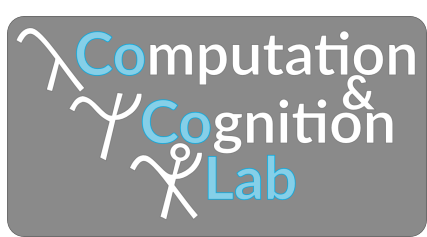




A cost and information-based account of epistemic *must*



Judith Degen <jdegen@stanford.edu> Justine Kao <justinek@stanford.edu>
Gregory Scontras <scontras@stanford.edu> Department of Psychology, Stanford University
Noah D. Goodman <ngoodman@stanford.edu>

Abstract

We explore the interpretation of statements with *must*, comparing these statements to utterances with weaker modals, or no modals at all. *Must*-statements coincide with less certain belief states; they are marked relative to bare statements, so they communicate a marked meaning, namely

Introduction: Strength, evidence, and modals

Arguing about semantics: How strong is *must*?

The epistemic necessity modal *must* is strong with respect to other modals, but remarkably weak when compared to bare statements without a modal: how could *must p* not entail *p*?

Compare: “It *must* be raining” vs. “It is raining”

? (?) gives birth to the “*must* is weak” mantra; ? (?) reject the mantra, arguing that *must* is a strong evidential; ? (?) shows how a strong semantics makes unreasonable claims about knowledge states of speakers, but picks up on the evidential meaning contributed by *must*

Evidentials in language: Why would you say so?

Must is often infelicitous when the speaker has direct evidence

Consider: (while standing in the rain) “It must be raining.”

Maybe *must* functions as an evidential, restricting evidence for *p* to not include maximally strong or trustworthy sources (?, ?, ?, ?)

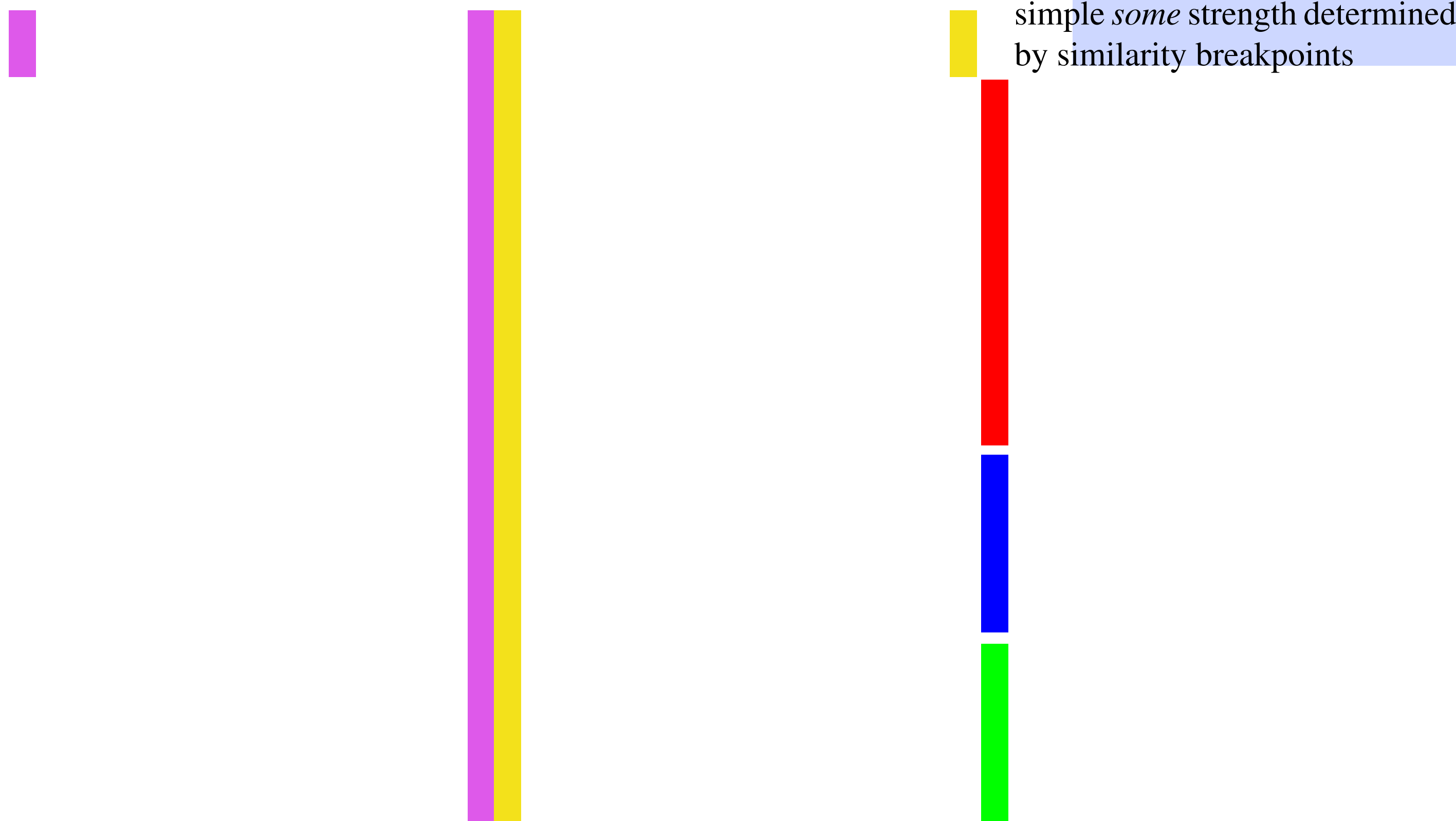
Choice between syntactic forms assumed to be driven by **meaning**

Study 2: Gradient Alternation Hypothesis

2 meanings: weak *sm* vs. strong *some* **Results** of MTurk ratings
(?, ?, ?, ?)

Methods. For each case, collected 10 similarity ratings of original to sentence with *some (of)* omitted to obtain a measure of *some*-NP strength (exploiting presuppositionality)

Results of fitting model to partitive and weak / strong simple *some*



meaning factors contribute to both weak/strong and simple/partitive difference – production pressures primarily to simple/partitive (strong dataset)

The test case: simple vs. partitive *some*

Alex ate	some	cashews.	[simple <i>some</i> ; shorter form]
Alex ate	some of the	cashews.	[partitive <i>some</i> ; longer form]
Alex ate	SOME	cashews.	[combined meaning contribution]

Hypotheses

1. **Meaning and production pressures operate in parallel in quasi-alternations** (when meanings of two forms are *similar enough*).
2. **Gradient Alternation Hypothesis:** Effects of production pressures are more pronounced, the more similar the meanings that the speaker intends to convey by using one of the two forms are.

Study 1: parallel pressures

Dataset 1237 cases of *some*-NPs (269 partitives, 23%) from Switchboard corpus after excluding cases that can only occur in one of the two forms (pronouns, singular count nouns, idioms)

Predictors entered in mixed-effects logit model predicting partitive use:

Meaning (discourse accessibility, ?, ?)	Production pressures	
	Availability	UID
- previous mention of NP	- frequency of head	- I(SOME NP head)
- topicality of <i>some</i> -NP	- animacy of head	- I(SOME previous word)
- modification of head	Alex ate some (of the) cashews	
- head type (mass/count)	previous word	NP head

$$I(\text{SOME} | \text{context}) = -\log(p(\text{some} | \text{context}) + p(\text{some of } DT | \text{context}))$$

Results

Meaning factors are strongest, but both **UID** factors and one **availability** factor affect partitive choice in predicted direction: more partitives with increasing information of SOME and decreasing availability of head.

Conclusion

- **Production pressures apply to the choice between forms that are not meaning-equivalent.** We conclude that, rather than production pressures applying after restriction of permissible forms by semantics, **the pressure to robustly communicate a core meaning applies in parallel with the pressure to find the most precise form to encode an intended meaning.**
- This is compatible with probabilistic approaches to form choice: the more similar the meaning (and associated inferences) of two forms, the more likely the choice between the two is to be affected by production pressures.