Gradable Adjectives, Vagueness and Optimal Language Use
A Speaker-Oriented Model

Ciyang Qing & Michael Franke, University of Amsterdam

Gradable Adjectives, Degrees and Scales

Gradable adjectives map individuals to degrees on degree scales.
E.g., [tall] = λx. height(x)
Open vs Closed Scales: whether endpoints are accessible.

(Kennedy2007:Vagueness-a)

Positive Forms
Composition with a silent morpheme pos
E.g. (pos tall] = 1x. height(x) ≥ θ,
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
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 φ(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 < θ, positive form cannot be used
  Only prior for literal listener: ρ0(d | N; θ) = φ(d).
• Otherwise belief update according to the semantics of “x is A”
  ρ0(d | A; θ) = φ(d | d ≥ θ) = \[\int_0^\theta \frac{1}{f_0} \phi(d) \, dd\].
  (Sub-)Optimal Language Use
• Expected (average) communicative success
  ES(θ) = ∫^\theta_0 ϕ(d) · ρ0(d | N; θ) 
  + ∫^\infty_\theta ϕ(d) · ρ0(d | A; θ) 
• (Sub-)optimal standard of comparison via soft-max
  Pr(θ) ∝ exp(A · ES(θ))
  (Luce1959:Individual-C)
• Speaker production via sampling a threshold from Pr(θ)
  σ(A | d) = p(d ≥ θ) = ∫_0^\theta Pr(θ) 
  (Lassiter2011:Vagueness-a)

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.
• “The tall man” (referential) vs “The man is tall” (descriptive)

Comparison to Rational Speech-Act Model
A different production rule in Lassiter & Goodman (2013):

\[\sigma(A | d, θ) = \exp(\lambda U(A, d, θ)) \cdot \exp(\lambda U(N, d, θ))\]

Utility of utterance u
\[U(u, d, θ) = \log(ρ0(d | u; θ)) - C(u)\]
Pragmatic listener: joint inference about degree and threshold:
\[ρ(d, θ | A) \propto ϕ(d) \cdot \text{Unif}(θ) \cdot σ(A | d, θ)\]
• 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

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 φ(d)

References: