How to find the rabbit in the big(ger) box: Reasoning about contextual parameters for relative adjectives under embedding

Helena Aparicio (MIT) & Roger Levy (MIT) & Elizabeth Coppock (Boston University) · XPrag 2019

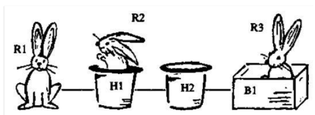
Background

1) The interpretation of relative adjectives (e.g., *tall*, *big*, *small*) is dependent on a contextually salient COMPARISON CLASS (CC) that determines the value of adjectival threshold (e.g. Kennedy 1999).

the tall building
CC: Buildings in the context

the tall glass
CC: Glasses in the context

2) When *the hat* is embedded in a 'Haddock description' like *the rabbit in the hat*, uniqueness w.r.t. hats is not required (Haddock, 1987); here, only uniqueness w.r.t. rabbit-containing hats.



the hat = ?? / *the rabbit in the hat* = R1

Are comparison classes for embedded adjectives restricted in the way that the descriptive content of the embedded definite appears to be?

Research Question

In *rabbit in the big bag*, what counts as a *big bag*?

Hypothesis 1: a bag that is big *for a bag with a rabbit in it*
Hypothesis 2: a bag that is big *for a bag*

the rabbit in the big bag

Hypothesis 1
CC = rabbit-containing bags
One possible threshold

Hypothesis 2
CC = bags
Two possible thresholds

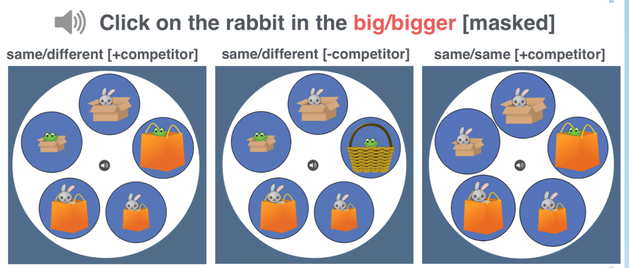
In general, when definite descriptions are interpreted relative to a restricted domain (e.g. *rabbit-containing* bags), must the comparison class for a gradable adjective contained within it shrink accordingly?

Predictions

- **Hyp. 1**: Embedded *big* is interpreted wrt **restricted** CCs. Then *big* should be **insensitive** to bags without rabbits.
- **Hyp. 2**: Embedded *big* is interpreted wrt **unrestricted** CCs. Then *big* should be **sensitive** to bags without rabbits.

Design/Procedure and Results

Design/Procedure: Referent selection with instructions partly masked by static noise. Participants clicked to hear audio.



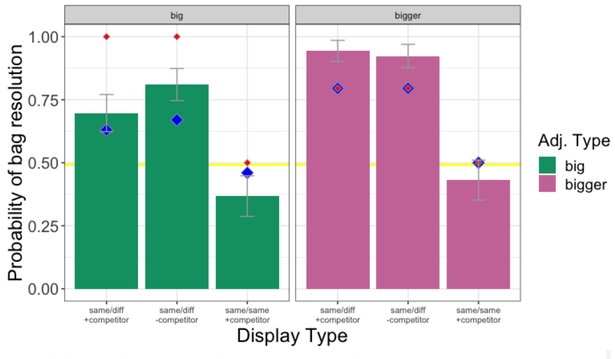
Manipulated variables:

- Display Type
  - Same or different animals in two smallest bags/boxes? (same[bags]/different[boxes] ⇒ a modifier on *box* would be uninformative as *the rabbit in the box* would suffice)
  - Is there a third, biggest bag (a competitor)?
- Adj. Type: *big* vs. *bigger*
- Bags vs. boxes (or equivalent) counterbalanced
- Random order within & between displays

Outcome variable: Selected image (< 1% chose non-targets)
– Target 1: rabbit in medium bag (lower left in displays above)
– Target 2: rabbit in medium box (top in displays above)

12 items, 24 fillers, 75 participants (Prolific and AMT).

Results. (Higher bar ⇒ more clicks to Target 1)



Model predictions: **Hyp. 1:** ◆ / **Hyp. 2:** ◆ (see next panel)

Main findings:

- Informativity is the biggest factor; if *the rabbit in the box* would have sufficed, the bag is preferred ($p < 0.05$).
- Moreover, appropriateness of *big* (not of *bigger*) is negatively affected by the presence of a competitor object ($p < 0.05$).
  ⇒ **Hypothesis 2** supported.

RSA Model

Literal listener: Assigns prior-based probability to every referent r in context C that satisfies the literal semantics of description d, assuming threshold θ for *big* (see semantics below).

$$L_0(r \mid d, C, \theta) \propto [\![d]\!]^{C,\theta}(r) \cdot P(r)$$

Problem: This distribution is undefinable if there is no referent in C satisfying d! Solution: special fail referent, with prior probability ϵ.

$P(r) = \epsilon$ if $r = \mathsf{fail}$; else uniform among $r \in C$

Speaker: Assigns probabilities to descriptions d given intended referent r, context C and threshold θ that reflect both the probability of the literal listener picking the referent and the cost (length) of d.

$$S_1(d \mid r, C, \theta) \propto L_0(r \mid d, C, \theta) - \mathsf{cost}(d)$$

Pragmatic listener: Assigns probabilities to referents r given (partially masked) description d, reflecting the probability of the speaker using (any resolution of) d to describe r in any context C with any threshold θ.

Marginalizing over C, θ, and N_2:

$L_1(r \mid d = N_1$ in the (Adj) [masked]$) \propto$
$\sum_C \sum_\theta \sum_{N_2}$
$S_1(d = N_1$ in the (Adj) $N_2 \mid r, C, \theta) \cdot P(r \mid C) \cdot P(\theta \mid C, d) \cdot P(C)$

Global context $\mathcal{C} = \{r_1, r_2, r_3, r_4, r_5\}$; $P(C)$ is uniform among $\wp(\mathcal{C})$.

Hypothesis 1: $P(\theta \mid C, d = N_1$ in the (Adj) $N_2)$

$$= \begin{cases} \text{uniform} & \text{among } \{\theta : \exists x \in C : [\![N_2]\!](x) \text{ and } \mathsf{size}(x) = \theta \text{ and } x \text{ contains an } N_1\} \\ 0 & \text{otherwise} \end{cases}$$

Hypothesis 2: $P(\theta \mid C, d = N_1$ in the (Adj) $N_2)$

$$= \begin{cases} \text{uniform} & \text{among } \{\theta : \exists x \in C : [\![N_2]\!](x) \text{ and } \mathsf{size}(x) = \theta\} \\ 0 & \text{otherwise} \end{cases}$$

Semantics for *the* and *big*

Bumford's (2017) semantic analysis of Haddock descriptions lets the uniqueness component of the inner definite take scope above *rabbit*, effectively disappearing in our setting (with max 1 bag per rabbit).

- $[\![$the rabbit in the big $N_2]\!]^{C,\theta}(r) = 1$ if and only if: r is the unique r' such that $[\![$rabbit in a big $N_2]\!]^{\theta,C}(r') = 1$
- $[\![$rabbit in a big $N_2]\!]^{C,\theta}(r) = 1$ if and only if: r is a rabbit in C that is in a b such that:
  - $[\![N_2]\!](b) = 1$
  - $\mathsf{size}(b) > \theta$

(Other approaches to Haddock descriptions are viable, but require higher-order RSA models to achieve qualitative accuracy.)

Semantics for *bigger*

$[\![$the rabbit in the bigger $N_2]\!]^{\theta,C}(r) = 1$ if and only if:

r is the unique rabbit in C that is in a b such that:

- $[\![N_2]\!](b) = 1$
- there is one b' in C such that $[\![N_2]\!](b') = 1$ and $\mathsf{size}(b) > \mathsf{size}(b')$
- and if *bigger* takes high scope with the definite: b' contains a rabbit *(0.5 probability of high scope)*

Conclusion

Hypothesis 2 is supported: In descriptions where uniqueness is computed with respect to rabbit-containing bags, what counts as *big* is computed with respect to all bags, rabbit-containing or not.

On Bumford's semantic approach to Haddock descriptions, such a finding is expected, as scope of *the* should not affect thresholds.

For a pragmatic view of Haddock descriptions, where they are based on a restriction of the context for the purposes of the inner definite (e.g. Muhlstein et al. 2015, see also e.g. Chambers et al. 2002), the consequence of our findings is that this context restriction does not affect all context-sensitive expressions equally.

Some open questions

- We have shown model results for a particular semantic approach to Haddock descriptions. What about pragmatic approaches (flexible contexts, context narrowing)?
  - Preliminary finding: Pragmatic approaches can yield qualitatively accurate results, using an L_2 who reasons about an S_2 who reasons about L_1.
- We have assumed a flat prior on contexts. Are there prior preferences for contexts of greater size, or greater homogeneity?

Selected References

Brown-Schmidt, Sarah & Michael K. Tanenhaus. 2008. Real-time investigations of referential domains in unscripted conversation: a targeted language game approach. *Cognitive Science* 32(4). 643–684.

Bumford, Dylan. 2017. Split-scope definites: Relative superlatives and Haddock descriptions. *Linguistics and Philosophy* 40(6). 549–593. https://doi.org/10.1007/s10988-017-9210-2.

Chambers, Craig G., Michael K. Tanenhaus, Kathleen M. Eberhard, Hana Filip & Greg Carlson. 2002. Circumscribing referential domains during real-time language comprehension. *Journal of Memory and Language* 47. 30–49.

Haddock, Nicholas J. 1987. Incremental interpretation and Combinatory Categorial Grammar. In *Proceedings of the 10 International Joint Conference on Artificial Intelligence*, vol. 2, 661–663. Morgan Kaufmann.

Kennedy, Christopher. 1999. *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison.* New York: Garland.

Muhlstein, Larry, Christopher Potts, Michael C. Frank & Roger Levy. 2015. Pragmatic coordination on context via definite reference. Poster presented at XPRAG 2015.