

Vagueness is Rational Under Uncertainty

Joey Frazee and David Beaver

University of Texas at Austin
 Department of Linguistics
 1 University Station B5100
 Austin, TX 78712-0198 USA
 {jfrazee,dib}@mail.utexas.edu

1 Introduction

In this paper we show that some of the properties of vague scalar adjectives are consequences of rational communication. Traditionally, theories of vagueness are evaluated through the lens of a single cluster of philosophical desiderata: how vagueness fits into a theory of truth (or even metaphysics), the sorites, and related issues. These are important, and we will have more to say about some of them, but we also seek to refocus the analysis of vagueness, moving away from consideration of abstract philosophical problems, and towards consideration of the problems faced by ordinary language users. It is our contention that it is by considering the latter that the former can best be understood. Using standard mathematical tools, we suggest that, given a certain model of communication, what have typically been taken to be puzzles in the truth conditions of vague expressions, should just be expected consequences of communication under uncertainty.

The most basic criterion for any successful semantic theory is that it should establish when sentences are true, and when they are false. This is a tall order as regards sentences involving vague predicates. As a result, the bar is sometimes set too low in the sense that some scholars settle for giving definite truth values for only a subset of conditions in the world, or for abandoning truth in favor of some notion of acceptability. We have no *a priori* argument against theories which attribute partial or fuzzy truth conditions to vague sentences,¹ and no *a priori* arguments against theories that are stated primarily in terms of acceptability, but we like and want to retain classical semantics. *Ceteris paribus*, a theory which attributes a definite, boolean truth value to a vague sentence is preferable to one which does not, and a theory which allows us to talk about the objective truth value of a vague sentence is preferable to one which relativizes the truth conditions to features of the utterance situation or speaker. We will not claim that it is always possible to specify objective boolean truth values for every vague sentence even with complete knowledge of the utterance situation and

¹ But we are claiming that vagueness is not a result of degrees of truth or likelihood. Vagueness rides atop distributions of certain facts. I.e. contra fuzzy logic, an individual is either *tall* or *not tall*, rather than *tall* to a degree x and *not tall* to a degree $1 - x$.

speaker, but we will offer a theory which implies that every vague sentence has a boolean truth value, and we will offer a theory of how speakers can and must approximate it.

2 A Free Variable Theory of Scalar Adjectives

Following Cresswell [1], we analyze the semantics of “Xena is tall” simply as saying that Xena’s height is greater than some standard of comparison or threshold degree d , where d is a free variable. Thus our analysis is semantically parallel to free variable accounts of e.g. modals [2] or quantifiers [3]. However, there is a crucial difference. Consider the analysis of quantifiers: if “Everybody is happy” is analyzed as $\forall x \in C, \text{happy}(x)$, with the restrictor C being a free variable, then it is normally assumed that while the hearer may sometimes be uncertain as to the value of C , the speaker has a particular value in mind, and cannot in any sense be wrong about that value. However, in our free variable account of vagueness, there is no such asymmetry: when a speaker says “Xena is tall”, neither the speaker nor the hearer has privileged access to d . Indeed, even if both speaker and hearer are clear on issues such as the comparison class for which Xena’s tallness is being considered, and on normative issues such as the utility or aesthetics of various heights for this class, they still cannot be certain as to the appropriate value of d . The best they can do, we will suggest, is form a probabilistic model of the value of d , based on their prior experience of how *tall* has been used.²

Probabilistic, statistical, and fuzzy theories of vagueness have previously been brushed aside and discounted because of the failure of rather naive straw-man statistical approaches to account for the properties of vague predicates. Arguments against probabilistic approaches to vagueness such as those of Fine, Klein, Parikh, and Kamp and Partee [4, 5, 6, 7] argue against the use for threshold values of point estimates such as averages, geometric means, or arbitrary probability densities.³ And they’re right; such analyses are inadequate. There is another way though. Instead of merely using statistics to calculate a point estimate and then

² Accounts of the semantics of scalar predicates fall into two broad categories: those in which the semantics refers directly to a standard of comparison or threshold degree (d , above), which may be derived normatively, and those in which the semantics refers to a comparison class. Our main claims could be stated either way, but for simplicity and concreteness we stick to an account based on explicit mention of a degree. In terms of the statistical model we advance, we take it that the comparison class would act as a prior in determining the distribution over values for the threshold, though we leave the details as an open issue. Likewise, we do not explicitly consider here how normative factors affect the threshold: it suffices for our main claims that while normative factors may constrain the threshold, they will typically not constrain the threshold to be a single, perfectly precise value.

³ Actually, the arguments of [4, 5, 6, 7] fall short in a second way: these authors do not distinguish fuzzy truth-valued from probabilistic accounts of the truth of vague sentences. But in fact many criticisms of fuzzy accounts do not extend to statistical approaches. For example, fuzzy accounts fail to explain the contradictoriness of

forgetting about the original statistics, we take the information conveyed by a vague sentence to be a statistical distribution. A vague meaning provides a conditional distribution over facts, a probability function dependent on the value of an unbound variable, a variable that is itself constrained by a pragmatically determined statistical distribution.

3 Communication Under Uncertainty

We can imagine, contrary to fact, a system of linguistic communication in which all expressions are precise, in principle assessable by language users as definitely true or definitely false, i.e. there are no vague expressions. In fact, much of semantics as well as information theory is aimed at developing ways of guaranteeing precision. However, we are taking a somewhat different approach and asking whether there are situations guaranteeing a form of imprecision, vagueness. We apply standard results from information theory to obtain a model of communication under uncertainty in which the following hold: (i) communication under uncertainty is represented as a game against a malevolent nature who's goal is to disrupt communication, (ii) even in the face of a malevolent nature, non-trivial exchange of information is possible, (iii) the information conveyed by vague expressions is not truth values or the precise facts upon which a relation over degrees is determined, but a distribution over such facts, and (iv) vagueness is rational under uncertainty. The discussion in this section is presented informally by way of example, but the concepts and results have counterparts in the formal theory of information. Readers familiar with the basics of information theory should be able to make the relevant connections through the citations provided.

Definition 1. *Information channel:* An information channel [8] is a system of communication between a speaker and a hearer for conveying and interpreting observations about objects. The input to an information channel is a speaker's beliefs or observations and the output is the utterance heard.

Suppose a speaker observes and believes that Xena's height is defined by a degree of height (e.g. 175cm) and that it exceeds some standard. Both of which *are* objectively true or false. The speaker communicates the observation by uttering "Xena is tall", and in this case the information channel is the one where the speaker inputs observations about Xena's height relative to this standard and the hearer receives these via the utterance "Xena is tall".

Such a communicative setup can be represented statistically [8] as a vocabulary of observations and utterances along with probability distributions relating them. For a speaker or hearer to know how to use the channel, i.e. to know the communicative conventions, they must have knowledge of the space of observations and utterances and the probabilities governing them.

"Xena is tall and she is not tall" (cf. fn. 1), whereas a statistical account predicts this unproblematically, provided the same degree distribution is assumed for both uses of *tall*.

Definition 2. *Mutual information:* Mutual information [8] is a measure of overlap between the distributions of some observations.

E.g. if a speaker knows Xena’s height as the uniform distribution over the interval (1.7m, 1.8m) but a hearer knows it as the uniform distribution over the interval (1.75m, 1.85m) then there is an overlap, a positive amount of mutual information, in the speaker and hearer’s distributions. The mutual information is also larger than it would have been if the hearer knew Xena’s height as the distribution over the interval (1.3m, 1.4m), but smaller than it would have been had they agreed exactly.

In order to be as precise as possible, rational speakers and hearers should strive to have channel distributions that overlap as much as possible [8]. Perfect overlap is unachievable under uncertainty, but maximizing mutual information is still the best option [9, 10, 11].

Definition 3. *Information channel game:* An information channel game [11] is a game between a speaker and a malevolent nature. In an information channel game, the speaker wants utterances to convey observations as precisely as possible, as measured by mutual information, but nature counters with noise making this difficult.

Suppose again that a speaker believes Xena’s true height is defined by a degree of height which exceeds some standard, and communicates this by uttering “Xena is tall”. Now, though, suppose that every observation the speaker has ever made was interfered with via a series of funhouse mirrors created by a malevolent nature. Thus, when the speaker says “Xena is tall”, he is likely not only to be wrong as regards Xena’s height, but also to be misinformed as to what counts as *tall*. That is, his distribution over heights as well as the distribution over the relevant standard of comparison have been distorted by the evil nature. The hearer will then recover something which is neither Xena’s true height nor what the speaker takes Xena’s height to be. For rational speakers and hearers to know how to use this channel is to have sufficient information to cope with the channel’s noisy distributions. Because the funhouse mirrors are not controlled by the speakers or hearers, though, the best that they can do is to jointly figure out the distributions of height and its standard that overlap as much as possible with the distributions that were distorted by nature, the true distributions plus some noise.

Notice that in the funhouse mirror game speakers and hearers do not play against each other. They play together, on the same team as it were, against the noise-introducing malevolent nature. They succeed when they maximize the overlap, the mutual information, of their communicative conventions, the channel distributions, under this uncertainty.

Proposition 1. *Communication under uncertainty can be modeled as an information channel game.*

Rational models of linguistic communication are often presented as Lewisian signaling games [12, 13] in which speaker and hearer play on separate teams

with separate (though perhaps related) utilities. Such games have proved very fruitful, but they bring with them the baggage of strong common knowledge and rationality (CKR) assumptions and also (in the most common applications) fail to recognize the importance of shared uncertainty in communication. Complete CKR is a non-starting position, neither possible nor wanted because of unavoidable uncertainty and variability in communicative situations. What is wanted is a model of communication that can accommodate the uncertainty, whether limited or extreme, that speakers and hearers have. If (i) what you believe or even know is like what I believe or know up to a certain point and (ii) you believe this and can make guesses about or estimate where we differ, then (iii) what really matters most is whether you can communicate according to what you believe or know plus some guesses about how that might differ from what I believe or know. Insofar as what we are interested in in this paper is completely cooperative communication in which the speaker and hearer share common goals, we might say that rational communication under uncertainty is like talking to a noisy version of yourself.

The result is reminiscent of Davidson’s notion of *charity* [14], which “counsels us quite generally to prefer theories of interpretation that minimize disagreement” [15, p. xix].⁴ Davidson makes it clear that rational communication cannot and should not eliminate disagreement; it should offer a way to make disagreement possible and useful. Substituting CKR with charity, the problem rational speakers and hearers face is to do what’s best with respect to certain beliefs plus some uncertain or noisy beliefs. The current paper can be seen as a way to refine Davidson’s proposal, by introducing a general method for working out what is *best* in a precise and motivated way. To wit, we treat the disagreement between interlocutors as if it were introduced by a malevolent third party, and understanding communicative acts as part of an optimal strategy against that party.

For example, a rational speaker who tries to communicate an observation of Xena’s height with an utterance of “Xena is tall”, faces the problem that there is potential or actual variation in speaker and hearer beliefs about Xena, height, Xena’s height, and standards of *tallness*. Letting these differences be our funhouse mirrors, the speaker that knows how to communicate over the channel can calculate the optimal channel distributions in the face of the uncertainty or noise created by the mirrors.

Assuming that the above model is apropos, a basic and welcome result is that even in the face of a malevolent nature, non-trivial exchange of information is possible. That is:

Proposition 2. *Communication under uncertainty is possible.*

It’s not obvious it should be so, although many scholars implicitly assume something like the above without argument; in an information channel game, the possibility of communication under uncertainty can and must be demonstrated. Here is where standard mathematical results have a voice. In an information

⁴ Thanks to Nicholas Asher for pointing out the parallel.

channel game, provided limited amounts of uncertainty, (i) a minimax solution exists, (ii) it's value is a positive amount of mutual information or overlap, and (iii) when the noise or uncertainty becomes too large, as should be expected, communication will fail [10, 11].

Suppose that the speaker observes Xena's height and says "Xena is tall", and the hearer knows that Xena is a warrior princess and has some guesses about the heights of warrior princesses and related standards of *tallness*. Then the hearer will come to know something about Xena's height even if there's some disagreement about what counts as *tall* for a warrior princess. This follows from the existence of a minimax solution with positive mutual information for limited amounts of uncertainty. On the other hand, if the hearer does not know what sort of entity Xena is, or thinks warrior princesses are a kind of flower, things will not turn out so well. For in this case there will be a very large mismatch between the speaker and hearer's distributions of Xena's height and the relevant standard or what counts as *tall*. In such circumstances it is quite possible that interpretation will fail, in the sense that the communicative act will not convey any information about Xena's height: chalk up a victory to malevolent nature. Fortunately, interlocutors commonly know enough about each other that they can make use of an information channel and succeed in meaningful information exchange.

There are two kinds of distributions that rational speakers and hearers must know to use an information channel: the prior probabilities and the conditional probabilities. In our example, the former are the distributions of Xena's height and *tallness* and the latter are distributions relating Xena's height to *tallness* and vice versa. The most important of these is the posterior distribution of Xena's height given an utterance of "Xena is tall". This allows a hearer to recover (estimates of) the speaker's observations.

Proposition 3. *The information conveyed⁵ in communication under uncertainty is the posterior probability distribution of the input observations.*

Typically in an information theoretic model the hearer decodes an utterance as follows: it is the most likely input observation given the output utterance. But this is problematic if a unique best observation does not exist, since selecting one reduces to choosing randomly from the posterior. As an alternative, we suggest not worrying about decoding at all. The information conveyed by communication under uncertainty is more generally the increased knowledge that a hearer has about the possible observations a speaker made. E.g. if a speaker says "Xena is tall", what a hearer knows for sure is a distribution describing the speaker's observation of Xena's height. This even holds in the case when there is no way for the hearer to recover a unique best threshold of height! The information conveyed

⁵ This is not a measure or quantity of information or information content such as entropy, self information, or mutual information. It is more akin to the notions of *semantic information* or *meaningful data* [see 16]. It is a way of drawing inferences or mapping from observations to unobserved but possible observations; not the facts but a tool that could be used to narrow in on them.

in the communication is not an exact observation of Xena's height relative to the standard. Rather, the information conveyed is whatever knowledge would, in the best case, allow decoding, and, in the worst case, allow the hearer to draw conclusions about what possible observations a speaker could have made of Xena's height. This takes us to propose a working definition of what it is for a claim involving a scalar predicate to be vague:

Definition 4. *Vague:* A claim involving a scalar predicate about an entity is vague *iff* for some range of observations of the value the entity takes on the scale, the information conveyed does not reliably determine whether those observations are compatible with the truth of the claim.

To see how this applies, consider again the predicate *tall*, which, as noted, is typically defined as a relation between a measure of height and a standard of comparison or comparison class. Suppose that the claim "Xena is tall" is made after an observation that Xena is 1.8m tall. Put in terms of information conveyed, the question is whether on the basis of hearing and accepting the claim "Xena is tall", a hearer would be in a position to say with complete certainty whether the observation that Xena is 2m tall is correct or not. And the answer is no, because the hearer has only a statistical distribution for the standard of height required for *tall* to be predicated of an individual, and there may be a non-zero probability that the standard is higher than 2m. Thus *tall* is vague.

We may contrast the case of *tall* with the case of *taller than 1.8m*. To the extent that the degree *1.8m* is itself precise, the expression *taller than 1.8m* turns out not to be vague under our definition. For any observation that Xena is X meters tall, for some X, it can reliably be determined whether the claim "Xena is taller than 1.8m" is compatible with the observation.

We note here that the inability of the information conveyed to reveal a precise observation creates the appearance that facts underlying vague expressions are indeterminate, gappy, or simultaneously true and false. But this is only an illusion much like an observation made through a funhouse mirror. There is a fact to the matter of an observation that underlies a vague sentence, but it is not part of the information conveyed.

At this point you may think that this definition reiterates a common definition of vagueness vis-a-vis borderline cases. Crucially it does not. This is clearer if we recognize that communication games include not just the input of an observed height and standard or comparison class but the context, the situation, the goals of communication, etc. If any of these, too, can only be known up to the posterior distribution, then the cases in which the information conveyed is unreliable for the determination of truth multiply. Vagueness, here, is uncertainty in the information conveyed and whether it can be used to determine the facts.

With the definition of vagueness in hand, we come to the central result:

Proposition 4. *Vagueness is rational under uncertainty.*

The rationality of vagueness follows directly from its definition and from constraints on maximizing mutual information with respect to uncertainty. If

there is any uncertainty between a speaker and hearer then the resulting channel distributions, the best or rational way for the speaker and hearer to use the channel, have a margin of error or non-negative categorization error [8]. I.e. the information conveyed by an utterance of “Xena is tall” will be a non-trivial distribution over Xena’s height. Vagueness is immediate.

4 Conclusions

Most theories of vagueness agree, as we do, on the point that vague expressions are uncertain or imprecise [17, 18, 19]. Where we’ve gone further is by giving a reason for the imprecision and saying how rational speakers and hearers cope with it and why they are happy to accept it or even prefer it. And from that, we can say why the sorites and epistemic uncertainty are not surprising.

Consider Xena and Callisto in a sorites sequence: if both are *tall*, then Callisto may be a hair’s breadth shorter than Xena and yet still *tall*. Without reference to heights or standards, this case seems to, incorrectly, confirm a general inductive step. In our model, the distribution by which this inference is made treats Xena and Callisto as all but indistinguishable; indeed, we may even be psychologically unable to distinguish the information conveyed about adjacent individuals in a sorites sequence. So if Xena is *tall*, we correctly predict an inference from the first to the second, but we also predict that the inference should not continue much further. According to the statistical model we employ, each distribution along the scale is (slightly) different, and to the extent that speakers and hearers reason with such models, it is predicted that statements made about individuals far apart on the scale convey different information, i.e. different distributions over possible observations. The confidence as well as the uncertainty that speakers and hearers have does not remain constant moving along the scale. Thus if we were to allow Callisto to be 30,000 hair’s breadths shorter than Xena, then the inference from “Xena is tall” to “Callisto is tall” will not necessarily be made with the same confidence. This is, of course, precisely what we observe.

Putting aside the sorites, consider our proposal in the light of previous linguistic accounts of vagueness. The logical revisions in supervaluationist [4, 7] and dynamic accounts [20] try to eliminate imprecision rather than bringing it to the fore. Consider the issue of higher-order vagueness: the vagueness which arises in defining precise regions of vagueness. If in a certain region on a scale truth values are indeterminate, then these new boundaries themselves fail to be precise. The awkward result is that imprecision remains as ever more regions of higher-order vagueness are added. Since we embrace imprecision within our model with statistical uncertainty, we never have to make a decision about where precision stops and imprecision begins.

It is also quite common to attack vagueness with more sophisticated ways of comparing individuals to standards of comparison or comparison classes [5, 19]. This hints at the role of uncertainty and may provide an adequate model-theoretic semantics, but these theories are silent regarding the source of the ordering relations. Are they the products of thought? Should we expect them to

be viable strategies in communication? We do not disagree with these modifications, but do believe they can be better motivated by looking at the properties of rational communication.⁶

Our main claim is that vague scalar adjectives convey insufficient information to decide the degrees and standards of comparison upon which they supervene. I.e. a vague predicate has truth conditions and it will be true or false but these facts are absent from the information conveyed in communication. This is similar to Williamson's [24, 25] epistemic view, which characterizes vagueness as a form of essential ignorance. When vague expressions are used, as speakers and hearers, we are only concerned with knowledge up to a margin of error; we are ignorant of the precise facts. Williamson's claims are merely observational, perhaps stipulative, but they can be substantiated by a formal model such as ours.⁷ The epistemic view, on its own, does not make it especially clear why ignorance bubbles up into communication (it's not obviously necessary; precise languages are staples in both philosophy and information theory) or why margins of error exist. In an information channel, though, we have ignorance because margins of error are unavoidable thus making ignorance rational under uncertainty. The ignorance in vagueness is strategic ignorance. Additionally, whereas Williamson [24, 25] gives a somewhat vague explanation for the widely accepted idea that vagueness supervenes on precision, we make it precise: vagueness is the product of limits on the information conveyed, information which supervenes on the precise facts in a statistical sense.

References

- [1] Max J. Cresswell. The semantics of degree. In Barbara Partee, editor, *Montague Grammar*, pages 261–292. Academic Press, New York, 1976.
- [2] Angelika Kratzer. The notional category of modality. In *Words, Worlds, and Contexts: New Approaches in Word Semantics*, pages 38–74. de Gruyter, Berlin, 1981.
- [3] Kai von Fintel. *Restrictions on Quantifier Domains*. PhD thesis, University of Massachusetts at Amherst, May 1994.
- [4] Kit Fine. Vagueness, truth and logic. *Synthese*, 30(3-4):265–300, 1975.
- [5] Ewan Klein. The semantics of positive and comparative adjectives. *Linguistics and Philosophy*, 4:1–45, 1980.

⁶ Pragmatic accounts have also contributed to our understanding of vagueness: we are friends of the idea that vagueness is tied to the purposes and practices of communication and that the uncertainties present in the communicative situation impact what speakers and hearers are willing to accept or find useful [21, 22]. There have also been other models of vagueness as language games. As we do, these [6, 23] derive vagueness from assumptions about communication, but our model is unique in being an information channel game and in providing a criterion for vagueness along with a notion of the information conveyed by vague expressions.

⁷ De Jaeger [23] also notes a deep connection between vague communication and epistemic uncertainty as suggested by a signaling game with correlated equilibria.

- [6] Rohit Parikh. Vagueness and utility: The semantics of common nouns. *Linguistics and Philosophy*, 17(6):521–535, 1994.
- [7] Hans Kamp and Barbara Partee. Prototype theory and compositionality. *Cognition*, 57:129–191, 1995.
- [8] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley, 2006.
- [9] Sergio Verdu. The exponential distribution in information theory. *Problems of Information Transmission*, 32(1):100–111, 1996.
- [10] David Haussler. A general minimax result for relative entropy. *IEEE Transactions on Information Theory*, 43(4):1276–1280, 1997.
- [11] Rudolf Mathar and Anke Schmeink. Saddle-point properties and nash equilibria for channel games. *EURASIP Journal on Advances in Signal Processing*, 2009(823513), 2009.
- [12] David Lewis. *Convention: A Philosophical Study*. Harvard University Press, 1969.
- [13] Anton Benz, Gerhard Jäger, and Robert van Rooij. An introduction to game theory for linguists. In Anton Benz, Gerhard Jäger, and Robert van Rooij, editors, *Game Theory and Pragmatics*. Palgrave Macmillan, 2005.
- [14] Donald Davidson. On the very idea of a conceptual scheme. *Proceedings and Addresses of the American Philosophical Association*, 47:5–20, 1974.
- [15] Donald Davidson. *Inquiries into Truth and Interpretation*. Oxford University Press, 2001.
- [16] Luciano Floridi. Philosophical conceptions of information. *Lecture Notes in Computer Science*, pages 13–53, 2009.
- [17] Delia Graff Fara and Timothy Williamson. Vagueness. In Delia Graff Fara and Timothy Williamson, editors, *Vagueness (International Research Library of Philosophy)*. Ashgate, 2002.
- [18] Chris Kennedy. Ambiguity and vagueness: An overview (to appear). In Claudia Maienborn, Klaus von Heusinger, and Paul Portner, editors, *The Handbook of Semantics*. to appear.
- [19] Robert van Rooij. Vagueness and linguistics. In Giuseppina Ronzitti, editor, *The Vagueness Handbook*. 2009 (to appear).
- [20] Chris Barker. The dynamics of vagueness. *Linguistics and Philosophy*, 25: 1–36, 2002.
- [21] Peter Lasersohn. Pragmatic halos. *Language*, 75(3):522–551, 1999.
- [22] Delia Graff Fara. Shifting sands: An interest-relative theory of vagueness. *Philosophical Topics*, 28(1):45–81, 2000.
- [23] Kris de Jaeger. A game-theoretic rationale for vagueness. *Linguistics and Philosophy*, 26(5):637–659, 2004.
- [24] Timothy Williamson. Vagueness and ignorance. *Proceedings of the Aristotelian Society*, pages 145–162, 1992.
- [25] Timothy Williamson. Imagination, stipulation and vagueness. *Philosophical Issues*, 8(Truth):215–228, 1997.