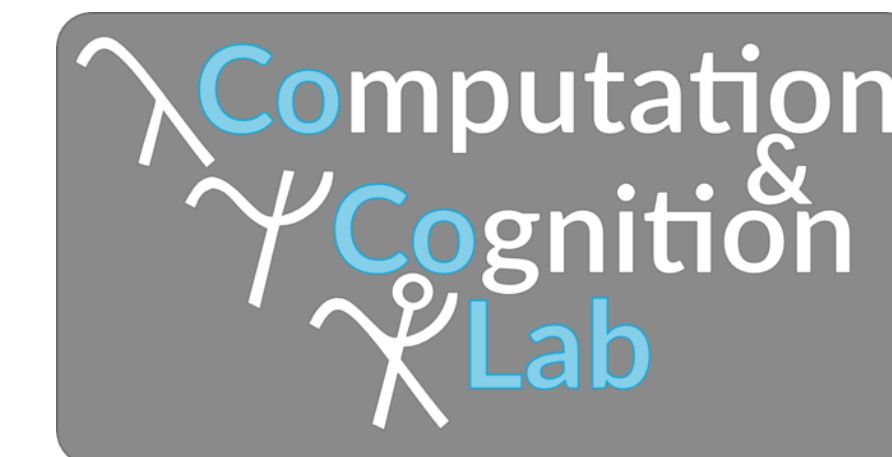# A cost- and information-based account of epistemic *must*

Judith Degen, Justine T. Kao, Gregory Scontras & Noah D. Goodman
Department of Psychology, Stanford University

Computation & Cognition Lab

Introduction

How do we communicate about beliefs, together with the evidence we use to form them? A case study: *must*

Must is strong, yet surprisingly weak

Compare: "It must be raining" *vs.* "It is raining"

Since Karttunen's (1972) **must is weak** mantra, four decades of semanticists cannot agree on the meaning of *must*; but *must* communicates more than strength

Evidentials in language: Why would you say so?

Consider: (standing in rain) "It must be raining"

Must q requires/signals **weak evidence** for *q* (von Fintel & Gillies; Lassiter; Matthewson, 2014)

Our account: M(arkedness)-implicature

The bare form conveys *q*, so listeners take marked *must q* to convey a marked meaning: the speaker arrived at *q* via an evidentially less certain route

We examine speaker & listener behavior as they relate to evidence strength, then implement a formal model of both within the Rational Speech-Act framework

Experiment 1 (n=40): Evidence strength

Rate probability of *q* (e.g., rain) given evidence (e.g., *You see that it's raining*), for 20 pieces of evidence (Fig. 1)

Strength estimates used in subsequent analyses

Experiment 2 (n=40): Utterance choice

How likely are **speakers** to use the marked *must p* utterance as evidence strength decreases?

Given a piece of evidence for *q*, then asked to **choose** one of four **utterances** to tell friend about the situation

– bare *q*, *must q*, *probably q*, *might q* –

Participants were more likely to choose the more marked *must q* form over the bare form as the strength of evidence decreased (Fig. 2; β=5.4, SE=2.4, p<.05)

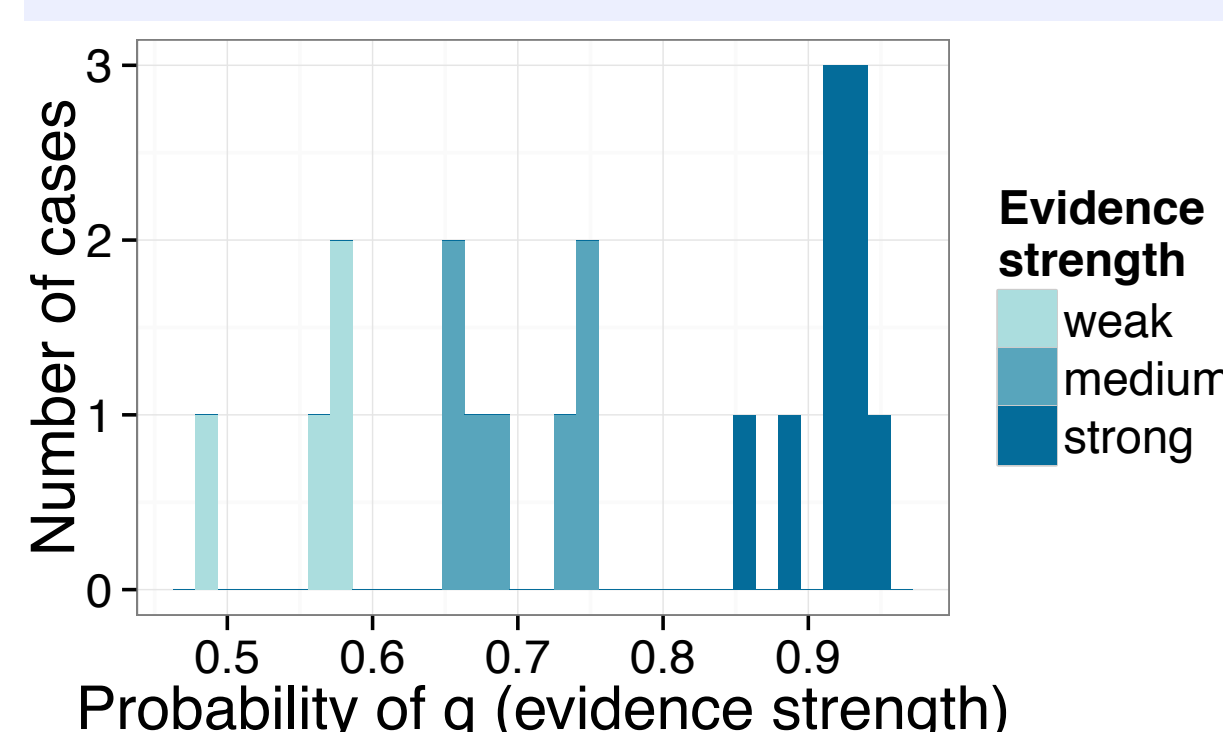## Empirical results/model predictions

Fig. 1: Experiment 1

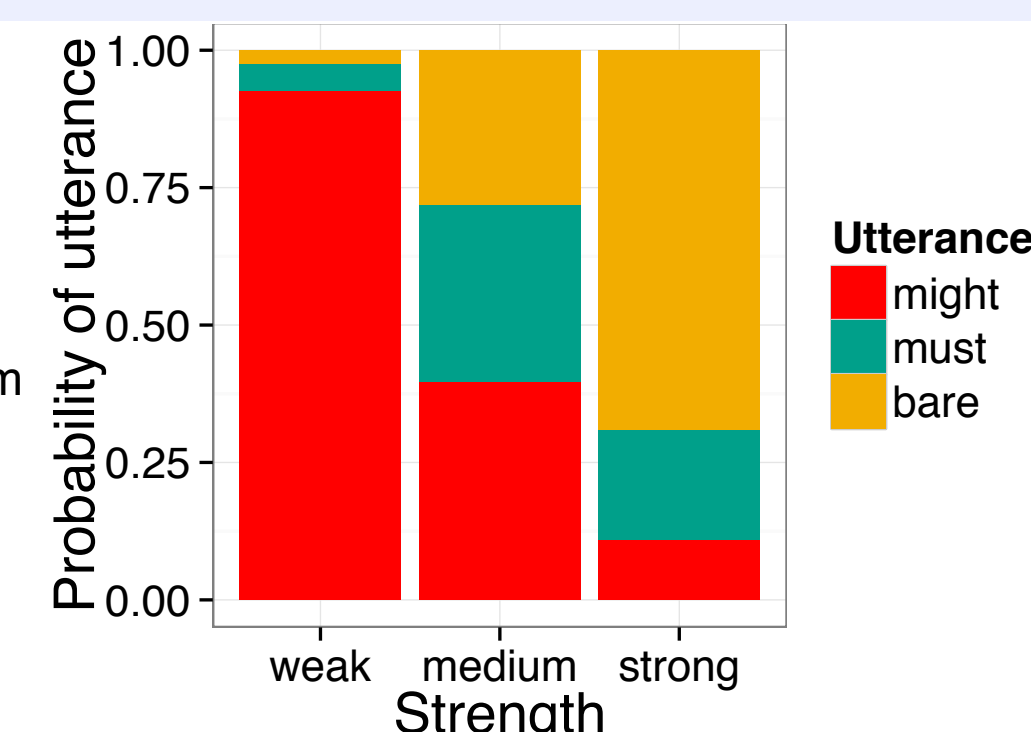### Fig. 2: Experiment 2

Fig. 3: Experiments 3a & 3b (inferred evidence strength)
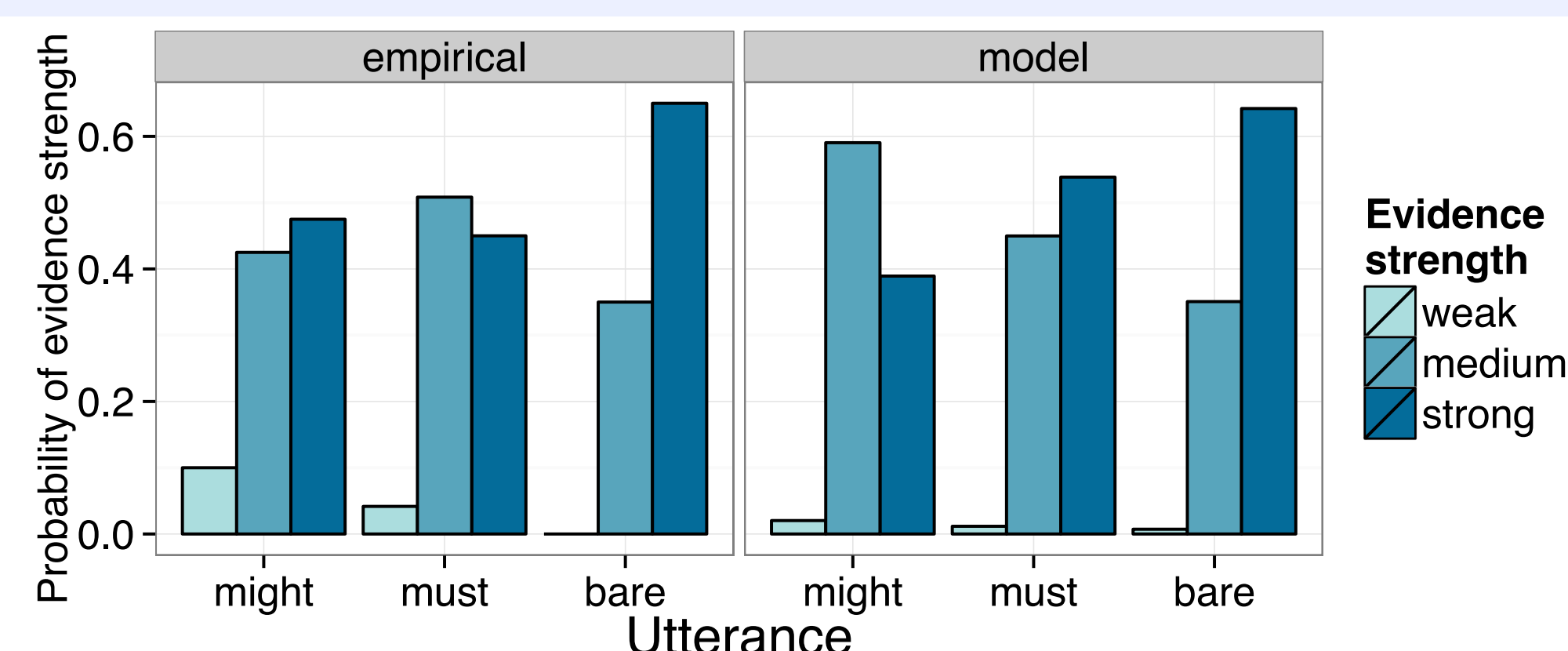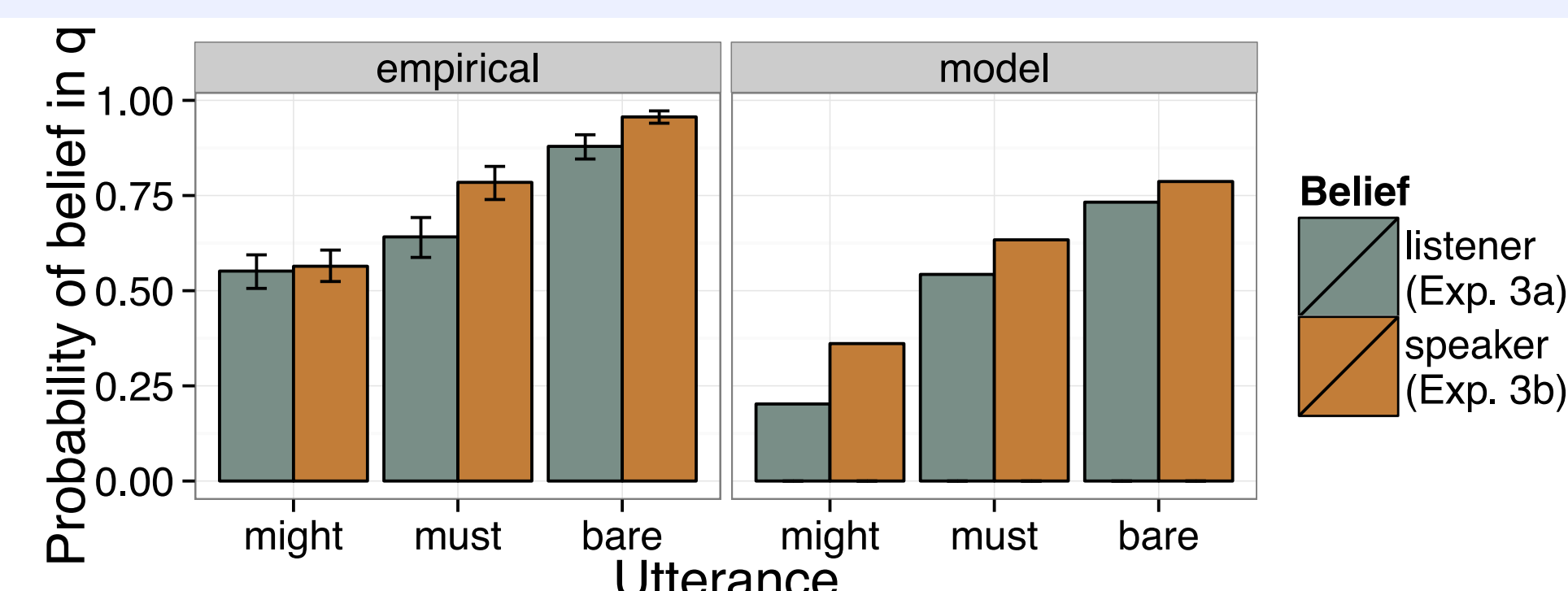

Fig. 4: Exps 3a & 3b (inferred degree of belief in *q*)

Exp. 3a (n=120): Utterance interpretation

How do **listeners'** beliefs about *q* and the strength of speakers' evidence for *q* depend on the utterance?

Given an utterance, then asked a) to **rate** the probability of *q*; and b) to select one out of five pieces of **evidence** that the speaker most likely had about *q*

Evidence strength lower after *must q* than after bare *q* (Fig. 3; β =-.08, SE=.01, t=-6.8, p<.0001)

Belief in *q* weaker after *must q* than after bare *q* (Fig. 4; β =-.21, SE=.02, t=-10.1, p<.0001)

Exp. 3b (n=60): Inferring speaker belief

How do listeners' beliefs about **speakers' beliefs** about *q* depend on the utterance?

Procedure as in Exp. 3a; asked to rate **speakers' belief** in *q* (e.g. *How likely is it that John thinks it's raining?*)

Listeners' estimates of speakers' beliefs stronger than their own (Fig. 4; β =.07, SE=.02, t=3.74, p<.001)

Model

We present a model extension within the **Bayesian Rational Speech-Act** framework (Frank and Goodman, 2012) using **lexical uncertainty** and **threshold semantics** (Lassiter and Goodman, 2013)

Model assumptions

- $p(q|bare) > \theta_b$, $p(q|must) > \theta_m$, $p(q|might) > 0$
- cost(*must q*) = cost(*might q*) > cost(*q*) (fit to data)
- Prior probability of *q = 0.5* (uniform)
- Probability of evidence *e* given *q* (fit to data)
- Probability of speaker belief *b* given evidence (Exp 1)

$$L_1(q,b,\theta,e \mid u) \propto S_1(u \mid q,b,\theta,e)\, P(b \mid e)\, P(e \mid q)\, P(q)$$

Listener infers that speaker has weaker belief given the costlier utterance *must q*. Given generative model of world state, evidence, and beliefs, listener then infers weaker evidence and lower probability of *q*, producing qualitatively similar results to those found in Exps 1 & 3

A new perspective on *must*

Weakened meaning derives straightforwardly from an M-implicature, which drives inference about evidence

References: von Fintel & Gillies (2010). *Must … stay … strong! Natural Language Semantics, 18.* Karttunen (1972). *Possible and must.* In *Syntax and Semantics.* Frank & Goodman (2012). Predicting pragmatic reasoning in language games. *Science, 336.* Lassiter (2014) The weakness of *must.* In *Proceedings of SALT 24.* Lassiter & Goodman (2013). Context, scale structure, and statistics in the interpretation of positive-form adjectives. In *Proceedings of SALT 23.* Matthewson (in press). Evidential restrictions on epistemic indefinites. In *Epistemic Indefinites.*