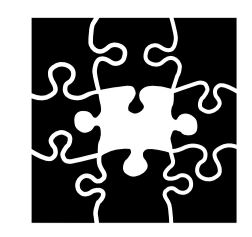


Gradable Adjectives, Vagueness and Optimal Language Use

A Speaker-Oriented Model

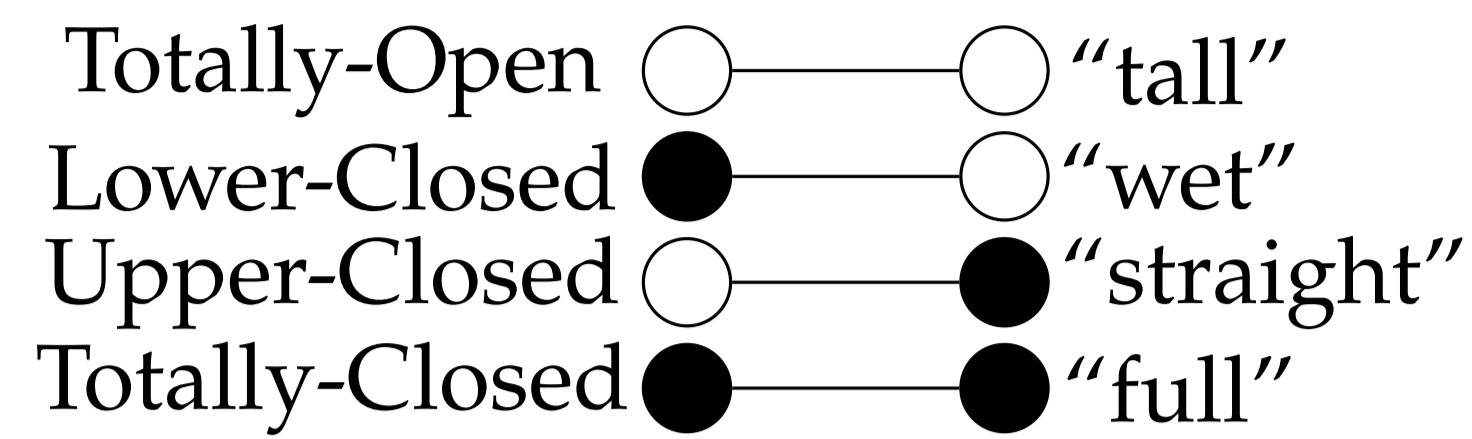
Ciyang Qing & Michael Franke, University of Amsterdam



INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Gradable Adjectives, Degrees and Scales

Gradable adjectives map individuals to degrees on degree scales.



E.g., $\llbracket \text{tall} \rrbracket = \lambda x. \text{height}(x)$

Open vs Closed Scales: whether endpoints are accessible.

(Kennedy2007:Vagueness-a)

Positive Forms

Composition with a silent morpheme pos

E.g. $\llbracket pos \text{ tall} \rrbracket = \lambda x. \text{height}(x) \geq \theta$,
 θ : standard of comparison (threshold)

1. How is θ contextually derived?

Relative/Absolute Adjectives

Relative (e.g., *tall*): vague standard
 Absolute (e.g., *full*): rigid standard

2. Why Relative vs Absolute?

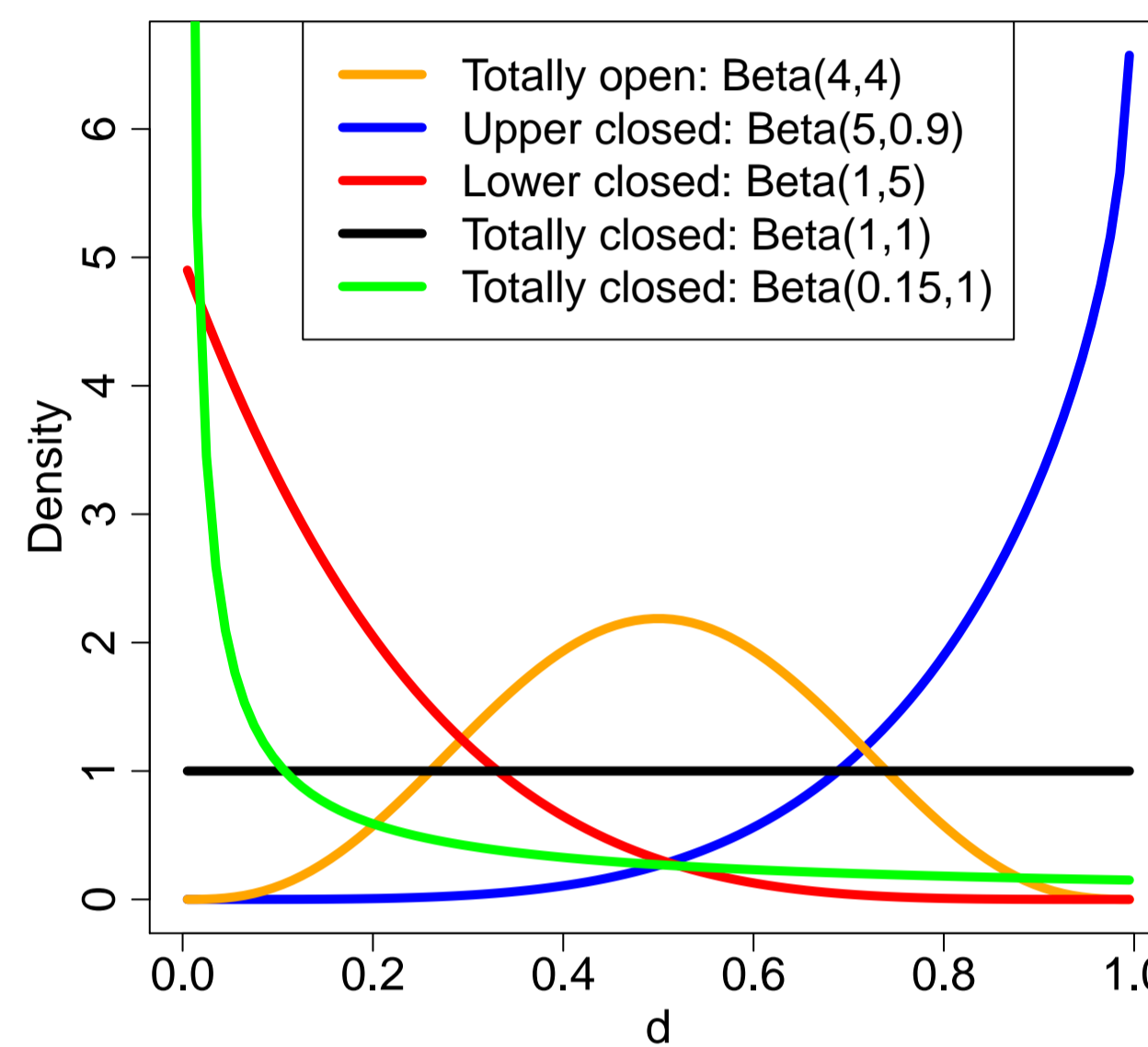
Interpretive Economy (ibid.)

Make maximal use of conventional meanings
 (Use endpoints as θ , if available)

Problem: Why is using endpoints optimal?
 “An optimization principle left unsupported by a theory of optimization” (Potts2008:Interpretive-E)

Communicative Efficiency

- Comparison classes as prior distribution $\phi(d)$ over degrees.
- Goal of communication: Using “ x is A ” truthfully to effectively convey the degree of each x in the comparison class.



Consider a threshold θ , for individual x with degree d :

- When $d < \theta$, positive form cannot be used
 Only prior for literal listener: $\rho_o(d | N; \theta) = \phi(d)$.
- Otherwise belief update according to the semantics of “ x is A ”
 $\rho_o(d | A; \theta) = \phi(d | d \geq \theta) = \frac{\phi(d)}{\int_{\theta}^{\infty} \phi(d) dd}$

(Sub-)Optimal Language Use

- Expected (average) communicative success
 $ES(\theta) = \int_{-\infty}^{\theta} \phi(d) \cdot \rho_o(d | N; \theta) dd + \int_{\theta}^{\infty} \phi(d) \cdot \rho_o(d | A; \theta) dd$
- (Sub-)optimal standard of comparison via soft-max
 $\Pr(\theta) \propto \exp(\lambda \cdot ES(\theta))$ (Luce1959:Individual-C)
- Speaker production via sampling a threshold from $\Pr(\theta)$
 $\sigma(A | d) = p(d \geq \theta) = \int_{-\infty}^d \Pr(\theta) d\theta$ (Lassiter2011:Vagueness-a)

Scales and Priors

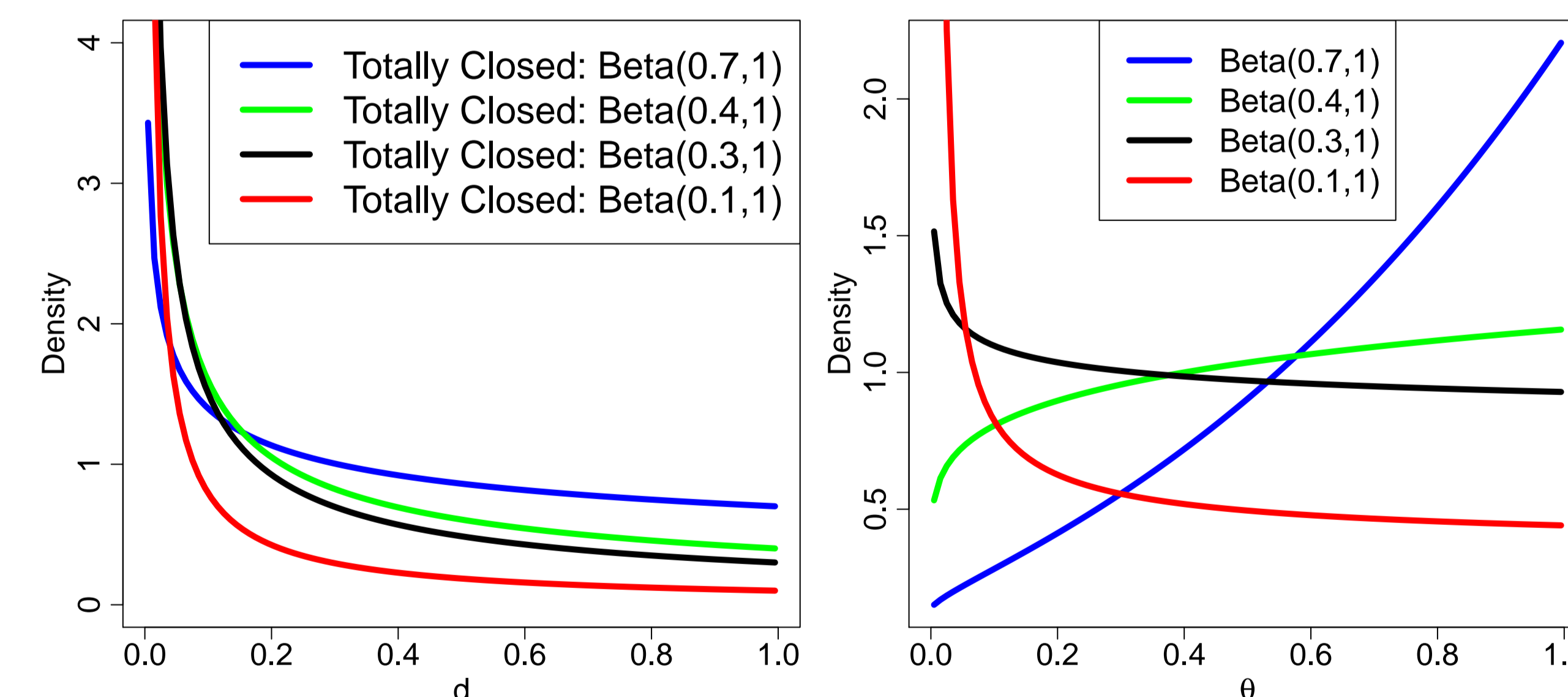
Prior distribution $\phi(d)$ is constrained by the scale

- Open/Closed: whether sufficient prior on endpoints
- In reality, uncertainty about $\phi(d)$:
 - The comparison class is often implicit
 - People seldom know the exact $\phi(d)$

Stability of Optimal Threshold (Relative vs Absolute)

Prototype prior $\phi(d)$, with some uncertainty:

- Open priors: slight change of $\phi(d) \Rightarrow$ optimal θ changes.
 Uncertainty about the optimal θ . (Relative)
- Closed priors: slight change of $\phi(d) \Rightarrow$ optimal θ remains.
 Rigid optimal θ . (Absolute)



Comparison to Rational Speech-Act Model

A different production rule in Lassiter & Goodman (2013):

$$\sigma(A | d, \theta) = \frac{\exp(\lambda U(A, d, \theta))}{\exp(\lambda U(A, d, \theta)) + \exp(\lambda U(N, d, \theta))}$$

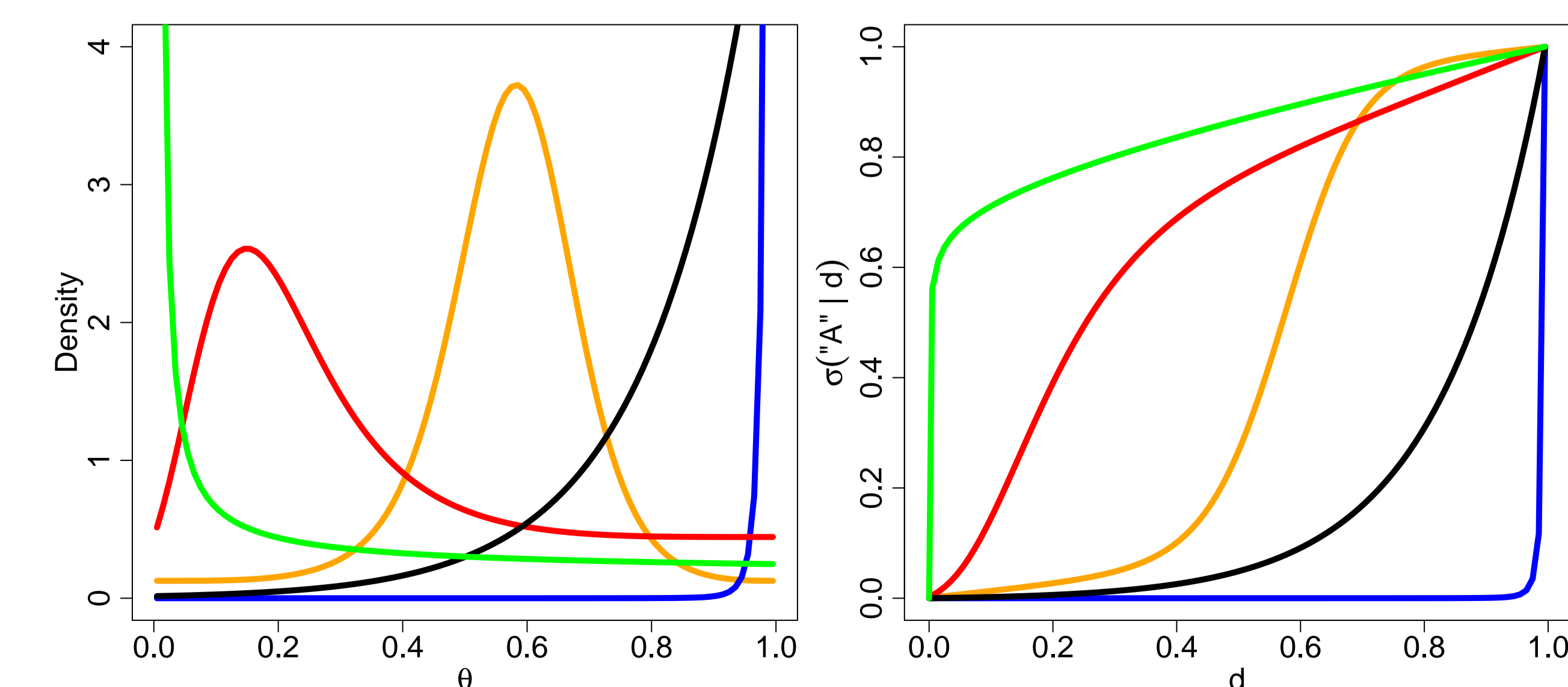
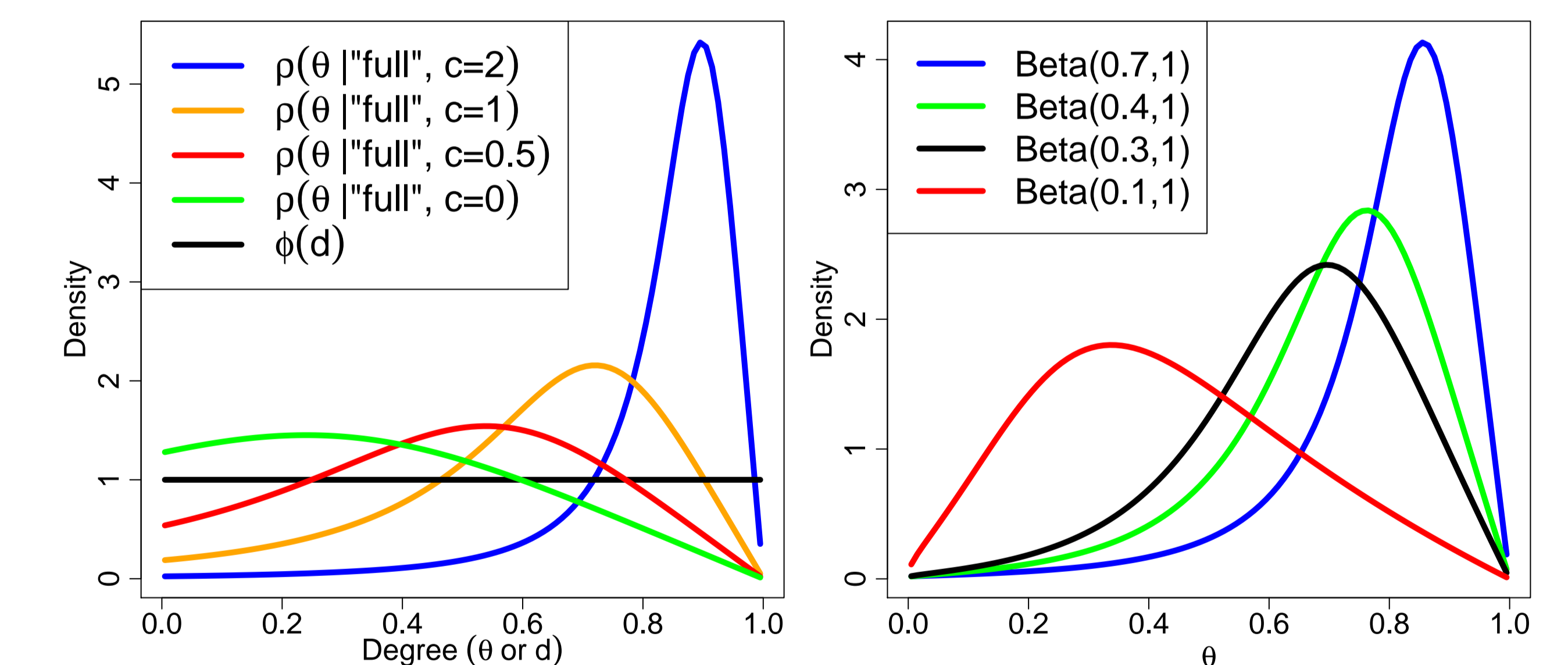
Utility of utterance u

$$U(u, d, \theta) = \log(\rho_o(d | u; \theta)) - C(u).$$

Pragmatic listener: joint inference about degree and threshold:

$$\rho(d, \theta | A) \propto \phi(d) \cdot \text{Unif}(\theta) \cdot \sigma(A | d, \theta).$$

- Listener assumes speaker knows θ , but is uncertain himself
- No predictive production model
- Prediction crucially relies on costs; no relative/absolute distinction when there is uncertainty about degree prior



Comparison to Previous Evolutionary Approaches

Potts (2008) considers *coordination* of θ and treats endpoints as most salient.

- Coordination of θ is not the direct purpose of communication.
- Endpoints need not be most salient to be optimal.

Franke (2012) considers *referential use* of gradable adjectives.

- “The tall man” (referential) vs “The man is tall” (descriptive)

Conclusions

- The “vagueness pattern” of gradable adjectives can be explained via (sub-)optimal descriptive language use.
- Relative vs Absolute: stability of optimal threshold under uncertainty about the degree prior $\phi(d)$

References: Franke, M. (2012) in: *Amsterdam Colloquium 2011* * Kennedy, C. (2007) *Linguistics and Philosophy* 30 * Lassiter, D. (2011) in: *Vagueness in Communication* * Lassiter, D. & Goodman (2013), N.D. in: *Proceedings of SALT 23* * Luce, D. (1959) *Individual Choice Behavior: A Theoretical Analysis*. * Potts, C. (2008) *Manuscript*.