If you don't have anything nice (or interesting) to say, don't say anything at all

Hannah Rohde & Michael Franke

Abstract

Problem: Many studies emphasize the role of real-world knowledge in language processing. Such emphasis, however, risks sidestepping another key contribution of communication – its use as a channel across which speakers convey newsworthy and informative messages.

Proposal: To revisit the role of *unpredictability* in language, we contrast participants' estimates of the knowledge and likely utterances of an individual. Although plausible situations may be predictable as beliefs about the real world, they are not necessarily predictable as messages for an individual to choose to convey.

Hannah **thinks** that Andy drank _____ cups of coffee last week.
Hannah **announced** to me that Andy drank _____ cups of coffee last week.

Results: Study1 elicits fill-in-the-blank responses, which are shown to pattern with previously collected real-world estimates (*Andy is a man from the US. How many cups of coffee do you think Andy drank last week?* Schöller & Franke 2017), but condition (**think**/**announce**) yields no main effect or interaction. In Study2, participants' forced-choice responses show the predicted effect of condition, whereby **announce** yields higher values than **think**. Intuitively, "good" sentences describe situations that are suitably plausible while still being rare enough to be interesting.

1. Goal

We test the role of **newsworthiness** in language processing by distinguishing between expectations about:
- Speakers' beliefs [e.g., real-world knowledge]
- Speakers' choice of what to say [content selection]

2. Real-world knowledge

Surprisal at implausible words (Kutas & Hillyard 1980; Hagoort et al. 2004)

The Dutch trains are — yellow / white / sour

yellow < {white, sour}

Surprisal tuned to comprehenders' knowledge about the world

Language, Cognition and Neuroscience ›

Harry Potter and the Chamber of *What*?: the impact of what individuals know on word processing during reading

Melissa Troyer ✉ & Marta Kutas
Received 13 Feb 2018, Accepted 10 Jul 2018, Published online: 20 Aug 2018

→ **Situation plausibility matters**

3. Content selection

Expectations for Informativity

- Maxim of Quantity (Grice 1975)
- Inclusion of disambiguating descriptors in reference (Dale & Reiter 1995)
- Omission of inferable information (Brown & Dell 1987)

→ **Newsworthiness matters**

4. Expecting the unexpected

Hypothesis: Estimates of what a speaker knows should differ from estimates of what the speaker will say.
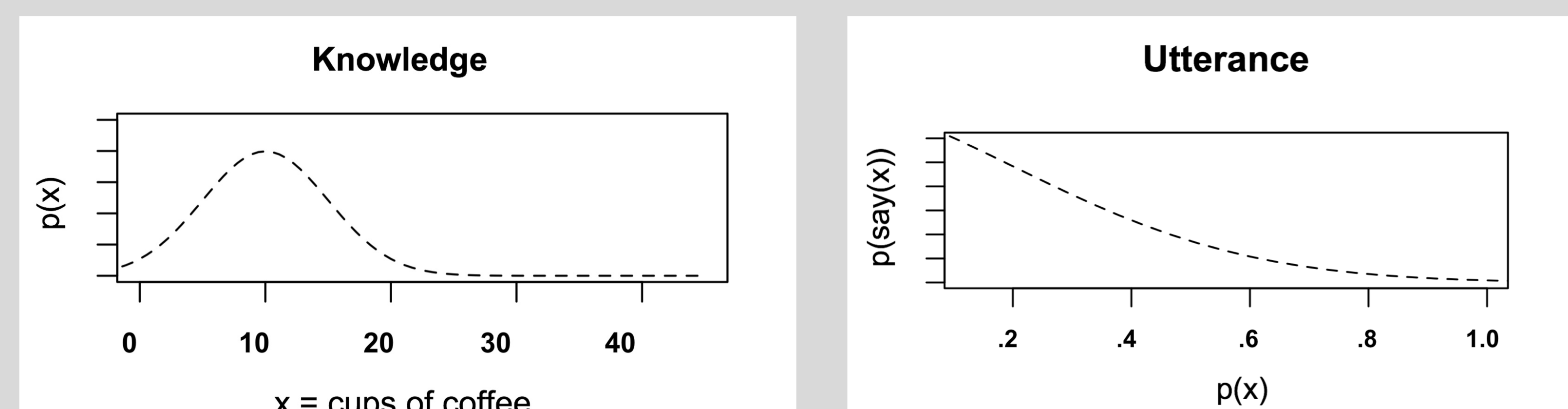
As listeners, we expect speakers to talk about situations that are:
- Plausible (*At CUNY, I saw a unicorn)
- Newsworthy (*At CUNY, I saw a poster)

→ **Values that approximate real-world knowledge for think; more extreme values for announce**

5. Think/Announce manipulation

Previous task provides *a priori* real-world estimates for a set of situations (Schöller & Franke 2017)

Andy is a man from the US.
How many cups of coffee do you think Andy drank last week?

→ **mean = 11.1**

Current materials: 12 scenarios adapted from Schöller & Franke; 2nd individual introduced as thinker/speaker

Andy is a man from the US. Andy has an aunt, Hannah.
Hannah [**thinks**/**announced to me**] that Andy drank __ cups of coffee last week.

6. Study 1, fill-in-the-blank task

Methods: Mturkers (N=31) type in a value for each item

**Raw means:** 32.7 (**think**) vs 41.2 (**announce**)

Analysis:
Linear mixed effects model with fixed effects for Schöller & Franke's real-world estimates, condition, and their interaction
→ only a main effect of real-world estimates ($p < 0.001$)
→ no main effect of condition ($p = 0.49$) and no interaction ($p = 0.34$)

a priori means, log-transformed
(from Schöller & Franke 2017)

Problems:
- Non-uniform response scales
- Newsworthy values can be large or small
- Outlier removal in a task eliciting newsworthy values
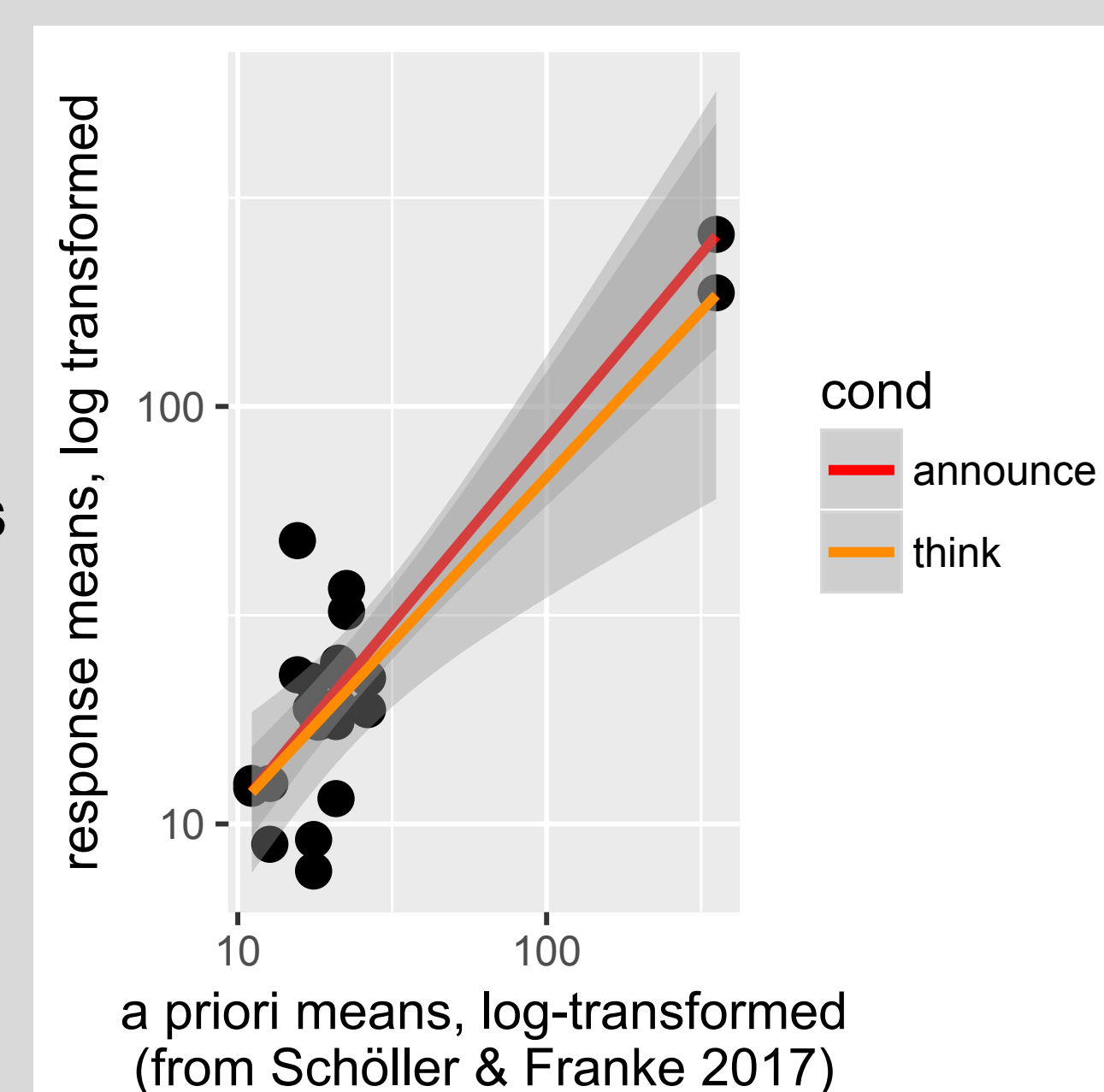
7. Study 2, forced-choice task

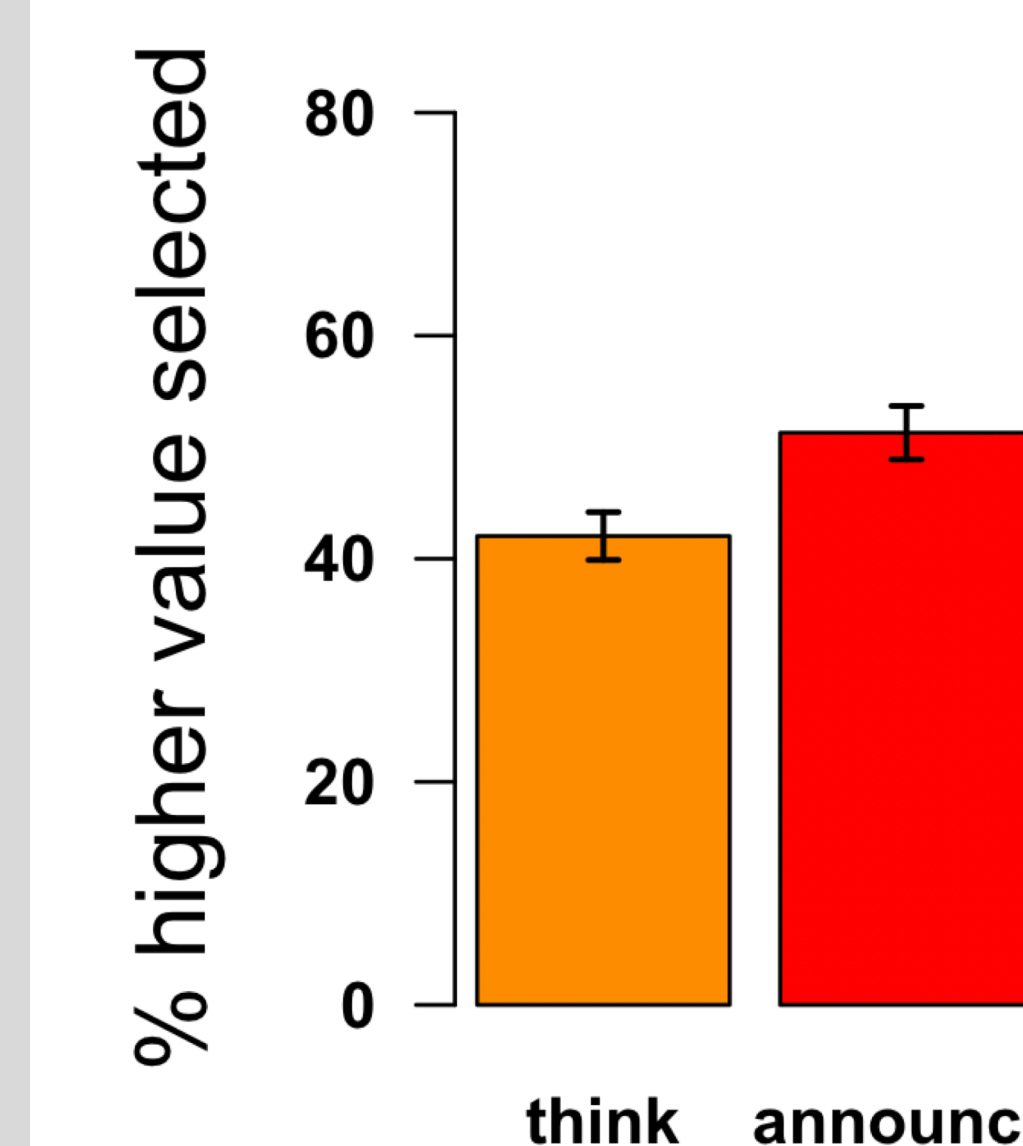**Methods:** Mturkers (N=90) select one of two choices for each item
- Lower value (Study 1 mean + 1/5 sd)
- Higher value (Study 1 mean + 4/5 sd)

Andy is a man from the US. Andy has an aunt, Hannah.
Hannah [**thinks**/**announced to me**] that Andy drank __ cups of coffee last week.
◎ 14   ◎ 20

Analysis:
Logistic mixed effects model for binary Lower/Higher response
→ **main effect of condition ($p < 0.01$)**

→ **Difference between expectations of what speakers know versus what they say**

→ **Situation probability influences message probability, but not directly**

8. Conclusions

Summary
Not all possible messages are worth uttering. The results here suggest that knowing that a speaker has chosen to utter a message can induce expectations for newsworthy content in that message. Rather than transparent mappings between situation probability and utterance probability, psycholinguistic models should distinguish real-world knowledge from content selection and surface realization.

Going forward
- What inferences arise from uninformative utterances? (see Kravtchenko & Demberg 2015)
- Are smaller-than-expected and larger-than-expected values equally newsworthy?
- In the **think** condition, why 42% higher value? Does encountering a sentence about someone's thoughts suggest newsworthy content about their incorrect thoughts?
- Online effects? Are appropriately newsworthy utterances <u>easy</u> to process?

References

Brown, P. & Dell, G. S. (1987). Adapting production to comprehension: The explicit mention of instruments. *Cognitive Psychology*, 19, 441–472.

Dale, R., & Reiter, E. (1995). Computational interpretations of the Gricean maxims in the generation of referring expressions. *Cognitive Science*, 19, 233–263.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics*. Vol. 3. New York: Academic Press, 41-58.

Hagoort P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science* 304, 438-41.

Kravtchenko, E. & Demberg, V. (2015). Semantically underinformative utterances trigger pragmatic inferences. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*.

Kutas, M., Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207, 203-5.

Schöller, A. & Franke, M. (2017). Semantic values as latent parameters: Testing a fixed threshold hypothesis for cardinal readings of few & many. *Linguistic Vanguard* 3.1.

Troyer, M. & Kutas, M. (2018). Harry Potter and the Chamber of What?: the impact of what individuals know on word processing during reading. *Language, Cognition and Neuroscience*.

Contact: hannah.rohde@ed.ac.uk