A theoretically motivated quantitative model
for the interaction between vagueness and implicatures

Alexandre Cremers – Vilnius University

Leffel et al. [4] observed a puzzling contrast between the implicatures of relative and minimum standard adjectives, which they attribute to the fact that the former, but not the latter, are vague:

(1) John is not very tall.
$\not\rightsquigarrow$ John is tall

(2) The antenna is not very bent
$\rightsquigarrow$ The antenna is (somewhat) bent

(2) gives rise to the expected structural implicature, by competition with the simpler and more informative alternative *not bent*, but this implicature is absent in (1), unless *very* is stressed. [4] remark that no height can both clearly satisfy *tall* and clearly falsify *very tall*, making the candidate strengthened meaning of (1) akin to *borderline contradictions* such as “tall and not tall” ([6]). By contrast, since *bent* can be interpreted strictly, one can choose a degree arbitrarily close to 0 in order to fully satisfy both *bent* and *not very bent*. [4] propose to generalize [2]'s notion of *innocent exclusion* so that the EXH operator block such borderline contradictions. Doing so captures the initial observation, but we argue that implicatures' sensitivity to vagueness is unlikely to be encoded semantically. Instead, we propose a pragmatic model which makes explicit the intuition of [4], but derives the contrast using the standard definition of EXH: (1) does not give rise to an implicature because *tall but not very tall* is only compatible with a very narrow range of heights, and if the speaker and listener assign slightly different thresholds to *tall*, the heights they consider “tall but not very tall” may not overlap and communication may fail.

Model description: We factor the speaker's uncertainty about the interpretation by implementing higher-order vagueness in the model: not only is there uncertainty about θ, but the distribution of θ is itself uncertain. We adopt [7]'s implementation of supervaluationism in RSA: the utility of a message is its average utility across all possible threshold distributions. Since utility diverges to $-\infty$ as the probability of the message being true approaches 0, a message must be true under all possible interpretations to be usable. In line with the grammatical view of implicatures ([1]) and recent work in the RSA framework ([3]), implicature derivation is treated as a disambiguation problem between parses with and without EXH. We adapt [3]'s Global Intentions model, which differs from the supervaluationist treatment of underspecification: the speaker chooses the pair (message, parse) which best conveys their intention. In particular, this decision rule does not prevent the speaker from using a message u when one of its interpretation is false or likely false (e.g., *not very tall*). Piecing everything together, the model captures the observation in (1) as follows: upon hearing *not very tall*, the pragmatic listener knows that—in principle—the speaker could mean either the exhaustive or literal interpretation. However, no matter which height the speaker had in mind, the exhaustive interpretation has a very low expected utility (across all possible vague denotations for *tall* and *very tall*): in supervaluationist terms, no height makes EXH*[not very tall]* supertrue. By contrast, the literal interpretation is compatible with low heights under any reasonable threshold for *very tall*. The listener therefore concludes that the speaker likely meant the literal interpretation, and that John is somewhat short. Concretely, we assign the following truth-conditions to vague messages, where θ and θ+δ are the thresholds for POS *adj* and *very adj* respectively, h the degree to convey, and Θ a set of parameters describing the distribution of θ and δ:

$[\![\text{POS adj}]\!]^{h,\Theta} = P(\theta < h|\Theta)$; $[\![\text{not POS adj}]\!]^{h,\Theta} = P(\theta \geq h|\Theta)$; $[\![\text{very adj}]\!]^{h,\Theta} = P(\theta + \delta < h|\Theta)$
$[\![\text{not very adj}]\!]^{h,\Theta} = P(\theta + \delta \geq h|\Theta)$; $[\![\text{EXH not very adj}]\!]^{h,\Theta} = P(\theta < h \leq \theta + \delta|\Theta)$

We follow [5] in assuming an additional ambiguity between POS and MIN for *late*, but for reasons of space we skip details regarding this aspect of the model (it doesn't play a crucial role in predictions). Our L_0 listener is parametrized by Θ and a parse i. The speaker S_1 selects the pair (u, i) such that u under parse i maximizes expected utility (across all parameter sets Θ). L_1 jointly infers h and i using Bayes' rule, with uniform prior on all parses compatible with u.

$L_0(h|u, i, \Theta) \propto P(h)[\![u]\!]^{h,i,\Theta}$ $\quad U_1(u, i|h) = \int \log L_0(h|u, i, \Theta)P(\Theta)d\Theta - c(u)$
$S1(u, i|h) \propto \exp(\lambda U_1(u, i|h))$ $\quad L_1(h, i|u) \propto P(h)S_1(u, i|h)$

Implementation and Evaluation: We tested our model on [4]'s Exp 1, which compared relative *tall* and minimum standard *late*. Because we are not interested in explaining vagueness *per se*, only its interaction with implicatures, we fitted a hierarchical Stan model on data from the affirmative constructions *adj* and *very adj* to obtain the distribution of Θ_{tall} and Θ_{late} empirically. As a first approximation, we treat within-participant fuzziness as indicative of first-order vagueness, and between-participants variance as second-order vagueness: we assume that each participant instantiates a single Θ, and the population variance reflects the distribution of Θ. From the fitted hyperparameters of the distribution of Θ, we computed L_1's posterior on EXH as a function of $(\lambda, c_{adj}, c_{not}, c_{very})$, and fitted participants' responses to *not adj* and *not very adj*, assuming that the acceptability of a message u in this experiment is its expected truth given Θ and a pragmatically inferred probability $P(\text{EXH})$. The Θ fitted for each participant from their responses to *adj* and *very adj* was fed to a new hierarchical model with parameters $(\lambda, c_{adj}, c_{not}, c_{very})$, predicting behavior on *not very adj*. Fig. 1 shows that the model correctly predicts participants' behavior with median by-participant parameters (λ=1.5,c_{adj}=2.0,c_{not}=2.6,c_{very}=2.1). The posterior probability of the exhaustive interpretation is lower with *tall* than with *late* (CIs [.14,.19] vs. [.36,.39]). Crucially, Fig. 2 shows that $P(\text{EXH}|\textit{not very late})$ usually increases with rationality, while $P(\text{EXH}|\textit{not very tall})$ always falls to 0.

Discussion: By combining results and intuitions from the theoretical literature with recent advances in RSA models, we were able to capture the whole range of behaviors in the experimental data. Qualitatively, the model correctly predicts that *not very tall* does not convey "tall but not very tall", while this interpretation can be very salient for *not very late*. We can show that the decision to use supervaluationism for vagueness and Global Intentions for implicatures is crucial: treating vagueness and implicatures uniformly under a single disambiguation rule fails to capture the contrast between *tall* and *late*.

Acknowledgment: This research was funded by the European Social Fund under the measure 09.3.3-LMT-K-712 "Development of Competences of Scientists, other Researchers and Students through Practical Research Activities".

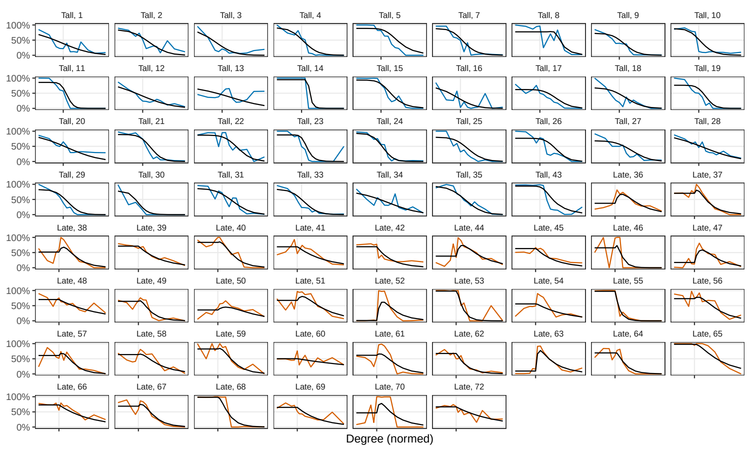

Fig. 1: Individual participants' acceptability of *not very adj* (colored line) and model fit (black line). The implicature translates as reduced acceptability for low degrees.

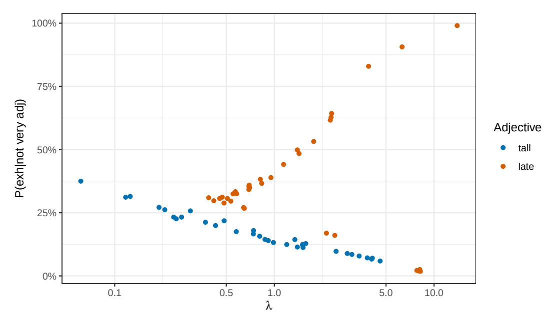

Fig. 2: $P(\text{EXH}|\textit{not very adj})$ as a function of participants' fitted rationality (log-scale).

References

[1] G. Chierchia, D. Fox, and B. Spector. Scalar implicature as a grammatical phenomenon. In *Semantics: An International Handbook of Natural Language Meaning*, volume 3, pages 2297–2331. Mouton de Gruyter, Berlin, 2012.

[2] D. Fox. Free choice disjunction and the theory of scalar implicature. In U. Sauerland and P. Stateva, editors, *Presupposition and implicature in compositional semantics*, pages 71–120. Palgrave Macmillan, New York, NY, 2007.

[3] M. Franke and L. Bergen. Theory-driven statistical modeling for semantics and pragmatics: A case study on grammatically generated implicature readings. *Language*, 96(2):77–96, 2020.

[4] T. Leffel, A. Cremers, N. Gotzner, and J. Romoli. Vagueness in Implicature: The Case of Modified Adjectives. *Journal of Semantics*, 36(2):317–348, 2019.

[5] C. Qing. Zero or minimum degree? Rethinking minimum gradable adjectives. *Proceedings of Sinn und Bedeutung 25*, 2021.

[6] D. Ripley. Contradictions at the borders. In R. Nouwen, R. van Rooij, U. Sauerland, and H.-C. Schmitz, editors, *Vagueness in communication*, pages 169–188. Springer, 2011.

[7] B. Spector. The pragmatics of plural predication: Homogeneity and non-maximality within the rational speech act model. In A. Cremers, T. van Gessel, and F. Roelofsen, editors, *Proceedings of the 21st Amsterdam Colloquium*, page 435, 2017.