

Evaluating linking functions for eye movements in the visual world

Casey Pablo Butcher (cbu42@stanford.edu) and Prof. Judith Degen (jdegen@stanford.edu)

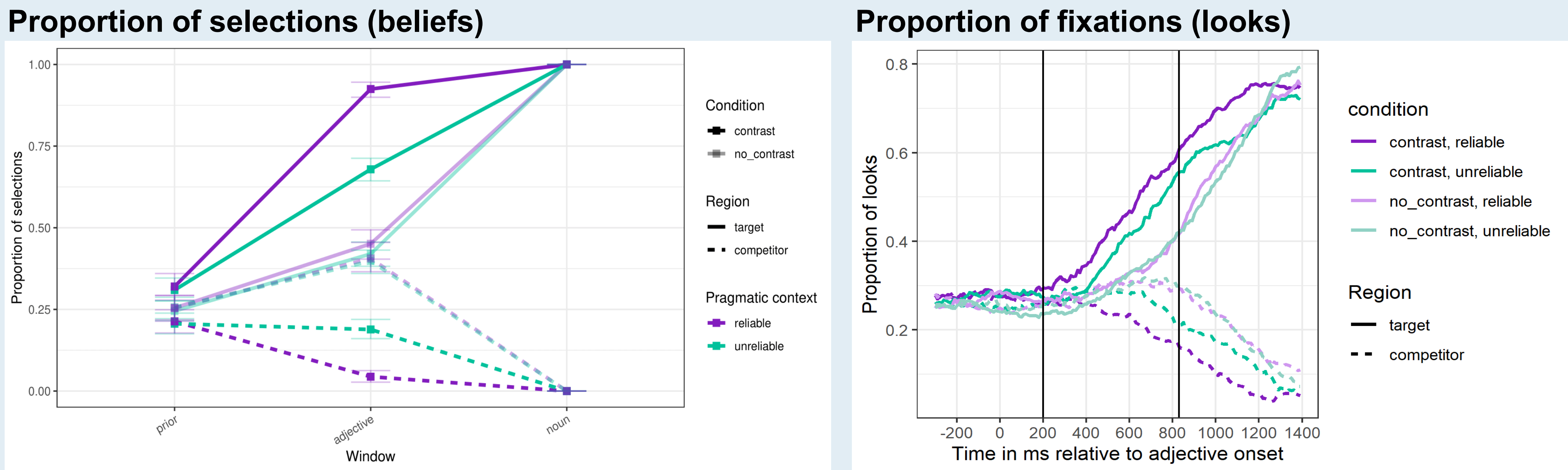
Department of Symbolic Systems, Stanford University // Department of Linguistics, Stanford University



Introduction

- In visual world eye-tracking paradigms (VWP), a common linking assumption underlying referential tasks is that within a window of interest, the proportion of looks to a referent reflects the degree of belief that that referent is the target. (Allopenna et al., 1998)
- Formally, for a referent *r* and an utterance *u*:  
 $p_{\text{empirical}}(r \mid u) \propto p_{\text{belief}}(r = \text{target} \mid u)$ .
- This linking assumption has been tested only few times before, with conflicting results as to whether it holds. (Qing, Lassiter, & Degen, 2018; Degen, Kursat, & Leigh, 2021)
- To test the linking assumption, we replicated the second experiment in Ryskin et al., 2019 using an incremental decision task to derive a distribution of explicit beliefs (referent selections) to correlate with eye movement data.
  - We hypothesize that we will find a positive correlation between looks to a referent and degree of belief (as indicated by selections in the referential task) that that referent is the target.
- We also tested the original paper's hypothesis that listeners suspend contrastive inference after exposure to pragmatically unreliable speakers. (Ryskin et al., 2019)

Result 1: Replicating Ryskin et al. (2019) – Is a listener’s pragmatic inference susceptible to a speaker’s infelicitous use of scalar adjectives?



At the adjective stage, fewer participants selected the target when their speaker (instructions) was unreliable. This is especially true in the contrast condition. Fixations, unlike selections/clicks, are plotted continuously (every 10 ms). Resembles the graph of selections even within condition.

- Listeners indeed suspend contrastive inference after exposure to improper use of scalars (as seen in the adj. window, where contrast is relevant).
- We fit a mixed effects logistic regression model to predict target selections at each window. The table below summarizes critical effects:

Predictor	Prior	Adj.
cond:	-	-
pragContext:	0	+
cond:pragContext:	0	-

“+” indicates a significant positive effect.  
“-” indicates a significant negative effect.  
“0” indicates no evidence of effect.  
Cond predictor refers to contrast condition.  
Noun window excluded since all select target (no variance to fit).

- Unexpectedly, we observe a contrastive inference effect in prior window (effect: -0.416, *p*<0.001).
- Negative main effect of contrast condition (fewer target selections in the absence of contrast) observed in adj. Positive effect of pragmatic context (more target selections with reliable speaker). Output of adj model to right.

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [‘glmerMod’]
Family: binomial ( logit )
Formula: target ~ ccond * cpragContext + (1 + ccond | workerid) + (1 | noun)
Data: dc_adj

Random effects:
Groups Name Variance Std.Dev. Corr
workerid (Intercept) 0.3573 0.5977
ccond 0.6212 0.7882 -0.90
noun (Intercept) 1.0043 1.0022
Number of obs: 2159, groups: workerid, 63; noun, 40

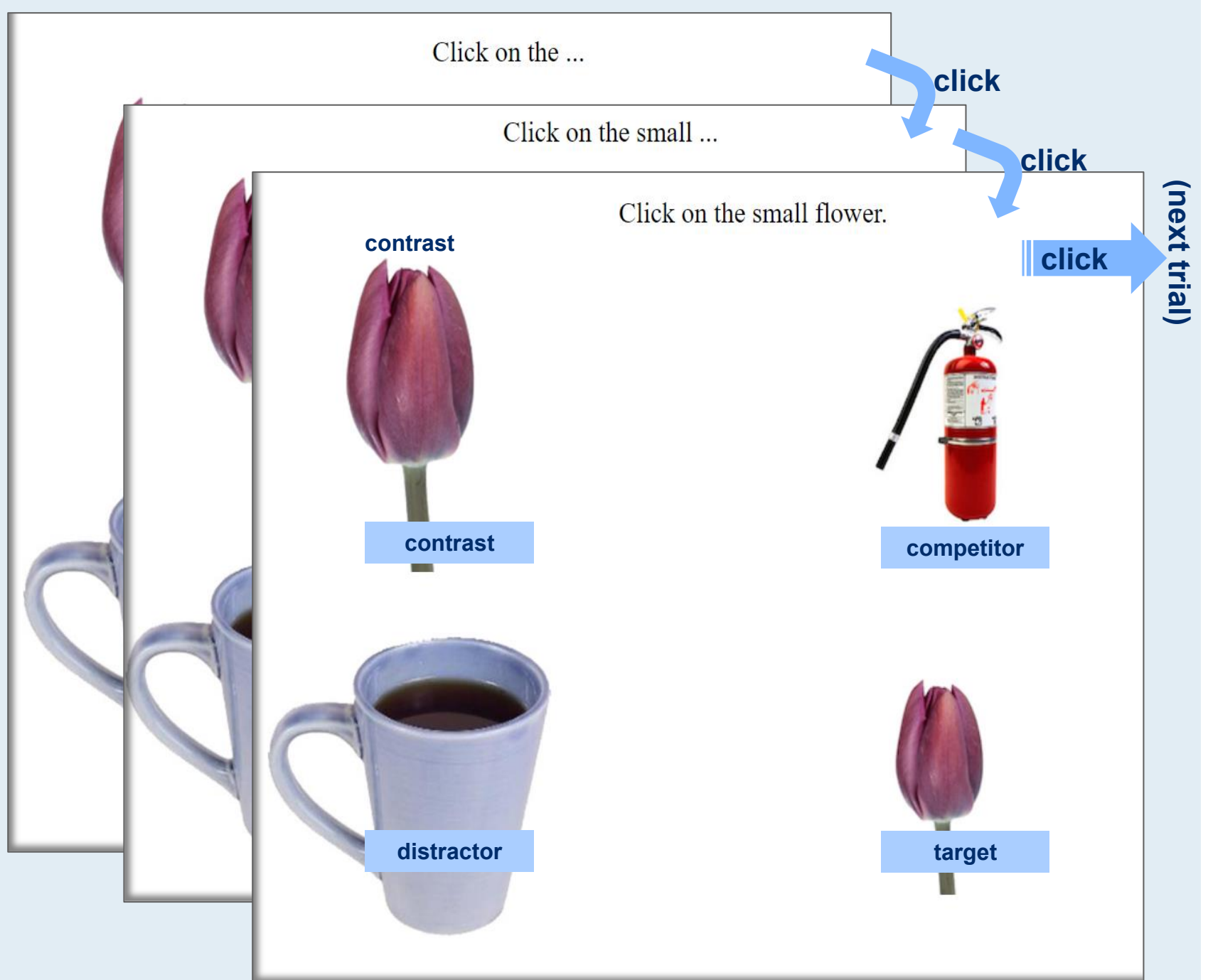
Fixed effects:
(Intercept) 1.2426 0.1935 6.420 1.36e-10 ***
ccond -2.3098 0.3682 -6.274 3.53e-10 ***
cpragContext 1.0129 0.2163 4.684 2.82e-06 ***
ccnd:cpragContext -2.1128 0.3618 -5.839 5.24e-09 ***
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
(Intr) ccond cpragcn
ccond -0.232
cpragContext 0.148 -0.143
ccnd:cpragcn -0.164 0.169 -0.688

Output of the mixed effects model for the adjective window.
```

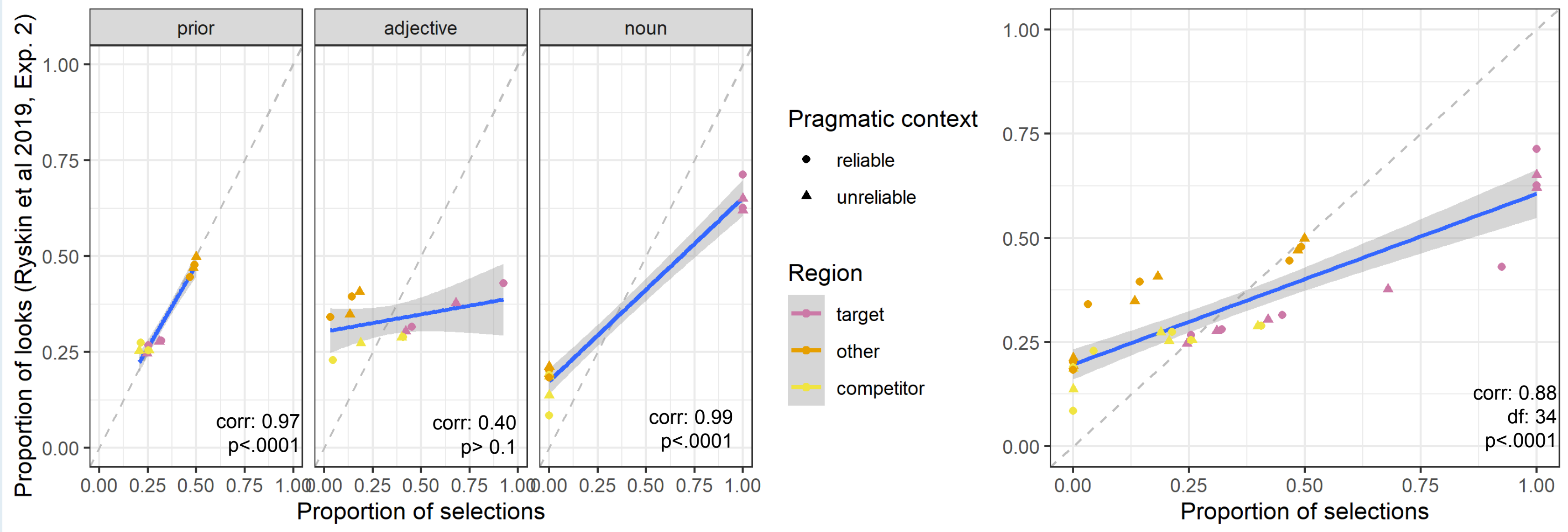
Methods: Incremental Decision Task

- Task: Click on the referent described by a partial utterance.**
- With each click, the partial utterance unfolds until complete (see right).
  - Critical trials were modified with an adjective. (e.g. “Click on the big pickle.”)
- Independent variables:
- Contrast condition (present/absent).
  - Pragmatic context (reliable/unreliable) randomly assigned to a participant at the onset of experiment.
- Participants (N=63):
- Excluded non-native English speakers and participants with more than a 5% error rate.
- Trials (N=300):
- 40 critical trials; 80 train trials; 180 filler trials.
  - Excluded trials in which the participant ultimately clicks the wrong target.

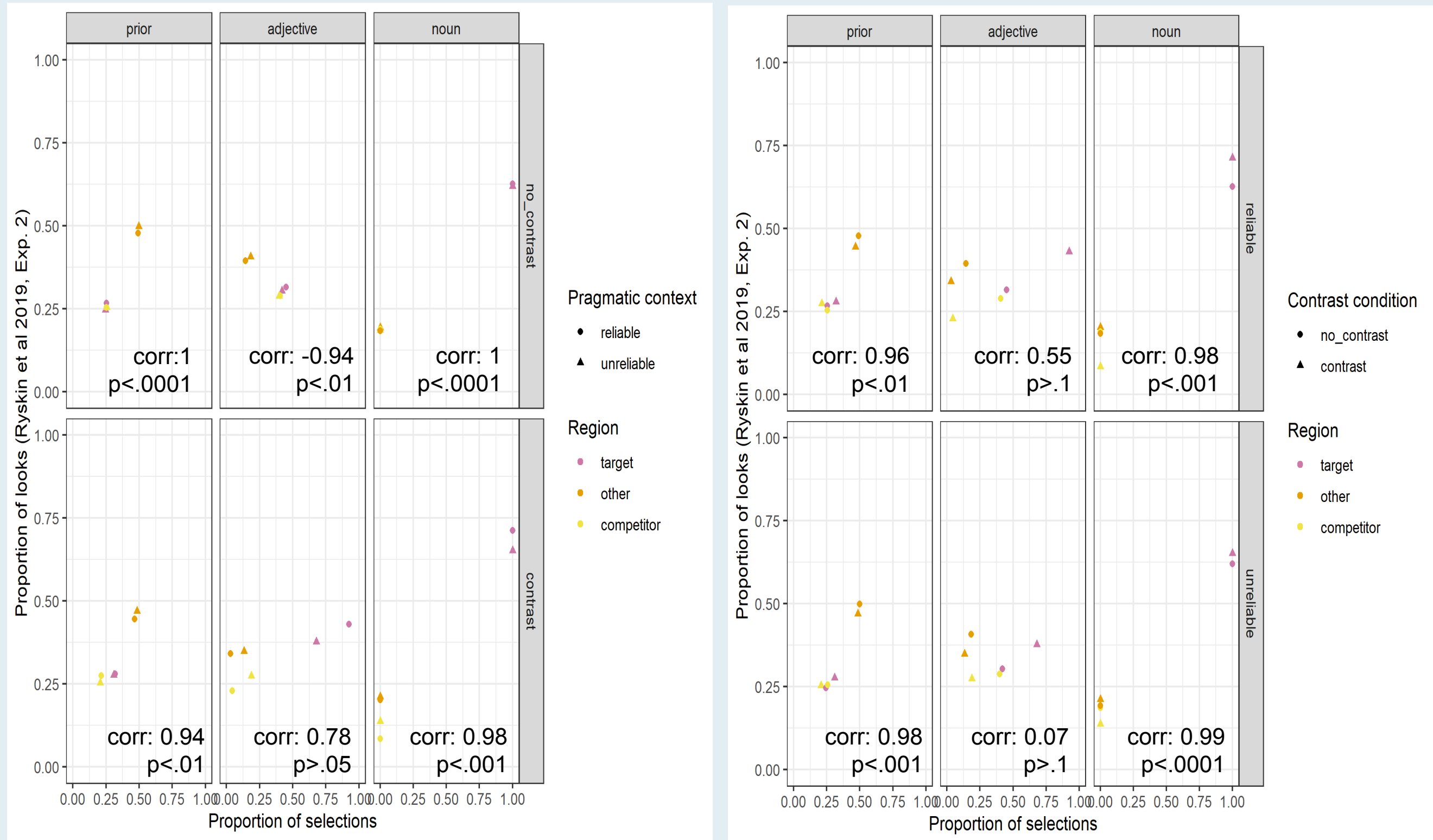


The critical trial above is of the contrast, reliable, and modified conditions: a large flower contrasts the target “small flower”; the use of “small” in the instructional utterance correctly implicates the target; the utterance uses an adjective.

Result 2: Is the proportion of looks to a referent indeed reflective of the listener’s belief that the referent is the speaker’s target?



Effect of explicit beliefs varies by window: insignificant in adjective window.



```
Call:
lm(formula = prop_looks ~ cprop_selections + cprop_selections:window +
  cprop_selections:Region, data = tomodel)

Coefficients:
(Intercept) 0.330641 0.009005 36.718 < 2e-16 ***
cprop_selections 0.929733 0.115502 8.049 5.51e-09 ***
cprop_selections:windowadjective -0.815819 0.115506 -7.063 7.48e-08 ***
cprop_selections:windownoun -0.438121 0.110413 -3.968 0.000417 ***
cprop_selections:Regionother -0.127463 0.065863 -1.935 0.062430 .
cprop_selections:Regioncompetitor 0.050725 0.073823 0.687 0.497291
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.03953 on 30 degrees of freedom
Multiple R-squared: 0.9406, Adjusted R-squared: 0.9307
F-statistic: 94.98 on 5 and 30 DF, p-value: < 2.2e-16
```

- Significant correlation between looks and beliefs (0.88).
- We fit a linear model to predict proportion of looks (summary below).
- Yielded a main effect of explicit beliefs in the predicted direction (estimate = 0.929733).
- Effect of explicit beliefs varies by region: beliefs have less explanatory power for distractor looks.
- Overall, the correlation plots show greater "other" looks across the board than predicted by selection proportions.

Conclusions

- Listeners suspend their contrastive inference when exposed to improper scalar usage by speakers, especially when an item contrasting the target is present (e.g. big flower and small flower present at once in referent set).
- Explicit beliefs (selections) that a referent is the target are a powerful predictor of the proportion of fixations to that referent, supporting the linking assumption proposed by Allopenna et al.
- An additional analysis of the linking assumption on a by-item basis could further substantiate its efficacy in VWP.

References:

Allopenna, P.D., Magnuson, J.S., & Tanenhaus, M.K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *JML*.

Degen, J., Kursat, L., & Leigh, D. (2021) Seeing is believing: testing an explicit linking assumption for visual world eye-tracking in psycholinguistics. *Proceedings of the Annual Meeting of the Cognitive Science Society*.

Qing, C., Lassiter, D., & Degen, J. (2018). What do eye movements in the visual world reflect? A case study from adjectives. *Cognitive Science*.

Ryskin, R., Kurumada, C., & Brown-Schmidt, S. (2019). Information integration in modulation of pragmatic inferences during online language comprehension. *Cognitive Science*.