Notes from 1st brainstorming

Barbara and Michael

February 8 2018

1 Background: theories and experimental data

2 Models

Rather than redo, outwit or refute previous experimental contributions, **our goal is to chart new territory**. The starting point should be new predictions made by conceptually interesting probabilistic models, ideally extensions of the optimal-θ model or the RSA model for GASs (Lassiter and Goodman, online first; Qing and Franke, 2014a,b). Two ideas came to mind.

2.1 Lexical uncertainty about absolute GASs

Lexical uncertainty models (Bergen, Levy, and Goodman, 2012, to appear; Potts et al., 2016) assume that the listener is uncertain about the lexical meaning that a speaker might bring to the conversation. We consider uncertainty about the lexical meaning of absolute GASs: do they receive a relative (prior dependent) or absolute (pure standard) interpretation. We combine lexical uncertainty with the GA-model of Lassiter and Goodman (online first) (exactly what Tessler and Franke, 2018, did too):

$$L_1(x, \theta, \mathcal{L} \mid u) \propto S_1(u \mid x, \theta, \mathcal{L}) \cdot P(x) \cdot P(\theta) \cdot P(\mathcal{L}) \tag{1}$$

$$S_1(u \mid x, \theta, \mathcal{L}) \propto \exp\left(\alpha \cdot \ln L_0(x \mid u, \theta, \mathcal{L}) - \text{cost}(u)\right) \tag{2}$$

$$L_0(x \mid u, \theta, \mathcal{L}) \propto \mathcal{L}(u, x, \theta) \cdot P(x) \tag{3}$$

A lexicon is a map $\wr: u, x, \theta \mapsto \{0; 1\}$ which gives a (Boolean) truth-value for any utterance u of some GA, degree x and threshold θ. Only absolute gradable adjectives are lexically uncertain in the way described above. Model variants could distinguish cases where speakers maintain a single rule (all absolute GASs are prior-dependent/pure-standard) or between-item flexibility (e.g., *full* is prior dependent; *bent* is not).

This model is likely to make interesting **novel predictions about task effects** that other stories are unlikely to offer anything beyond hand-wavy explanations. Generally speaking, observations from previous trials/encounters could shift beliefs about the speaker's likely lexicon. If speaker's have been observed to use an absolute GA to refer to non-absolute degree x, listeners should update their lexical beliefs accordingly and be more likely to interpret a future use of this GA (or others, depending on the model variant) as relative-standard (prior-dependent). Also, interpretation tasks which display multiple utterances at the same time ((implicitly:) by the same speaker) could show interesting effects of jointly conditioning the model with all observed utterances (as observed by Tessler and Franke, 2018).

[mf: models need to be formulated precisely, implemented and predictions checked; this is all just intuitive guesses about potential model predictions]

2.2 Uncertainty about the prior (or the comparison class)

If speakers and listeners are uncertain about the prior over degrees $P(x)$, we are also bound to see potentially interesting **predictions about response dynamics as a function of prior exposure**. Suppose that items to be judged or chosen for interpretation are presented individually in each trial, or at the same time (like in stuff from the Chicago group [mf: insert ref]), this task manipulation will likely have effects on participants' construction of the comparison class / the relevant prior distribution. For example, for absolute GASs it might matter whether the end-point degrees have already been observed or not: as long as there is uncertainty about how likely these belong to the comparison class,

priors with little density on these degrees are reasonably likely, thus shifting predictions about θ "further away from the end-points". In general, the more extreme instances are observed, the more median instances should count as "neither this nor that".

3 Materials and envisaged pilot

4 Future music

Two **big issues** to ponder:

- how to derive predictions about response reaction times from prob-models?
- how to link model predictions to data from some suitable EEG study?

References

Bergen, Leon, Roger Levy, and Noah D. Goodman (2012). "That's what she (could have) said: How alternative utterances affect language use". In: *Proceedings of the 34th Annual Meeting of the Cognitive Science Society.*

— (to appear). "Pragmatic Reasoning through Semantic Inference". to appear in *Semantics & Pragmatics*.

Lassiter, Daniel and Noah D. Goodman (online first). "Adjectival vagueness in a Bayesian model of interpretation". In: *Synthese*.

Potts, Christopher et al. (2016). "Embedded implicatures as pragmatic inferences under compositional lexical uncertainty". In: *Journal of Semantics* 33.4, pp. 755–802.

Qing, Ciyang and Michael Franke (2014a). "Gradable Adjectives, Vagueness, and Optimal Language Use: A Speaker-Oriented Model". In: *Proceedings of SALT 44*. Ed. by Jessi Grieser et al. elanguage.net, pp. 23–41.

— (2014b). "Meaning and Use of Gradable Adjectives: Formal Modeling Meets Empirical Data". In: *Proceedings of CogSci*. Ed. by Paul Bello et al. Austin, TX: Cognitive Science Society, pp. 1204–1209.

Tessler, Michael Henry and Michael Franke (2018). "Not unreasonable: Carving vague dimensions with contraries and contradictions". Manuscript.