Speak before you listen: Pragmatic reasoning in multi-trial language games

Les Sikos, Noortje Venhuizen, Heiner Drenhaus, and Matthew W. Crocker (Saarland University)

Background

- Rational Speech Act (RSA) theory [1] has been successfully applied in many communicative settings
- Follow-up studies using **one-shot web-based** language games suggest that listeners may not behave as pragmatically as originally suggested [2-5]

Goal:

Investigate whether pragmatic behavior is enhanced through increased exposure to the task:

- Increased number of trials
- Exposure to both speaker and listener task

Example Critical Stimulus and Tasks

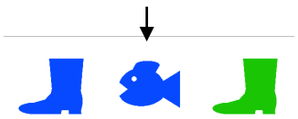

Figure 1: Example of a critical visual context, which contains a *pragmatic referent* (blue boot), a *color competitor* (blue fish), and a *shape competitor* (green boot). The arrow only appears in the Speaker task.

- **Speaker:** You want someone to pick out the object indicated by the arrow, but you can only use one word … Which word would you say? [“**blue**” / “**fish**”]
- **Listener:** You hear the word “**blue**” / “**fish**” … Which object do you think is being referred to?
- **Salience:** You cannot understand the message … Which object do you think is being referred to?

Methods and Results

2x2 Design:

- **Exposure**: First-trial vs All-trials
- **Block Order**: Listener-first vs Speaker-first

- **Listener** block: 24 trials (6 critical, 12 fillers, and 6 Salience trials)
- **Speaker** block: 18 trials (6 critical, 12 filler)

To replicate previous findings from one-shot language games [4,5]:

- First Speaker trial is a critical trial
- First Listener trial is a critical shape-word trial

Behavioral Results

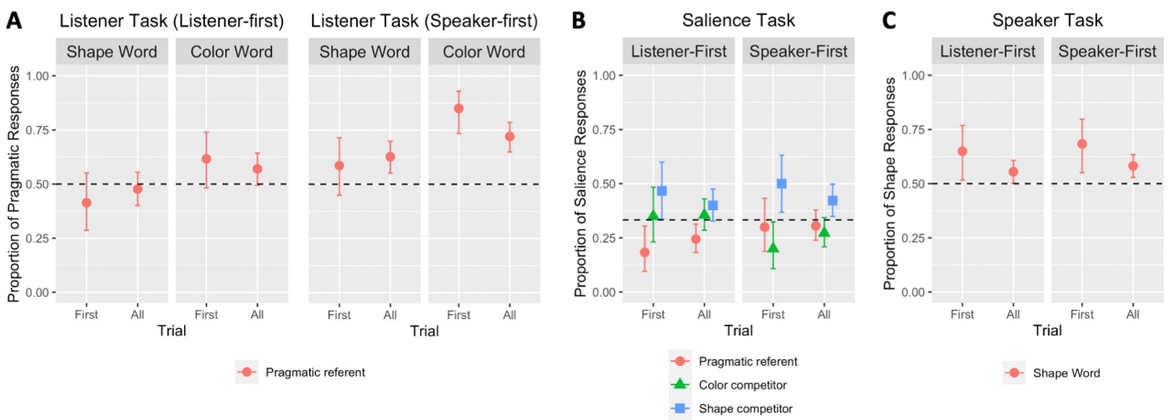

Figure 2. Human Judgments in the Listener task (**A**), the Salience task (**B**), and the Speaker task (**C**). We plot the proportion of responses by block order (Listener-first, Speaker-first) and the observed word (shape, color). Error bars represent binomial 95% confidence intervals and the dashed lines represent chance.

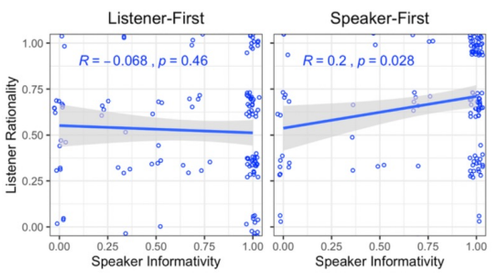

Figure 3: Correlation of Speaker Informativity and Listener Rationality (R: Spearman’s rho correlation).

Model Evaluation

RSA: $P(target|word,C) = \frac{P(word|target,C)P(target)}{\sum_{r \in C} P(w|r,C)P(r)}$ **LL:** $P(target|word,C) = \frac{[\![word]\!](target)P(target)}{\sum_{r \in C}[\![word]\!](r)P(r)}$

		Listener-First					Speaker-First				
Dataset	**Model**	r	R^2_{adj}	t	p	cocor-p	r	R^2_{adj}	t	p	cocor-p
First-trial	RSA	0.90	0.79	6.59	$< .0001$	$\} .43$	0.99	0.98	23.07	$< .0001$	$\} < .0001$
	LL	0.94	0.87	8.65	$< .0001$		0.85	0.69	5.08	$< .001$	
All-trials	RSA	0.99	0.97	26.64	$< .0001$	$\} < .0001$	0.98	0.95	21.56	$< .0001$	$\} < .0001$
	LL	0.90	0.79	9.46	$< .0001$		0.80	0.62	6.19	$< .0001$	

Table 1. Results from the Listener task comparing RSA to the baseline literal listener model (LL). *r*: Pearsons’s correlation; cocor-*p*: *p*-value for comparison of overlapping dependent correlations.

Discussion and Conclusions

- We replicate previous findings [2-5] that listeners show limited pragmatic behavior in the one-shot task
- Limited evidence that increasing number of trials results in more pragmatic responses
- Listeners show increased pragmatic reasoning after first playing the role of the speaker
- In the Speaker-first condition (only), a participant’s tendency to be an informative speaker predicts their degree of pragmatic behavior as a listener

- Results confirm the observation put forward by [5] that the high correlation between RSA’s predictions and listener behavior reported in one-shot experiments [e.g., 1] is primarily driven by non-pragmatic factors
- However, our results also suggest that the role of RSA’s pragmatic component, which reasons about informative speakers, is particularly enhanced when listeners have experience as the speaker.

Good speakers become better listeners

References

[1] Frank & Goodman (2012). *Science, 336* (6084).

[2] Frank, Emilsson, Peloquin, Goodman & Potts (2016). *PsyArXiv (f9y6b)*.

[3] Franke & Degen (2016). *PLOS ONE, 11* (5).

[4] Qing & Franke (2015). In: *Bayesian natural language semantics and pragmatics*.

[5] Sikos, Venhuizen, Drenhaus & Crocker (2021). *PLOS ONE, 16*(3).