104063.

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. In *Linguistic Approaches to Bilingualism* (Vol. 9, Issue 3, pp. 398–426). John Benjamins.

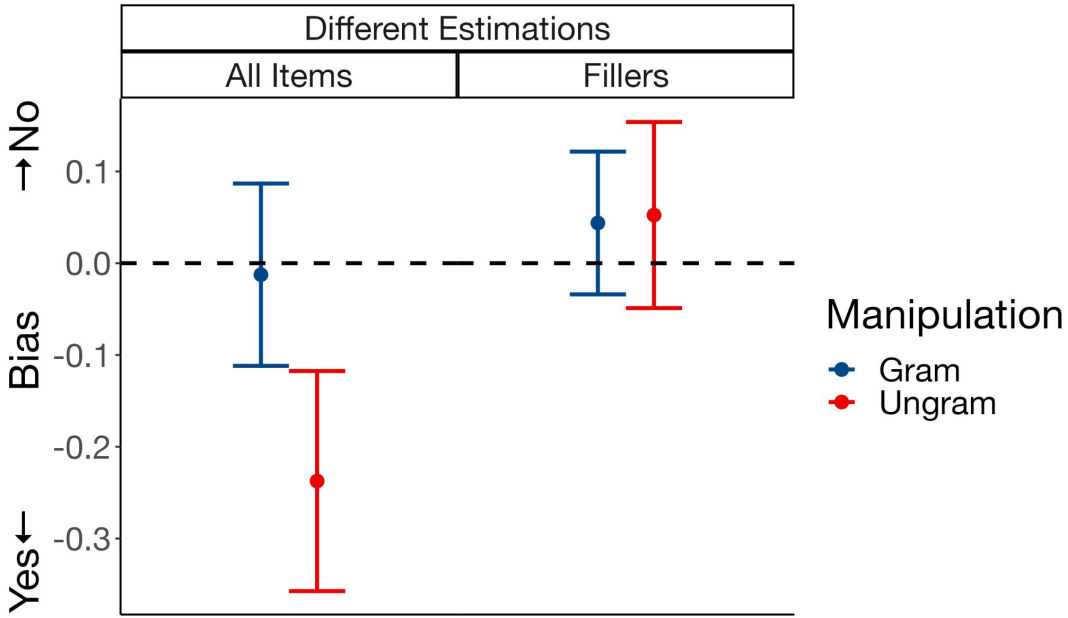
Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Reason*, *4*(2), 61–64.

Nicenboim, B., & Vasishth, S. (2016). Statistical methods for linguistic research: Foundational ideas—Part II. *Language and Linguistics Compass*, *10*(11), 591–613.

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.

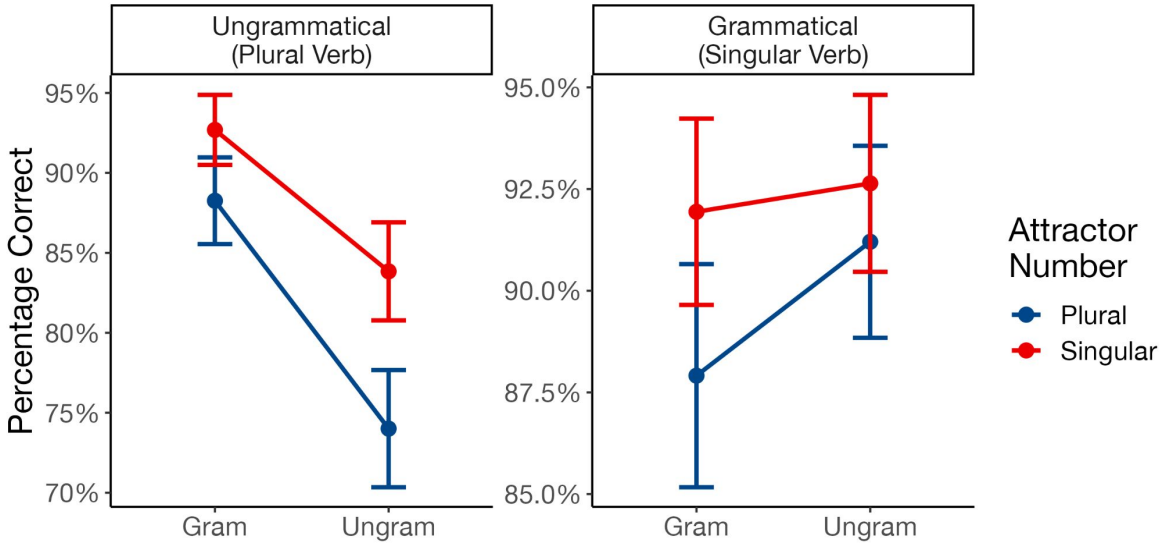
Appendix A: Our Exp with Different Bias Estimations

- In “Towards Ungrammatical” Bias, we had more ‘yes’ bias.



Appendix B: Our Exp with Original Manipulation Grouping

- Even when we look at “Ungram” as more ‘yes’ bias,
  and “Gram” as equi-bias, our results do not follow from bias-informed theories



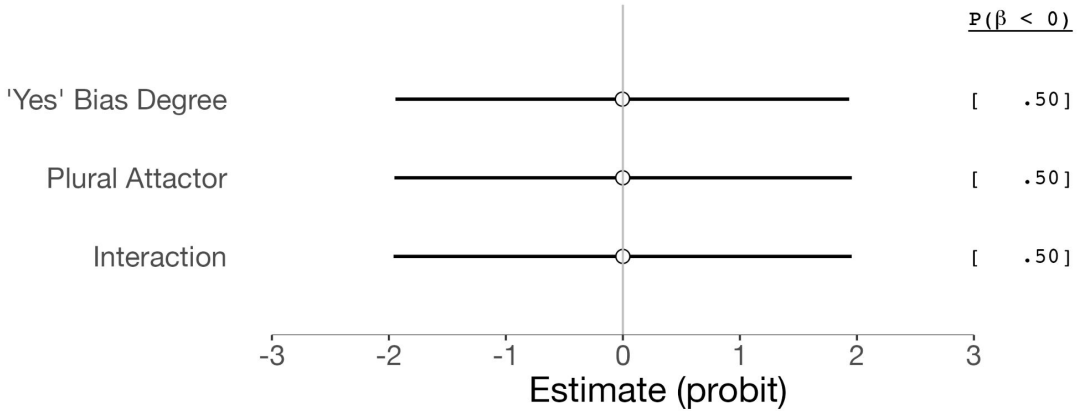
Appendix C: Model Specifications

- Packages: `cmdstanr` and `brms`
- Priors: Reasonable Agnostic Priors

$$\text{Intercept} \sim Normal(0,1)$$
$$\beta \sim Normal(0,1)$$
$$\sigma \sim Normal(0,1)$$
$$\rho \sim LKJ(2)$$

- Sum contrast coding

Bias is continuous, no coding.
+0.5 for Plural Attractor
-0.5 for Singular Attractor
(+0.5 for Ungrammatical)
(-0.5 for Grammatical)



Appendix C: Model Specifications

- Formula & Predictors:
  - Continuous Response Bias Value
  - Attractor Number
  - The interaction
  - Trial Number (log)

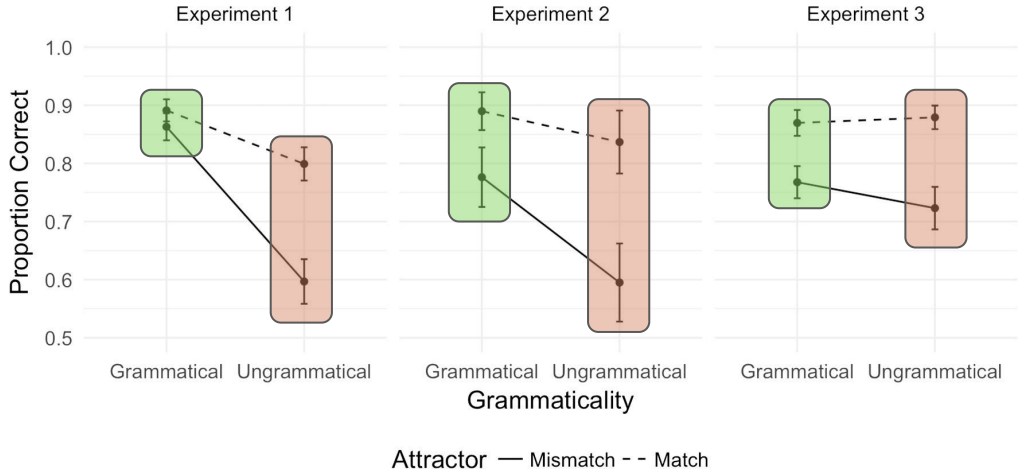
```
response_yes ~ bias * attractor_number + log_trial +
                    (bias * attractor_number + 1 | subject) +
                    (attractor_number + log_trial + 1 | item)
```

Appendix D: Bias Estimations

- How to calculate bias?

$$-\frac{Z(\text{Hit Rate}) + Z(\text{False Alarms})}{2}$$

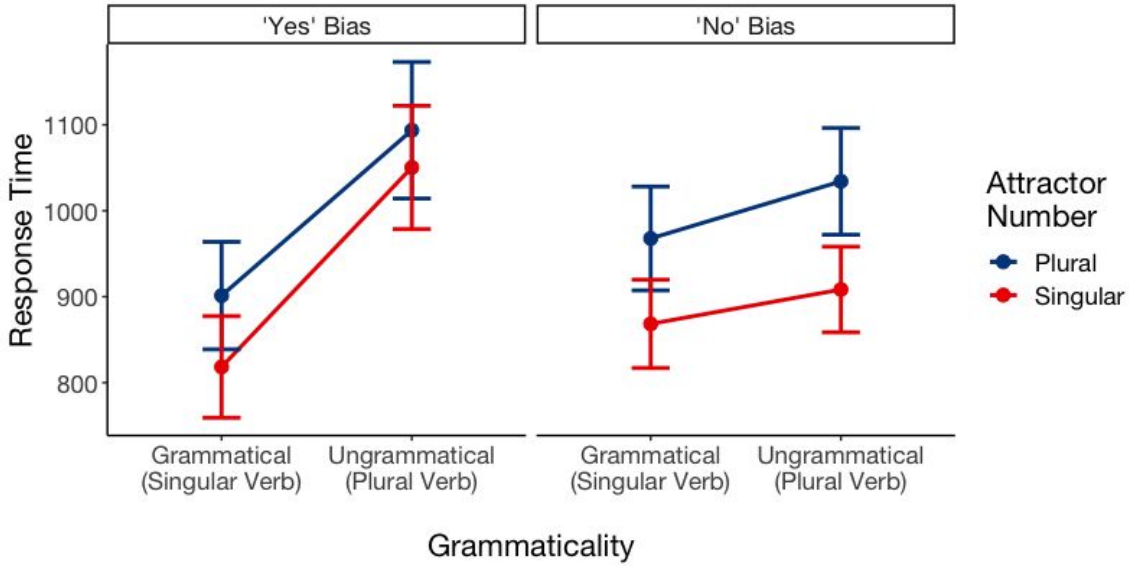
- What happens when we use ALL Items?



(Macmillan & Creelman, 2005)

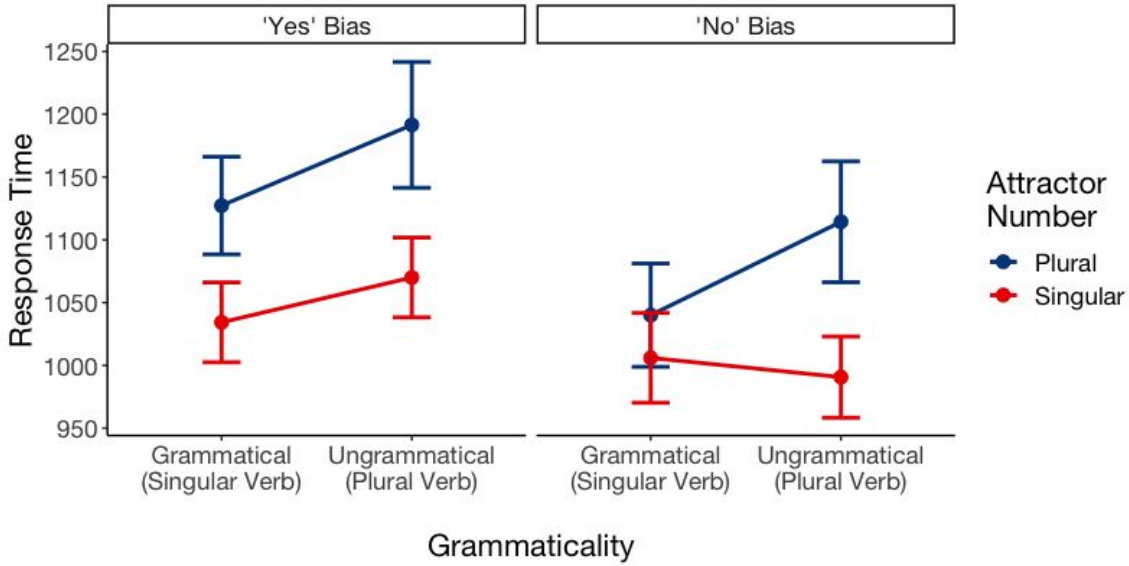
Appendix E: RTs?

- What do we see in vanilla attraction experiments?
  - Overall slowdown for ungrammaticals
  - Additional slowdown for plurals in ungrammaticals
- What does bias-informed analysis expect?
  - No slowdown for ungrammaticals
  - Same contribution from plurals in both grammatical and ungrammatical



Appendix E: RTs?

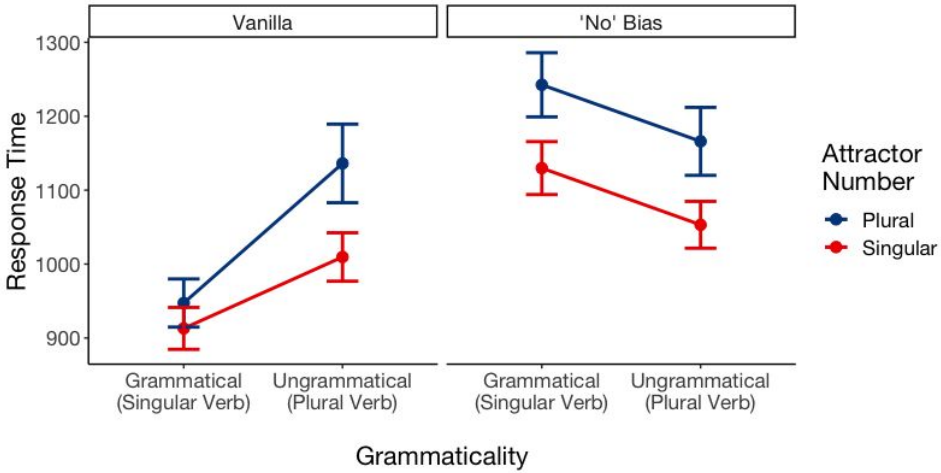
- Our experiment RTs close to prediction, but not quite.
- Hammerly et al's? RTs look close to the prediction as well.



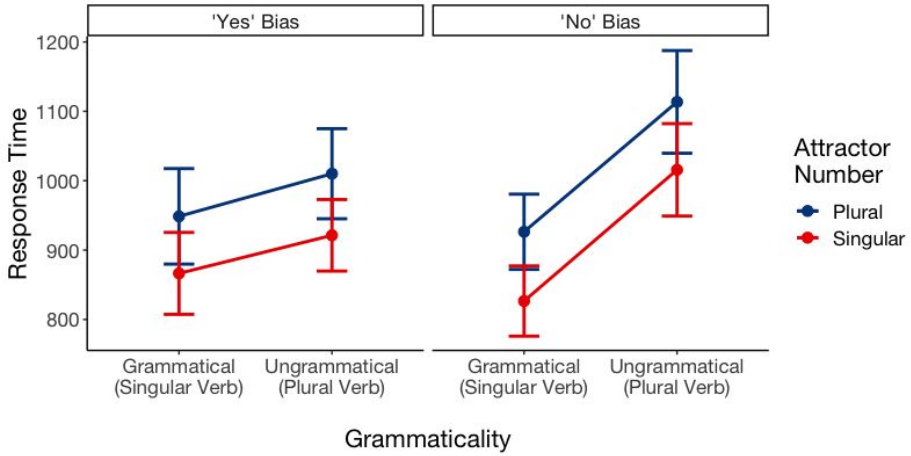
Appendix E: RTs?

- Maybe our bias estimation is actually not good?
  - Hammerly bias(all) predictions ✅
  - Our bias(all) predictions ❌
- Their bias estimation: Their Acceptability & RT ✅, Our Acceptability & RT ❌
- Our bias estimation: Their Acceptability & RT ❌, Our Acceptability & RT ✅

Hammerly et al.

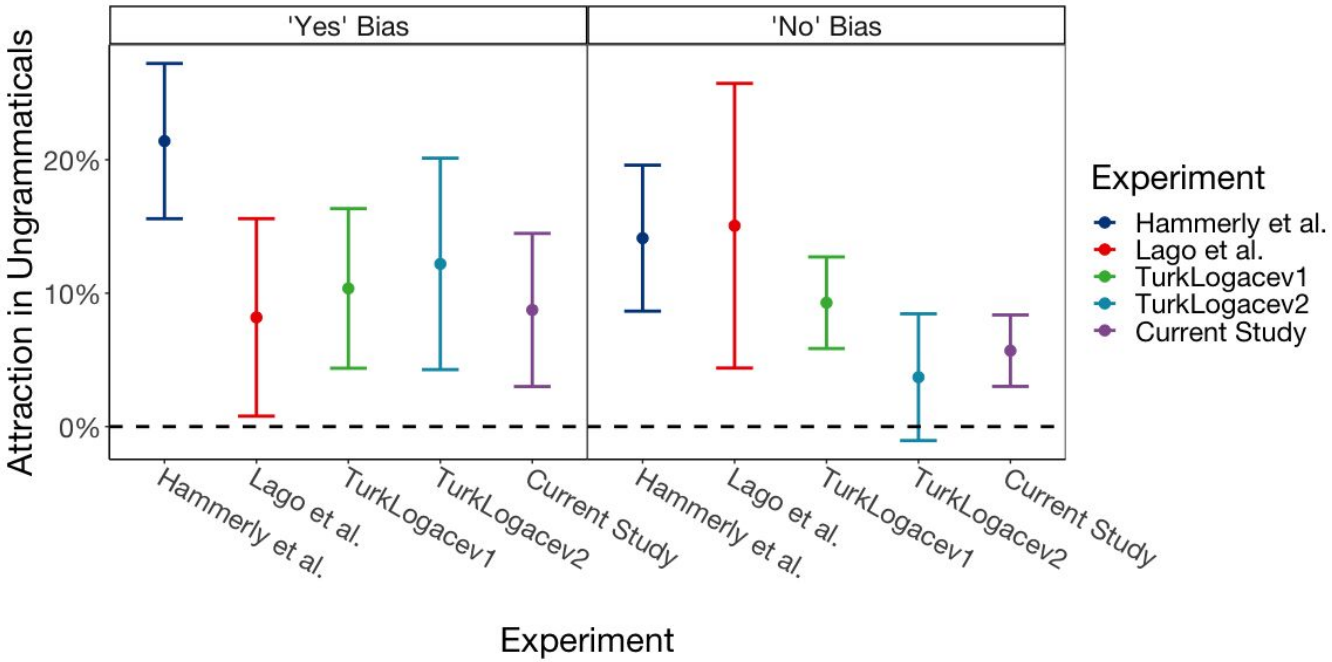


Our Experiment



Appendix F: Stable Attraction in Ungrammaticals

- Attraction effects are persistent in ungrammatical sentences.
  - independent of response bias and experiment.



Appendix G: Finicky Attraction in Grammaticals

- Attraction effects vary in grammatical sentences.