Probabilistic approaches to vagueness and semantic competency

Abstract

Wright (1975) holds that the following two theses are jointly incoherent: (T1) Rules determine correct language use. (T2) These rules are discoverable via internal reflection on language use. I argue that incoherence is derivable from (T1) alone and examine two types of probabilistic accounts that model a modification of (T1), one in terms of inexact knowledge, the other in terms of viewing semantic rules as reasons for linguistic actions. Both accommodate tolerance by breaking the link between justified assertion and truth, but incoherence threatens their conception of justified assertion (the 'relocation problem'). I argue that the rules-as-reasons approach can relocate sharp boundaries to a place where they are not only more tolerable, but to be expected.

1 Introduction

In his classic paper, Wright (1975) introduced a way of thinking about vagueness, namely *tolerance*, which has become a norm of philosophical discourse in the vagueness literature. Wright's paper draws no firm positive conclusions about how to resolve the problems tolerance creates for accounts of vagueness, however, it is still worth revisiting. The reason for this is that Wright clearly grapples with what might be called 'the relocation problem of vagueness'. Wright gives us a dilemma: either one must admit that our language competency is incoherent, or we must give up internal reflection as a method of discovering truths about our semantic competency. However, Wright is keenly aware that this dilemma runs deep. Even if one can finesse one's way out of the dilemma at one level, (say at the level of truth-conditions), a similar, equally unpalatable dilemma will occur again at another level (such as that of justified assertion).

In this paper, we will consider some examples of relatively recent approaches to the semantics of vague expressions which employ Bayesian probability calculus as a modelling tool. Our goal will be to explore whether such approaches help us to find a depth at which Wright's dilemma ceases to be unpalatable.

1.1 Tolerance and Incoherence

The *governing view* of understanding and meaning as defined and criticised in (Wright, 1975) is the combination of two theses:

(T1) The allowable/correct use of an expression is *completely determined by* a set of rules.

(T2) The properties of these rules may be discovered, not only via behavioural evidence, but also via an internal reflection on what *justifies* the application of a given expression.

(T1) identifies the role of the rules, and (T2) identifies how we might determine what properties these rules have. Wright acknowledges the kinds of worries that (T1) and (T2) might generate with respect to rule-following, however he focusses his main discussion on a tension between (T1) and (T2). I will follow him in this approach. According to (T2), we can gain insight into the rules that govern (correct) language use via internal reflection. For example, by reflecting on the use of vague predicates such as 'is tall', we might judge that, in every case, there will always be a small enough difference in height such that if 'tall' applies to someone, 'tall' will also apply to someone else who differs by only that amount. Whether we count these observations as rules for the application of 'tall' or "properties" of some underlying rules, as Wright has famously pointed out, we are led, via (T1) and (T2) into deriving contradictory instructions for predicate use. Given repeated applications of rules such as the ones above, we would be led to apply 'tall' to individuals whose heights we have independent justification for classifying as being not tall. The (properties of) rules (such as the one above for 'tall') generalise into what Wright coined as *tolerance*: that small differences between properties/features of objects, such as shades for colour terms, do not give rise to differences in whether those terms should be applied. Contradictory instructions cannot support consistent linguistic behaviour. But inconsistent linguistic behaviour cannot be either informative or allowable/correct, hence those rules cannot be ones which completely determine the allowable/correct uses of those expressions. (T1), Wright concludes, is inconsistent with (T2).

1.2 Wright on Degrees

Wright considers one way that a defender of the governing view might try to resolve the conflict between (T1) and (T2), namely, the introduction of degrees. Wright (1975, §IV) phrases his discussion in terms of "states of affairs justifying applications of expressions" and "degrees of such justice", and the degrees of justice he has in mind are those which employ logics of degree. By this, Wright does not simply mean any formalism with degrees, but one which makes use of degrees of truth. In other words, positions which view "the paradoxical reasoning as essentially depending upon the constraints of bivalence" (Wright, 1975, p. 349). Wright argues that the introduction of degrees gets the governing view no further "for among [the degrees of justification] we still have to distinguish those with which for practical purposes the application of the predicate is to be associated; otherwise we have not, in repudiating bivalence done anything to replace the old connection between justified assertion and truth" (Wright, 1975, p. 350).

Multiple-valued accounts do, however, allow for a weakening of tolerance. Based on the standard 'short' Sorites premise in (1)[1] we can provide a meta-semantic principle of tolerance in (2). $x \sim_P y$ means that x is sufficiently similar and/or indiscriminable from y in the relevant P respects. $\mathcal{I}$ is a semi-formal representation of an interpretation function:[2]

$$\forall x, y(P(x) \wedge x \sim_P y \rightarrow P(y)) \tag{1}$$

$$\text{For all } a, b, \text{ if } \mathcal{I}(P)(a) = 1 \text{ and } a\ \mathcal{I}(\sim_P)\ b, \text{then } \mathcal{I}(P)(b) = 1 \tag{2}$$

This principle will fail on a multi-valued account which will support only the weaker:

$$\text{For all } a, b, \text{ if } \mathcal{I}(P)(a) = 1 \text{ and } a\ \mathcal{I}(\sim_P)\ b, \text{ then } \mathcal{I}(P)(b) \neq 0 \tag{3}$$

Importantly, such accounts still support tolerance principles (ϵ is some arbitrarily small value, $n \in [0, 1]$):

$$\text{For all } a, b, \text{ if } \mathcal{I}(P)(a) = n \text{ and } a\ \mathcal{I}(\sim_P)\ b, \text{ then } \mathcal{I}(P)(b) = n \pm \epsilon \tag{4}$$

However, this is precisely what Wright argues is still problematic. Now we can ask, what is the smallest value, m, in $[0, 1]$ such that $\mathcal{I}(\phi) = m$ and ϕ is assertible? If there is a value, then there is a shift from m to $m - \epsilon$, such that a formula such as $P(a)$ with value m is (correctly) assertible, but a formula $P(b)$ with value $m - \epsilon$ is not. In other words, the following a higher order tolerance principle in terms of correct assertion fails in tension with (T2):

$$\text{If } P(a) \text{ is correctly assertable and } a \sim_P b, \text{then } P(b) \text{ is correctly assertable.} \tag{5}$$

If there is no such value, m, paradox arises at the higher level, since it implies that a formula with value 0 is assertible.

To re-emphasise the point, a major insight of Wright's paper is that it recognises how deep the vagueness problem goes. If we ease away tolerance problems at one level, we get what I have called the *relocation problem*: Removing/easing away a sharp boundary in one place (e.g. at the level of truth conditions) simply relocates it somewhere else (e.g. at the level of correct use/assertibility).

I do not wish to dispute Wright's conclusion regarding multiple truth-value positions. However, a major goal of this paper will be to outline a notion of degree that does get us further. I will argue that, even though Wright's argument succeeds in showing that a degree-based governing view is incoherent, when we weaken (T1) and incorporate an alternative notion of degree, the same conclusion does not follow. Crucially, however, this will mean facing up to the relocation problem. We must accept that sharp boundaries will not go away entirely. The trick will be to relocate them somewhere where sharp boundaries are acceptable.

In §2, I point out that (T2), need not be the source of incoherence, since incoherence can be generated from (T1) alone. I then discuss two ways of adjusting/altering (T1) that seem to avoid incoherence altogether. The first, our exemplar for which will be Lassiter (2011), weakens (T1) to a claim about inexact knowledge about correct application of predicates. The second, which I will go on to defend as the basis of addressing the relocation problem, adjusts (T1) to view rules as merely guiding the application of predicates, not as determining such application. As such, on the second view, rules can be seen as reasons. Our exemplar for this second approach will be Sutton (2013, 2015). I will argue that a shift in the understanding of the role of rules, what I shall call the *rules-as-reasons* approach, is not incoherent. We can and do have conflicting reasons for action without any resultant inconsistent

behaviour. In §3, I shall outline two probabilisitic models which reflect the two ways of reinterpreting (T1). In §4, I shall argue that both models avoid incoherence at the first level by switching to a communication centred stance and thereby severing 'the old connection between justified assertion and truth'. Nonetheless, we still do not escape the relocation problem. On probabilistic views, the relocation problem arises at the level of justified/correct assertion. I will then assess how this problem can be addressed and argue that, by shifting to a rules-as-reasons perspective, we can better understand reasoning about the use of language, and so reveal some interesting conclusions about semantic competency. In short, I will suggest that only by seeing rule as reasons may we be able to embrace a form of tolerance without its intolerable consequences.

2 Incoherence without Tolerance

In order to motivate adjustments of (T1), I will argue that incoherence lies not only in the tension between (T1) and (T2), as Wright had it, but in (T1) itself. If (T1) itself can be shown to be the source of the trouble, irrespective of (T2), then we will have a good reason to adjust (T1), retain (T2), and then see if principles of tolerance still generate incoherence.

2.1 Troubles with (T1)

With a little unpacking, we can see that (T1) is what lies behind the orthodox approach to semantics: to understand a sentence/utterance, is to know under what conditions it would be true.[3] The sets of rules referred to in (T1) are meant to completely determine allowable/correct uses of expressions. Arguably, how allowable an application of an expression is, or whether a use of an expression is correct, need not be tied to whether that use of the expression was one in which something true was said. For example, in English, it may be in some sense incorrect to describe someone's (natural) hair colour using 'orange' as opposed to 'strawberry blonde', 'ginger', or even 'red'. Nonetheless, the spirit of the governing view is to associate the correct application of expressions with the *true* application of expressions. Hence, we can substitute (T1) with a related idea of the governing view, albeit one put in terms of truth-conditions: competent speakers know under what conditions expressions would apply truly.

Rewording (T1) in terms of competency and knowledge is not equivalent to an inclusion of (T2) in (T1). No internal reflection on the rules of language use needs to take place in speakers for it to be true that competent speakers know the truth conditions for sentences/utterances. Young children can be competent speakers and thereby know the truth-conditions of utterances (as much as adult competent speakers do), however, it is doubtful that young children need to be able to internally reflect upon the rules of usage in order to have this knowledge. Furthermore, although (T1) merely states that there are rules that govern language use, implicit in this claim is that competent speakers know (in some sense of 'know') these rules and follow them. On the governing view, knowing these rules is equivalent to knowing the truth-conditions for expressions of one's language.

The problem for (T1), understood in terms of knowledge of truth conditions, is that vague terms seem to be the

very kind of expression for which such knowledge fails. Irrespective of whether one believes that vague expressions have sharp boundaries, the blurred nature of vague predicates means that competent speakers tend to be uncertain as to the exact truth-conditions of that expression.[4] In other words, we have the argument below. For now, I take (P1) to be true. If it were not, then this would imply that we would all be systematically erring from some ideal language use ability. This is not objectionable in itself, and as we shall see, some ways of adjusting the governing view amount to just this claim. (P2) is both a statement of the governing view in terms of truth conditions and an explication of (P1). (P3) is the observation that vague terms are sources of gaps in our knowledge of truth conditions.

(P1) There are fully competent speakers of languages.

(P2) If a speaker, S, is fully competent with a predicate F, then S knows under what conditions F would be truly applied.

(P3) For vague predicates, no speakers know (entirely/exactly) under what conditions F would be truly applied.

(C1) No speakers are fully competent.

This argument should not be surprising or contentious. It simply points out that one has a hard time defending a notion of semantic competency based upon an all-or-nothing account of knowledge of truth-conditions, where truth conditions are based upon a bivalent view of truth. If nothing else, the vast literature on vagueness and the sorites in the last one hundred years should be testament to a recognition of this tension. However, as indicated by this description of the tension, one way to defend the governing view is to finesse one's account of what it takes to *know* the truth-conditions of a vague predicate. Indeed, there are well-known epistemic approaches to vagueness that are fairly true to the governing view (but proposed well after Wright's paper), that adjust (P2), and thereby our understanding of (P1). See, for example, (Sorensen, 1988; Williamson, 1994).

Since (P1) contradicts (C1), the (simplistic) governing view is incoherent. Importantly though, if we accept (P2) as approximating, or following from, (T1), then this incoherence can be derived without (T2). This indicates that the trouble with the governing view may well stem from (T1), not from (T2). In further support of this claim, by the end of this paper, I will have suggested an interpretation for tolerance that is consistent with a revised form of (T1).

In summary: A simple-minded version of the governing view faces major challenges defending (T1) in the light of data regarding our knowledge of the truth-conditions of vague predicates. However, if one thinks that vagueness is at root epistemic, then a marginal adjustment to the governing view can be made. Rather than exact knowledge of truth-conditions, one can argue that we have inexact knowledge of truth-conditions.

2.2 Inexact Knowledge

The reasoning from (P1-3) to (C1) is not a knock-down argument against all forms of the governing view. Those who defend the governing view can make minor adjustments to some of the premises to avoid contradiction. For example, Williamson's *epistemicism* (Williamson, 1994) could be read as a response to the kind of worry that the

above argument generates. His positive view could be interpreted as a shift of understanding with respect to (P1). This amounts to replacing the explication of (P1), namely (P2), with (P2*), and therefore also replacing (P3) with (P3*):

(P2*) If a speaker, S, is fully competent with a predicate F, then S knows under what conditions F would be truly applied, but only within a margin for error.

(P3*) For vague predicates, no speakers know under what conditions F would be truly applied without any margin for error.

With these adjustments, the contradiction is blocked. The task of this paper will not, however, be to respond directly to Sorensen's or Williamson's epistemicist accounts of vagueness. Some have found the unknowable semantic facts posited by epistemicism to be implausible (see, for example, Smith, 2008). Worries about implausibility have led to the development of communication-centred versions of the inexact knowledge approach (Frazee & Beaver, 2010; Lassiter, 2011; Lassiter & Goodman, 2015) that openly try to relieve some of the discomfort that Epistemicism creates. We will consider these communication centred approaches in §3.

2.3 Rules as Reasons

Rather than attempting to maintain the governing view whilst being forced to reinterpret our understanding of competency, we can instead negate the statement of the governing view made in (P2). If the original argument is sound, then from the contradiction between (P1) and (C1), we may infer the negation of (P2) (which amounts to a denial of the governing view's (T1)):

(C2) Speakers can be fully competent without knowing the truth conditions for the predicates they are competent with.

If (C2) is correct, this leaves room for an account on which being competent with an expression can allow *not knowing*, in some possibly restricted sense, under what conditions a predicate would be true. If we maintain some role for rules in accounting for linguistic competence, we can maintain the governing view's notion of competency: being competent is, in some sense, knowing the rules of the game. However, we can drop the governing view's thesis that these rules *determine* correct usage.[5] If rules do not determine correct usage, then a weaker alternative is that they simply guide it. This suggestion is comparable to one made by Travis in his discussion of vagueness and the sorites:

> "we suppose that the first member [of the sorites series] is unmistakably red. That, plus indistinguishability, gives us excellent reason to say the same of the next member. Such is transmitted not much diminished to the next member, and so on. As we go through the series, however, the force of reason transmitted by this process diminishes. The reason is that a different, though parallel chain of reasons is transmitted from the last member of the series ... Since the series is long, there may be many cases where one reason clearly outweighs the other. The

first and last members are two such. But there may also be many cases where weighing yields no clear results." (Travis, 1985, p. 361).

If rules merely guide correct usage, rules can be seen as reasons for linguistic actions. For example, there may be different rules informing the usage of a term in a particular case, and an agent's role is to weigh up which rule to apply. In other words, rules could be seen to provide reasons whether or not to apply the expression in each particular case. Being competent with that expression will then be a matter of rationally balancing and weighing up these reasons in order to arrive at a linguistic action.

Similar proposals to the rules-as-reasons approach have been made in the philosophical literature on vagueness, most notably Raffman (2014) and Rayo (2008). Rayo mentions, and Raffman emphasises, that our competency demands that we make *arbitrary* decisions to stop along (forced march) sorites series, or when making judgements in borderline areas (Égré (2011) defends a similar position which will be addressed in more detail in section 4). Rayo stresses that tolerance "is a feature of our ability to use linguistic representations", not "a semantic principle governing the correctness of our assertions". Furthermore, that "Language-mastery is not a matter of applying semantic rules; it is a matter of making semi-principled decisions about how to partition the context set in light of past linguistic usage and the location of the gap" (Rayo, 2008, p. 366). Both of these lines of thought are compatible with, and perhaps even complementary to the final position to be defended here. Some further discussion will be made in §4.2.

2.4 Tolerance and semantic competency

There is a strong link between accepting or weakening (T1), and how we should view tolerance principles. To be sure, in assessing the consistency of (T1) and (T2), what gets the governing view into difficulties is that, by (T1), tolerance principles (which are discoverable via internal reflection on linguistic practices) *determine* correct linguistic practice. To this extent, in evaluating different conceptions of semantic competency (which will be a focus of this paper), it will also be important, as a test of these conceptions, whether, paired with tolerance principles, they lead to incoherence.

In §3, two proposals for modelling these adjustments/revisions of (T1)/(P2) will be presented, both of which employ Bayesian probability calculus. The first adopts the strategy from §2.2 in that it treats out linguistic knowledge as inexact. The second will exemplify the rules-as-reasons approach.[6]

3 Probabilistic Approaches to Vagueness

A long standing manoeuvre in the vagueness literature is to appeal to degrees of truth. Such a move does not qualify as an adjustment of competency in either of the ways just outlined. Semantics based on degrees of truth replace two values with a multitude. All that this implies for the governing view of competency is that competent users grasp degrees-of-truth conditions rather than truth conditions.

Probabilistic approaches differ from degrees of truth approaches. In §3.1, I will give a brief overview of some extant probabilistic approaches, and will connect these to the inexact knowledge and rules-as-reasons conceptions of semantic competency. In §§3.2-3.3, we will consider examples of probabilistic accounts that fall under each of these conceptions.

3.1 Overview of probabilistic treatments of vagueness

A number of probabilistic treatments of vagueness have been suggested over the past one hundred or so years. The two early examples are Borel (1907/2014)[7] and Black (1937), but more recently, interest in probabilistic approaches were reignited starting with Edgington (1992, 1997). Edgington's approach is not strictly probabilistic, however. Rather, Edgington draws structural parallels between Bayesian probability calculus governing uncertainty, and a suitable logic for vagueness. She argues that this logic, the values for which are *verities* (degrees of closeness to clear truth), captures vagueness better than degree-functional approaches such as fuzzy logic (for a response and an incorporation of both uncertainty and degrees of truth, see MacFarlane (2009)). In fact, Edgington (1997) explicitly argues that vagueness cannot be reduced to uncertainty (for responses to these claims, see, Sutton (2013); Lassiter & Goodman (2015)).[8] However, Edgington's work has had an impact on the development of other probabilistic accounts such as Lassiter (2011); Lassiter & Goodman (2015); Sutton (2013). In a slightly different tradition, and seemingly arrived at independently (via work into prototype theoretic representations), vagueness has also been treated probabilistically in Hampton (2007) (however, the treatment of vague concepts via prototypes (represented, more or less, in probabilistic terms) goes back at least as far as Rosch & Mervis (1975)). Also, Hampton's work has been developed and elaborated upon in the Conceptual Spaces paradigm by Decock & Douven (2014).

Broadly speaking, these accounts fall into two main types. One approach is to describe either a threshold, or a set of sharpened interpretations for a vague predicate, and then use probability theory to model uncertainty about the location of this threshold (or which precisification is being expressed). This approach forms the basis for Lassiter (2011); Frazee & Beaver (2010); Lassiter & Goodman (2015); Larsson & Fernández (2014); MacFarlane (2009).[9] An alternative is to treat categories/concepts as inherently graded, and so use probability theory to characterise the probability that a speaker would apply a predicate in a given case. This second approach is, arguably, of a type with Borel's and Black's account (Égré & Barberousse, 2014).[10] It is more explicitly defended in Hampton (2007); Sutton (2015).[11]

These two types of theories also match, approximately, with the inexact knowledge, and rules-as-reasons approaches detailed in §2. Employing either a set of sharp precisifications, or a set/range of threshold values and defining a probability distribution over them is one way to provide a formal basis for defining inexact linguistic knowledge. One assumes that some sharp semantic value/threshold is in play in a communicative context, and then treats agents and tracking and approximating this value by updating their belief spaces. An example of this approach is Lassiter (2011) who adopts the spirit of the epistemicist response to the competency argument, but attempts to make it more palatable. On Lassiter's view, vague utterances do have exact truth conditions at any given

context, however competency consists of being able to reduce uncertainty about these truth conditions sufficient for one's communicative purposes. This departs from the governing view in that the precise truth-conditional content of an utterance needn't be grasped by competent users. However, it remains close to the governing view in that the possible (precisified) contents of an utterance can still be characterised in terms of truth conditions. After Lassiter's account, I call this the *Probabilistic Linguistic Knowledge* (PLK) approach.

If an account adopts a different probability model which encodes a notion of graded judgements without the use of thresholds, then one need not commit to there being a fact of the matter or a 'true' boundary in any particular context. It is not that agents are approximating/reasoning towards some precise language, they can have a fully correct, competent grasp of a meaning and yet still be uncertain whether or not a predicate applies in a given case. An example of this approach is Sutton (2015). This account retains part of the governing view of competency, since it maintains that one grasps what is expressed by utterances containing vague predicates. However, it departs from the governing view by denying that what one grasps are truth conditional contents. The motivation for this move is, in part, given by a demand for a plausible account of semantic learning (also see van Eijck & Lappin (2012)). In overly simple terms, learners are exposed to too impoverished a data set to determine the truth conditions for an expression, but can nonetheless become competent users of that expression. Therefore, what speakers learn (what they grasp), must be something other than a truth-conditional content. Instead, vague predicates are viewed as encoding uncertainty.

Put another way, to be hesitant in borderline cases is compatible with being competent in a language where 'being competent in a language' means fully grasping a set of rules. For some set of rules, being truly torn about whether or not 'red' applies in a borderline case is part of what being competent with 'red' requires.[12] Van Eijck & Lappin (2012) provide a helpful gloss on this notion. One can conceive of this uncertainty as the probability that a competent speaker would, all else being equal, assent to a predication/would make e.g. a 'red' judgement. Contents are characterised by correlations between utterance types and states of affairs. Linguistic judgements are formed on the basis of these correlations. Call this the *Probabilistic Judgements and Correlations* (PJC) approach.

Another account that should be mentioned is Égré (2009, 2011). Égré suggests that probabilities may be taken to reflect that "if $p(P(x)) > p(P(y))$, [...] the expected frequency of judgements of the form 'x is P' is greater than the expected frequency of judgements of the form 'y is P'" (Égré, 2011, p. 76) at least relative to a constrained categorization task. I return to Égré's account, and a comparison between it and my own proposal in section 4.2.

In the remains of this section, I will present examples of the PLK and PJC approaches. In §4, I will argue that both models can escape Wright's charge of incoherence, although I will suggest that to do so, PLK must adopt something closer to the rules-as-reasons approach.

3.2 The PLK Approach

Although there are more recent examples of PLK we will focus on Lassiter's (2011) account. Lassiter & Goodman (2015) make significant improvements on the account in Lassiter (2011) by incorporating a pragmatic iterated

reasoning model (see, amongst others, Franke (2009); Lassiter & Goodman (2014)). This allows, for example, an integration of the impact of the comparison class that is underdeveloped in Lassiter (2011). However, with respect to the presence of probabilistic thresholds towards which agents reason, the accounts remain very close.[13]

Lassiter's version of PLK, in part, builds on the work of Stalnaker (1978) and Barker (2002) who argue that utterances can carry metalinguistic as well as factual/worldly information. For example: if George is unfamiliar with the standards for tallness in Daniel's linguistic community, but she knows that Jess is 188cm in height, then an utterance of 'Jess is tall' can help George narrow down what the standards for height are in Daniel's linguistic community (how the likes of Daniel use 'tall').

Barker's account is solidly within the dynamic semantics paradigm (see, for example, (Groenendijk & Stokhof, 1991)). However, Lassiter persuasively points out that such accounts should be enriched with a probabilistic notion of gradience. Just as there is a good case for being able to represent varying levels of uncertainty in our beliefs about the world, especially as a result of updating our beliefs on the basis of a non-wholly-reliable source, there is also a case for introducing uncertainty about how words are being used.

Lassiter defines a probabilistic belief space that can reflect both worldly and metalinguistic uncertainty. Roughly, worldly uncertainty comes out as a probability distribution over possible worlds. Metalinguistic uncertainty is captured as a probability distribution over precisifications of natural language terms. For example, 'tall' could be understood in terms of a threshold for heights. Different values for the threshold could be different ways to precisify 'tall'. Metalinguistic uncertainty with respect to tall is characterised as a probability distribution over precise languages in which the threshold for 'tall' is set at different values. However, worldly and metalinguistic information are not independent of one another.

Formally, Lassiter defines a probabilistic belief space (W, L, μ), and a probability function $\mu : (W, L) \rightarrow [0, 1]$ where W is a set of possible worlds and L is a set of possible languages. The probability an agent assigns a possible world will then be the sum of the probabilities of the world-language pairs it occurs in, *mutatis mutandis* for a possible language. Utterances using vague terms will then be interpreted as Bayesian updates on the probabilistic belief space.

Here is an example. For simplicity, assume that our model contains just one possible world (so we have no worldly uncertainty), which is characterised as a proposition about Jess's height.

$$w_1 = \{height(jess){=}188cm\}$$

However, we may be uncertain about what the standard for *tall* is. Assume that our probability space contains five sharp interpretations of tall:

$$tall_1 = \lambda x.height(x) \geq 150cm$$

$$tall_2 = \lambda x.height(x) \geq 160cm$$

$$tall_3 = \lambda x.height(x) \geq 170cm$$

$$tall_4 = \lambda x.height(x) \geq 180cm$$

$$tall_5 = \lambda x.height(x) \geq 190cm$$

Then μ is a function that assigns probabilities to world, language pairs. Below, (made-up) values have been chosen that reflect that we take the most probable threshold for *tall* to be fairly great in height, but not too great.

$$\mu = \{\langle\langle w_1, tall_1\rangle, 0.05\rangle, \langle\langle w_1, tall_2\rangle, 0.15\rangle, \langle\langle w_1, tall_3\rangle, 0.3\rangle, \langle\langle w_1, tall_4\rangle, 0.4\rangle, \langle\langle w_1, tall_5\rangle, 0.1\rangle\}$$

From this we can calculate the probabilities of w_1 being actual, and of each sharp meaning of 'tall' being the one in play (which, in this case, are both pretty trivial given that we have only one possible world):

$$\mu(w_1) = 1$$

$$\mu(tall_1) = 0.05,\ \mu(tall_2) = 0.15,\ \mu(tall_3) = 0.3,\ \mu(tall_4) = 0.4,\ \mu(tall_5) = 0.1$$

The probability that 'Jess is tall' is true, is then calculated as the sum of the probabilities of the world, language pairs in which it is true that $tall_n(jess)$, weighted against the probability that the world in the pair is the actual world. There is only one world language pair in which it is false that Jess, at 188cm in height, is tall, namely $\langle w_1, tall_5\rangle$. This gives:

$$\begin{aligned} prob(\text{Jess is tall}) &= 0.05 \times 1 + 0.15 \times 1 + 0.3 \times 1 + 0.4 \times 1 \\ &= 0.9 \end{aligned}$$

This is not to say that there is no further role for Bayesian update on Lassiter's semantics. An agent can receive new information, either about the world (via perception, being told something, etc.), or about her language (via the way some expression is used). This information can then be used to update the distribution over world-language pairs and so affect the probability value she calculates for further utterances.

The contents of vague utterance types are not, on Lassiter's view, best understood as truth conditions (or propositions characterised by truth-conditions), but in the spirit of dynamic semantics, as updates on belief spaces/contexts. However, truth-conditional content is nonetheless present in Lassiter's account. If one has a probability distribution over the sharpened understandings of a vague natural language expression that could be in play, then there must be some one in play. If it were genuinely indeterminate which understanding of *tall* is in play, then it would not make sense to talk of the probability of any one understanding in play. To see this, consider a dice roll. If there is an equal one-sixth chance of getting any one of the values of the dice, then there must, on any roll, be some value that turns up. If we expand the possibility space to allow for the dice to land perfectly balanced on one edge (and so indeterminate between results), it would be sensible to talk of the probability of getting an indeterminate result within the model. However, if the results space is simply $\{1, 2, 3, 4, 5, 6\}$, then employing

a probability distribution over this space in one's model ties one to the commitment that any roll will result in one of these values. Similarly, if we have a probability distribution over what possible language is in play, then for any utterance, there is a possible language in play. Furthermore, if one wanted to try to incorporate genuine indeterminacy into Lassiter's account, unlike in the dice roll case where the dice could land on its edge, it would be unclear what to incorporate into the probability space to do so. However, Lassiter does not intend his account to allow for indeterminacy: if there is a probability value for the truth of some utterance, then that utterance is true or false *simpliciter*. Indeed, a motivation for PLK is that one can replace the need for indeterminacy in semantics via the incorporation of uncertainty.

PLK is therefore not the same as the governing view (on which we have exact knowledge of rules that completely determine correct uses of expressions). Instead, on PLK, competent speakers are only required to have inexact knowledge of rules that nonetheless determine correct uses of expressions. Semantic competency amounts to tracking truth conditions, but allows for uncertainty in what these truth conditions are. Our linguistic knowledge is probabilistic, the linguistic facts are determined relative to a context. To this extent, we are assumed to have inexact knowledge of the linguistic facts.

In §4, it will be important to consider assertion conditions as well as the probability of an utterance being true. Lassiter suggests that the assertibility of an assertoric utterance could be connected to whether the probability of the truth of that utterance exceeds some threshold. This suggestion is tentative, however, given that Lassiter presents it as an option on the assumption that knowledge is the norm for assertion. A full discussion of this suggestion will be made in §4. The incorporation of assertibility along these lines is fairly straightforward. Some, possibly context sensitive, threshold can be set such that an utterance which has a value below the threshold is not assertible. Given that the focus for this paper will be on rules for (correct) language use, I will take this tentative suggestion of Lassiter's to be one alternative up for evaluation.[14]

3.3 The PJC Approach

The example used to illustrate PJC will be Sutton's situation theoretic semantics for vague adjectives (Sutton, 2015), rather than Hampton's prototype theoretical approach (Hampton, 2007), however this should not be taken as a loaded choice. It should be noted that more comprehensive probabilistic situation theoretic semantics is provided in (Cooper et al., 2014, 2015), however, since Sutton (2013, 2015) explicitly models vagueness, the formal details given below will be based on (Sutton, 2015).

The PJC approach as applied to vagueness in (Sutton, 2013, 2015) has its formal foundations in Situation Theory (Barwise & Perry, 1983) (see Devlin (2006) for a compact overview), and Austinian accounts of truth (Austin, 1950/1979). Situations are formal representations of small concrete parts of the world. These can abstracted over in terms of common features into situation types. Situation types are therefore inherently more fine grained than full-blown possible worlds, however, it will not be a goal of this paper to decisively argue for either situation theoretic or possible worlds theoretic semantics. The advantages of situations and events is that they are more fine

grained than possible worlds and so are arguably more psychologically plausible entities than full blown worlds (see Cooper et al. (2015) for discussion). Nonetheless, possible worlds remain the dominant paradigm in formal semantics.[15]

The basic idea in situation theoretic approaches is that utterances are associated with situation types.[16] Austinian truth is a way of cashing out a correspondence theory for truth. Utterances refer to situations, sentences/utterance types are associated with situation types. An utterance is true if the situation it describes is of a type with the relevant situation type.

Just as Lassiter argues that dynamic, possible worlds models require enriching with gradience, there is room to incorporate gradience into situation theory too. Both Cooper et al. (2015) and Sutton (2015) use probability calculus to model this. Cooper et al. develop a probabilistic version of Type Theory with Records (TTR) (Cooper, 2012). A basic concept in TTR is a type judgement (roughly the judgement that some situation/object/event is of a particular type). This rich type theoretic approach allows that probabilities be assigned to type judgements. Sutton (2013, 2015), retains a simpler type theory and remains closer to Situation Semantics by describing the probability of a situation being of a described situation type, given a situation of an utterance situation type.[17]

Sutton's version of PJC adopts the distinction made in Situation Semantics between a *Discourse Situation Type* and a *Described Situation Type*. Formally, these are represented as abstractions over situations, where the content of the situation is given by an *infon*. For example, (6) represents the type of described situation in which John (j), is 180cm in height, and (7) represents the type of discourse situation in which John is described as tall by some agent.

$$\lambda[\dot{s}](\dot{s} \vDash \langle\!\langle \texttt{height=180cm}, \texttt{j}, yes \rangle\!\rangle) \tag{6}$$

$$\lambda[\dot{d}](\dot{d} \vDash \langle\!\langle \texttt{utters}, \dot{a}, \text{TALL}, \texttt{j}, yes \rangle\!\rangle) \tag{7}$$

In (6) and (7), $\dot{s}$, $\dot{d}$, and $\dot{a}$ are parameters (roughly the situation theoretic versions of variables), for *described situations*, *discourse situations*, and *agents* respectively. The $\vDash$ symbol is the type relation such that $s \vDash \sigma$ iff situation s is of type (supports the infon) σ. Infons ($\langle\!\langle ... \rangle\!\rangle$) contain a relation, arguments and a polarity which states whether the relation holds.

As in (8), the information carried by 'John is tall' is represented as a probability distribution over heights, h, where $k \in [0, 1]$. The conditional probability distribution represented is that of the probability of a situation being of a type in which John is some height, given that he has been described as tall (where values vary with different heights h).

$$p(\lambda[\dot{s}](\dot{s} \vDash \langle\!\langle \texttt{height=h}, \texttt{j}, yes \rangle\!\rangle) \mid \lambda[\dot{d}](\dot{d} \vDash \langle\!\langle \texttt{utters}, \dot{a}, \text{TALL}, \texttt{j}, yes \rangle\!\rangle)) = k \tag{8}$$

Along with the relevant priors (the values for which can be derived from a learning model as in (Cooper et al., 2015)), one can then apply Bayes' rule to calculate:

$$p(\lambda[\dot{d}](\dot{d} \vDash \langle\!\langle \texttt{utters}, \dot{a}, \text{TALL}, \texttt{j}, yes \rangle\!\rangle \mid \lambda[\dot{s}](\dot{s} \vDash \langle\!\langle \texttt{height=h}, \texttt{j}, yes \rangle\!\rangle))) = k \tag{9}$$

For some value of h, what (9) represents is the probability of a discourse situation type in which John is described as tall, given he is some height. How this may be interpreted is as the probability that an idealised competent agent would judge John to be tall, given that John was that height. It is important to make a distinction here, however. The probability that an agent would judge something/someone to be of the type $tall$, may, of course, be influenced by all sorts of factors extraneous to the agents competency. However, it is legitimate to invoke something along the lines of a competence/performance distinction. We can ask, all else being equal, whether a competent speaker, would, for example, more probably judge someone of a certain height to be tall or not, and this matter can be divorced, in principle, even if not in practice from other performance related factors. Thereby we can separate, on a theoretical level, semantic metalinguistic uncertainty (uncertainty about what words may be used to accurately describe a situation), from other extraneous factors which may impact correlations between discourse situation types and types of described situations.

The difference between non-metalinguistic and metalinguistic uncertainty with respect to 'John is tall' can then be linked to (8) and (9). The distribution in (8) represents non-metalinguistic uncertainty. From the discourse situation alone (so lacking further information), an agent may be uncertain what John's height is. However, in another situation, even if John's height is known, metalinguistic uncertainty may remain as to whether to describe John as 'tall' (or 'not tall'). This will be represented as the conditional probability distribution represented in (9) where certainty about describing John as 'tall' varies with values for height h (how uncertain an agent should be of describing John as tall, given that he is of some height). Better still, to use the gloss from van Eijck & Lappin (2012), this can be understood as the probability that an idealised agent would make a judgement with this predicate.

PJC thus differs from both degree interpretations of the governing view (on which rules determine correct uses of expressions), and from PLK (on which rules determine correct uses of expressions, but that our knowledge of them is inexact). On PJC, the rules one learns are themselves probabilistic (or, at least, represented probabilistically). In this sense, on PJC, in principle, we may have exact knowledge of inexact rules, namely, how probable the actions of some idealised competent speaker would be. Semantic competency amounts, not to tracking truth conditions (as in PLK and the governing view), but to tracking probabilistic dependencies that hold between uses of expressions/judgements, and properties that the things described have. Thus articulated, PJC can be understood as a model for the rule-as-reasons approach by way of a bridging principle (B) linking this metalinguistic uncertainty.

(B) There is a proportional relation between the probability that a competent agent would judge that ϕ and the strength of reasons an agent has to make a judgement that ϕ.

A basic justification for (B) may be grounded in semantic learning. The probability values an agent assigns to a judgement are grounded in how she has witnessed predicates being used. These abstractions over learning data provide the vast majority of the data an agent has for making her own judgements. There may be many ways to encode a mapping from probabilities to degrees of (semantic) reason for a linguistic action. For simplicity, we may assume a strength of reason space $\mathbf{R} = [0, 1]$ and a one-to-one mapping between values in this space and probability values.[18] Probability values encode, or can provide the basis for deriving, strengths of semantic reasons.

On the simplifying assumption, probability values directly encode strengths of semantic reasons. For example, for a judgement that ϕ, 1 marks as complete/full semantic reason to judge that ϕ, 0 as no reason to judge that ϕ, and 0.5 as a middling semantic reason to judge that ϕ. For metalinguistic uncertainty, probability values can thus be understood e.g., to represent the strength of reasons one has for describing John as tall (as opposed to not tall), given that he is of some height.

To follow is a highly simplified example for how Sutton's version of PJC works for 'tall'. Instead of world-language pairs, we will have a conditional probability distribution for the height of some individual, given that they have been described as tall.[19] This distribution, in PJC, models a generalisation an agent has made during their semantic learning process, based upon correlations that exist between predications of 'tall' and people's heights. An example is given in Table 1.

Table 1: Example probability distribution for heights, given predications of *tall*

height in cm (h)	$h < 175cm$	$175 \leq h < 185$	$h \geq 185$
$p(\lambda[\dot{s}](\dot{s} \vDash \langle\!\langle \texttt{height=h}, \texttt{j}, yes \rangle\!\rangle) \mid \lambda[\dot{d}](\dot{d} \vDash \langle\!\langle \texttt{utters}, \dot{a}, \text{TALL}, \texttt{j}, yes \rangle\!\rangle))$	0.1	0.4	0.5

These made-up values reflect high expectations for an individual being above average height, given a description of them as tall (in my country of residence, the average height for men is around 180cm). Expectations are a bit lower for values around the average height and are much lower for values below the average.

We also need to provide priors for height values which have been chosen to reflect a preference for heights around the average value. An example is given in Table 2.

Table 2: Example prior probability distribution for heights

height in cm (h)	$h < 175cm$	$175 \leq h < 185$	$h \geq 185$
$p(\lambda[\dot{s}](\dot{s} \vDash \langle\!\langle \texttt{height=h}, \texttt{j}, yes \rangle\!\rangle))$	0.25	0.5	0.25

Finally, we have a prior for an utterance situation being one in which 'tall' is used. This is tricky. It is doubtful that such values should play an important semantic role. I therefore adopt an ambivalent value of 0.5 ($p(\lambda[\dot{d}](\dot{d} \vDash \langle\!\langle \texttt{utters}, \dot{a}, \text{TALL}, \texttt{j}, yes \rangle\!\rangle)))) = 0.5$), making the prior for 'tall' equal the prior for 'not tall'.[20] To derive the probability that an idealised agent would apply 'tall' to individuals in the above height ranges is then simply a matter of applying Bayes' Theorem, which gives us the values shown in Table 3:

Table 3: Example posterior probability distribution for predications of *tall*, given heights

height in cm (h)	$h < 175cm$	$175 \leq h < 185$	$h \geq 185$
$p(\lambda[\dot{d}](\dot{d} \vDash \langle\!\langle \texttt{utters}, \dot{a}, \text{TALL}, \texttt{j}, yes \rangle\!\rangle)) \mid \lambda[\dot{s}](\dot{s} \vDash \langle\!\langle \texttt{height=h}, \texttt{j}, yes \rangle\!\rangle)$	0.2	0.4	1

So one has very little reason to call someone (in this case, a man my country of residence) below the average 'tall', little reason to call someone around the average 'tall', and excellent reason to call someone well above the average height 'tall'.

Since reasons will play a part in the final proposal given here, it is worth flagging the issue of how reasons

can be conceived of relative to (agent) internal and external criteria. The sense of 'reason' here is internal in the following sense: how much (semantic) reason there is for making a judgment in a given case turns on the beliefs of the agent, and the generalisations she has drawn from her semantic learning data. Values in the range [0,1] represent the extent to which the learner's data and beliefs support the hypothesis that some predicate is the correct one to apply in a given case. However, in the literature on philosophy of action, intuitions are split between those who think that reasons for action can be internally characterised, and those who externalise them (see Alvarez (2010) and references therein). As a simple example, suppose someone has a false belief that it is raining. On an internal conception of reasons for action, provided that this belief stands in the correct causal relationship with an appropriate desire, then the agent has a reason for, say, taking an umbrella with her when she goes out. However, on an externalist conception of reason, reasons are facts. (The content of) a false belief cannot be a fact, and so the agent has no reason for taking an umbrella (she just thinks that she does).

I do not wish to enter this debate on reasons, but a role in a theory of communication for both an internally, and an externally characterisable account of reasons does not seem implausible. The former surely plays some role in explaining the responses and behaviours of agents. The latter surely plays some role in explaining whether or not those actions were, in fact, correct/justified viewed from the perspective of agents as individuals in a complex and intertwined social network of actions and responses to each other's actions. Indeed, below, I shall argue that the relocation problem of vagueness can be eased if we include, in our theory, reasons for e.g. assertion that are internally accessible to an agent and reasons for assertion that are not necessarily internally accessible. In short, I will argue that normatively loaded sharp boundaries are not acceptable for internally accessible reasons for assertion, but can be accepted for external or non-internally accessible reasons for assertion. Easing the relocation problem will mean relocating sharp boundaries to a place where we can accept them as being without infringing upon the semantic competence of agents.

4 Revisiting the Problem for Degrees

As described in §1.2, Wright rightly points out that a switch to degrees of truth means that it is no longer the case that small changes in properties/features of objects have no effect on the justification one has for applying a predicate. With an all or nothing account of justification, our introspection in line with (T2) leads us to realise that there will always be a small enough change in an object such that the change cannot justify an expression ceasing (starting) to be applied where it wasn't (was) prior to the change. With degrees of justification (such as degrees of truth), provided one has a fine enough grain of degrees, all small changes in objects do (or at least may) affect the justice with which a predicate is applied. For any small change in an object, there can be a small change in justification. In other words, for a degree of truth view, we replace one tolerance principle (2) with another (4) (both reprinted from above):

$$\text{For all } a, b, \text{ if } \mathcal{I}(P)(a) = 1 \text{ and } a\ \mathcal{I}(\sim_P)\ b, \text{then } \mathcal{I}(P)(b) = 1 \tag{2}$$

$$\text{For all } a, b, \text{ if } \mathcal{I}(P)(a) = n \text{ and } a\ \mathcal{I}(\sim_P)\ b, \text{ then } \mathcal{I}(P)(b) = n \pm \epsilon \tag{4}$$

For Wright, however, this amounts to moving around bumps in the carpet (Wright, 1975, pp. 349-50), since the question can always be posed: what degree of justification is, on balance, sufficient for an application of a predicate? If there is such a level, then one of the following will hold, neither of which is acceptable:

(a) There is some a, b such that if $P(a)$ is correctly assertable and $a \sim_P b$, then $P(b)$ is *not* correctly assertable. In this case the tolerance of the predicate is lost (we must abandon (T2)).

(b) For all a, b, if $P(a)$ is correctly assertable and $a \sim_P b$, then $P(b)$ is correctly assertable. In this case we are no further forward than we were before, since our rule will give rise to incoherence, given that our rules are meant to determine correct usage.

What arguably gets Wright's argument going against the multiple-valued variant of the governing view is that it does not replace "the old connection between justified assertion and truth". Put another way, this variant of the governing view runs together two distinct claims, namely:

(GVi) Small changes in objects may affect the degree of truth of applying a predicate.

(GVii) Degrees of truth determine whether the predicate is to be correctly applied.

The governing view does not, of course, have to be committed to the claim that *all* changes in justification/degrees of truth make for changes in truth value. However, it is, by the lights of (T1), committed to the view that there are some rules, albeit possibly ones which can be put in terms of degrees, which determine when an expression should be applied. Put another way, changes in objects may affect (degrees of) justification (understood here as degrees of truth), and some level of degrees of truth determine a non-graded notion: whether a predicate would or would not be correctly applied/correctly assertable of an object. Wright's argument exploits this determination commitment. After all, if rules put in terms of degrees determine how to apply an expression, then there must be some threshold of degrees above which one should apply the expression and below which one should not.[21]

An obvious move is precisely to try to break "the old connection between justified assertion and (degrees of) truth", and so doing, pull apart (GVi) and (GVii). In §4.1, we will see that both PLK and PJC do precisely this. However, in §4.1.1 and §4.1.2, I argue that PLK and PJC still generate a relocation problem: even if 'correctness' of application is not understood in terms of truth, there is some notion of correctness which sharp boundaries are relocated to.

In §4.1.3, however, I will suggest a way that a rules-as-reasons approach such as PJC has further room for manoeuvre. I will suggest that such probabilistic approaches can be related to the internalist and externalist conceptions of reasons mentioned in §3.3. Crucially, having reason for an assertion in an internalist sense only imperfectly tracks having reason for an assertion in an externalist sense. Semantic competency can be defined in an internalist sense. However, this does not make us immune from error (acting without reason) in an externalist

sense: Our internally defined competency makes us externally error prone. As we shall see, far from abandoning (T2), we need internal reflection to operate as communicative agents, but contra (T1), there is more to say about justified/correct application of predicates than these internal considerations can provide.

4.1 Evaluating Probabilistic Approaches to Vagueness

PLK does not appeal to degrees of truth. On Lassiter's version of PLK, if one knows that Jess is 188cm in height, then the probability that 'Jess is tall' is true is the sum of the probabilities of the world-language pairs in which it is true that Jess is 188cm in height (weighted by the probabilities of the world in each pair). However, since there is some standard for height in play for each context, it is implied by PLK that there is a (possibly unknown) *de facto* threshold for truth of an utterance.

On the probabilistic rules-as-reasons approach, PJC, probability values are understood as the probability that a competent speaker would apply a predicate (given a situation of a certain type). This does not imply that there is a threshold for the truth of an utterance. In this sense, PJC is not committed to there being, in a context, some set of conditions that sharply divides true instances of applications of predicates from false ones, and so does not assume that we are tracking (albeit inexactly) those conditions in that context.

Nonetheless, there are similarities between PLK and PJC. For both, instead of (GVi) and (GVii), we have (PRi), but not (PRii):

(PRi) Small changes in objects may affect the agent's subjective probability that the predicate applies (for PLK, that it truly applies).

(PRii) Probability values determine whether the predicate is correctly (=truly) applied.

Even in the limit cases, probability values in (PRi) do *not* determine whether or not a predicate applies since they are merely subjective estimations. As such, any probability value for ϕ is consistent with ϕ being true or false. Thus, PLK and PJC break "the connection between justified assertion and truth" in the following senses. On PLK, justification for applying a predicate comes via inexact linguistic knowledge (probabilistic estimations of truth conditions). Agents are tracking truth-conditions, but only inexactly, via subjective probabilities over a space of possible truth conditions. So it is possible that an agent can be justified in applying a predicate (has a high subjective probability that it truly applies), even when it does not truly apply. On PJC, the probability values (that is, the values on a scale of strengths of reasons) do not determine the truth value of the predication. Varying levels of semantic reasons, are compatible with either the truth or falsity of an assertion. If the probability of an idealised agent applying a predicate to an object is 0.99 (as opposed to its negation), then, all things being equal, one has excellent reason to apply that predicate. However, having an excellent reason for applying a predicate in a given case is wholly consistent with that predicate being true, false, or indeterminate of that object.

With respect to tolerance, PLK and PJC can accept a principle which bears obvious similarities in form to the

principles of the degree based governing view:[22]

$$\text{For all } a, b, \text{ if } p(P(a)) = n \text{ and } a \sim_P b, \text{ then } p(P(b)) = n \pm \epsilon \tag{10}$$

This weaker notion of tolerance is equivalent (mutatis mutandis with respect to conceptions of probability/degrees of truth etc.) to other proposals. For example, is is roughly equivalent to the "descriptive tolerance principle" in Égré (2009, 2011) and the closeness principle in Smith (2008). However, the latter is defined in terms of degrees of truth, not probabilities. No incoherence threatens on tolerance principles such as (10), since the probabilities in question do not themselves determine the truth conditions of $P(a), P(b)$ etc.

PLK could be interpreted as having an epistemicist flavour. There are assumed to be bivalent truth values for propositions, determined by whatever standard/threshold/sharp proposition is being expressed. Communicative agents merely imperfectly track these thresholds and so could, in principle, via their inexact knowledge, unknowingly move from truly asserting $P(a)$ to falsely asserting $P(b)$, even though they assign very similar probability values to both propositions. However, competence on an inexact knowledge view is not stated in terms of tracking truth conditions and behaving accordingly, but is stated in terms of tracking truth conditions closely enough.

On PJC, agents are assumed to be tracking the use of vague predicates by competent speakers. There is no commitment to there being some precise interpretation of a vague expression being expressed in every context. On PJC, one, more or less, precisely learns imprecise representations. As such, PJC can remain uncommitted to whether there are sharp thresholds between veridical uses of a predicate and non-veridical ones as a matter of one's semantic representations.

However, we must still consider whether the relocation problem arises for probabilistic approaches to vagueness. Even though subjective probabilities do not determine correct (=true) applications of a predicate, is there some other equally troubling conception of correctness that is determined by subjective probability estimations? Namely, we must test (PRiii):

(PRiii) Probability values determine whether one is justified to assert that a predicate applies.

4.1.1 Thresholds for assertion

On the assumption that knowledge is the norm for assertion, Lassiter (2011) tentatively assumes there to be a context-sensitive threshold for assertion in play. We might think, then, that any threshold for assertion must be at least as high as the *de facto* threshold for truth in that context. However, Lassiter is more interested in communication than in the *de facto* thresholds that his model implies:

> "If [the supposition that knowledge is the norm of assertion] is correct, then we should expect, as a *descriptive matter*, that a cooperative speaker A will typically assert things that *she thinks she knows*. On standard assumptions, A's subjective probability for a proposition p will rarely if ever reach 1; but p may have high enough probability that A thinks that she may profitably make an assertion calculated to communicate the information that p. How

high is judged 'high enough' will depend on various features of the context, such as the conversational stakes and perhaps even aspects of the speaker's personality. I will use α as a placeholder for the threshold of assertibility, however this is determined in particular contexts. Cooperative speakers will assert a proposition p only if the probability of p is greater than α, and the information that p is deemed relevant, useful, etc." (Lassiter, 2011, p. 142) (my emphasis)

If a threshold for assertion is only set because a speaker 'thinks she knows', or thinks that she has enough justification, then we cannot, from the value of the assertion threshold, infer whether or not all assertions of p, when the probability of p is above the threshold, are true. Furthermore, Lassiter commits to this being only a descriptive matter. As such, we cannot say that asserting p when the probability of p is above the threshold makes the assertion of p correct or justified, only that the agent thinks she is justified in making the assertion.

There is, however, a worry about the introduction of probabilistic thresholds, even at a descriptive level, when characterised in the above way. Assume $\mathcal{BK}_A(\phi)$ indicates that agent A thinks she knows that ϕ. For any threshold α, there will be some minute value ϵ such that for some a, b:

$$|p(F(a)) - p(F(b))| = \epsilon \wedge a \sim_F b \wedge \mathcal{BK}_A(F(a)) \wedge \neg\mathcal{BK}_A(F(b)) \tag{11}$$

Whether one thinks that one knows something is a matter that can be assessed, presumably, via internal reflection. Vague predicates would also, presumably, support the following tolerance principle. For all a, b:

$$(\mathcal{BK}_A(F(a)) \wedge a \sim_F b) \rightarrow \mathcal{BK}_A(F(b)) \tag{12}$$

Therefore, if thresholds are defined in terms of whether one thinks one knows something, then it looks as though we have counterexamples to a higher-order tolerance principle. The relocation problem has struck again.

Nonetheless, it is hard to see how, on a descriptive level, we can do without thresholds of some form, and still account for the fact that people do make assertions in some conditions, but not others. The relocation problem at this level is plausibly be addressed in terms of arbitrary mechanism (Raffman, 2014). There is, descriptively speaking, a point at which agents switch from asserting something is F to not doing so. There is, therefore, a threshold. For probabilistic accounts, this threshold can be described very naturally in terms of probability values. What a probabilistic account should not do, however, is assume that the value of this threshold can be rationally, introspectively accessed. Nonetheless, we may still ask what the connection between these probability values and correct/justified assertion is. I return to this in §4.1.3.

4.1.2 Degrees of justification of assertions

Wright points out that, on a degree view:

"[degree views do not support] the following principle: if b is marginally less F than a, then if the less misleading description of a is 'F', the less misleading description of b is 'F'. Yet if this principle is false, there must, in any

> Sorites type series, be a last case of which we are prepared to say that if we *had* to describe it as either F or as not-F, the better description would be 'F'." (Wright, 1975, p. 349)

Yet, worryingly, probabilistic approaches to vagueness do not, arguably, support the above principle either. Take, for example, PJC (and continue to suppose that the probability values assigned to ϕ on PJC map directly to a scale marking the strength of reasons one has for an assertion). If $p(F(a)) = 0$, there is no reason to assert $F(a)$; if $p(F(a)) = 1$, there is full reason to assert $F(a)$; and if $p(F(a)) = 0.5$, there is middling reason to assert $F(a)$. Suppose that $a \sim_F b$, $p(F(a)) = 0.51$, and $p(F(b)) = 0.49$.[23] Now, unless we say that one can have more/a better reason to assert $F(b)$ than $F(a)$, we seem to be committed to (Diff):

(Diff) If $p(F(a)) > p(F(b))$, then $F(a)$ is a better description of the relevant situation than $F(b)$

Yet this looks like the introduction of a threshold, not for correct assertion, but for the last case in which one assertion can be better than another. Namely, there must now be a last case of which we are prepared to say that if we *had* to describe it as either F or as not-F, the better description would be 'F'. In other words, there is some ϵ such that $F(a)$ is a better description than $\neg F(a)$ even though $|p(F(a)) - p(\neg F(a))| = \epsilon$.

However, this is not as damaging as a threshold for *correct* assertion. On any scale of reasons there will be a tipping point. Depending on the granularity that an agent can track the strengths of reasons they have for an assertion, there will be a point at which some reasons outweigh others even if only marginally. This does not imply that having more reason to make one assertion than another determines the correctness of either assertion. One might have more reason to apply *tall* to someone of average height than to someone of short stature, but this does not imply that it would be correct to assert that either individual is *tall*.

4.1.3 Rules as reasons: Internalist and externalist reasons

Even if we adopt a Raffman-style, mechanistic interpretation of probabilistic thresholds for assertion on a descriptive level, the question remains, for probabilistic views, what should be said about when one is justified in making an assertion.

It is helpful, at this point, to return to the distinction from §3.3 between reasons conceived of in an internalist way and reasons conceived of in an externalist way. Tolerance principles, recall, are meant to be discoverable from behavioural evidence and an internal reflection on what justifies the application of an expression in a given context. The rules-as-reasons adjustment of (T1), namely, that the allowable/correct use of an expression is guided but not determined by such principles, in fact prevents a probabilistic rules-as-reasons theorist from committing to a probabilistic threshold that determines correct assertion. Probability values represent strengths of reasons for making assertions, and agents can, with some granularity, assess whether some reasons are better than others. These strengths of reasons are what are available via internal reflection (grounded, for example, in semantic learning). What A PJC theorist must reject is that strengths of reasons *determine* when a particular action is correct.

If the reasons accessible from internal reflection do not determine conditions on correct assertion, then one route for a PJC theorist is to claim that the determiners of such factors are externally characterised. It is not in the scope of this paper to provide an account of how such conditions are determined, but externally characterised conditions on assertion could be governed, for example, by the beliefs and presuppositions of others and how they intersect with those of the agent, conventions, and complex social and/or power relationships between agents. If $\mathcal{A}(\phi)$ indicates that ϕ is correctly assertible where *correctness* is determined by some conditions that are not necessarily retrievable via behavioural evidence and internal reflection, this means that the following may hold on PJC. For some small value ϵ, and for some a, b:

$$|p(F(a)) - p(F(b))| = \epsilon \wedge a \sim_F b \wedge \mathcal{A}(F(a)) \wedge \neg\mathcal{A}(F(b)) \tag{13}$$

However, this does not make agents incoherent. What (13) states is that there may be cases where small differences in internally accessible reasons for asserting $F(a)$ and $F(b)$ can correlate with large differences in whether one can correctly or successfully pull-off some particular linguistic action such as the assertion of $F(a)$ and $F(b)$ (plus, presumably, some intended perlocutionary effect). Critically, however, it cannot be a condition on the semantic (or pragmatic) competence of the agent that she is aware of that which she cannot be expected to know via inferences based on behavioural evidence and internal reflection on language usage.

Since (13) holds on PJC, we do have sharp boundaries. The relocation problem persists. Nonetheless, the move made available by a rules-as-reasons approach is that these sharp boundaries have been relocated to a place outside of the semantic competency of an agent. Agents can make decisions to act based on the strengths of their internally characterised reasons and still fail in their communicative endeavours due to externally characterised reasons (such as the beliefs and presuppositions of others, conventions, and complex social and/or power relationships) that competent agents may not reasonably be expected to track or, indeed control.[24]

It might, however, be a part of an agent's competency that she does not sail too close to the wind when balancing reasons that guide her assertions and other linguistic actions. To assert $F(a)$ when one has only marginally better reason for doing so than for asserting $\neg F(a)$ is not good practice when one's reasons for actions are mere guides or heuristics for when an action would be correct. To do so would be to invite situations of a type characterised by (13), and, most likely, miscommunication.

The question is, does the condition that an agent avoids judgements on the basis of *marginally* better reasons constitute an instance of the relocation problem? When it comes to weighing up *internal* reasons for linguistic action, how marginal is too marginal? A promising option for this internal version of a tolerance challenge would be to adopt a view based on semi-principled or arbitrary decisions (Rayo, 2008; Raffman, 2014). Furthermore, a combination of Bayesian meaning representations combined with a rules-as-reasons approach may allow for a more nuanced account along these lines. Our introspective access to the strengths of reasons we have for performing linguistic action is, at best, fairly coarse-grained. That is to say, we cannot always know whether one reason is

stronger than another in borderline cases of whether to assert F or not-F. Furthermore, the reason we find sorites series so compelling, is, arguably, because we are cognitively set up to ignore small amounts of uncertainty (I suggest in Sutton (2013) that this derives from a requirement for cognitive efficiency). This coarseness of grain could be modelled via modulation of the bridging principle from Section 3.3 whereby probabilities would map to (ranges of) values on a scale representing strengths of reasons in line with a probability distribution. (This would have the effect of introducing some indeterminacy into the system such that a probability value would not always map to the same value on the scale of strengths of reasons.) Strengths of reasons are what competent speakers must weigh up in performing linguistic actions. I leave the details of this for further work, but, in simple terms, one could find an appropriate stochastic mapping between probability values and strengths of reasons such that decisions would be, at least in part, principled (one would have e.g. more reason to assert $F(a)$ than not-$F(a)$), but the stochastic nature of this mapping relation would predict that the switching point between an agent judging $F(a)$ or not-$F(a)$ could change across situations this making the switching point appear arbitrary.

4.2 Comparison with other proposals

A question that immediately arises is whether an inexact knowledge approach such as PLK could make the same move. In principle, the answer is yes. As long as justification for assertion is not determined by the value of an agent's probabilistic judgement. This would, however, amount to a rules-as-reasons interpretation of PLK with respect to justified assertion.

Returning to Raffman (2014) and Rayo (2008), rules-as-reasons accounts seem to fit well with the idea that tolerance is "a feature of our ability to use linguistic representations" and to "make semi-principled decisions". As we saw above, an account based on arbitrariness fits well here too. Acting on our internally accessible reasons, we should, at some point, cease to make positive judgements when approaching borderline areas, but exactly where should be arbitrary/only semi-principled. However, these arbitrary stopping points are descriptive and not-normative. If we wish to discuss normative questions such as, *was that assertion justified*, we must, in order to relocate boundaries to a non-threatening area, distinguish internally accessible reasons for action from reasons for justification that need not be internally accessible.

The case of permissible assertion is considered in Égré, albeit in terms of *obligation* and its dual *permissibility*.[25] On Égré's account, obligation, $\mathcal{O}$, to judge ϕ has a probability threshold, as does permissibility, $\mathcal{P}$, to judge ϕ. For example, suppose that the permissibility threshold is defined $\mathcal{P}(P(x)) \leftrightarrow p(P(x)) > 0$ and the obligation threshold is $\mathcal{O}(P(x)) \leftrightarrow p(P(x)) \geq 1$ (i.e. equals 1). A relocation problem arises for Égré's (2009; 2011) account, except that now we have two sharp boundaries. Between $p(P(x)) = 0$ and $p(P(x)) > 0$ one goes from it not being permissible to judge that x is P to it being permissible to judge that x is P. Also between $p(P(x)) = 1$ and $p(P(x)) < 1$ we jump from it being obligatory to judge that x is P to it being not obligatory to judge that x is P. This means that the following applies to the account in Égré (2009, 2011) regardless of the precise thresholds one

sets.

$$P(a) \text{ is permissible/obligatory and } a \sim_P b, \text{and } P(b) \text{ is } \textit{not} \text{ permissible/obligatory} \tag{14}$$

A response is made by Égré (2011) to this jarring result, namely, that "we could imagine that one ought to judge something P or not P provided it is *sufficiently* unambiguously P or *sufficiently* unambiguously not P." (Égré, 2011, p. 79). However, if *sufficiently unambiguous* is interpreted precisely, we will not ease the sharp boundaries problem. If *sufficiently unambiguous* is not interpreted precisely, this amounts to a proposal to make the thresholds for $\mathcal{O}$ and $\mathcal{P}$ relative to a predicate vague. At the very least, in this latter case, we should expect a theory of the vagueness of *these* concepts that does not deliver sharp boundaries at the next order up. As Égré argues, (p. 85), there may be benefits to having two sharp boundaries over one. Nonetheless, as he goes on to note, an account of higher-order vagueness is required to give a fuller account of this problem.

What we need to do, is to relocate sharp boundaries to a place where we should expect them to be and in which we can accept them as being. On Égré's approach, this would be to make explicit that the internally accessible reasons for action, such as the probabilities agents track for when competent speakers apply predicates, cannot fully determine the deontic forces for assertion which I call "correctness" and what Égré calls "obligation", but can only act as a heuristic. One would furthermore need to argue that standards of obligation (or permissibility) should not be expected to be (fully) accessible for a speaker to be competent. To do this would be to explicitly adopt a rules-as-reasons approach.

Finally, I will consider the proposal made by Smith (2008). Smith's framework is not probabilistic. Instead, he defends fuzzy plurivaluationism. However, his response to the relocation problem of vagueness is worth considering, since it could be adopted by a probabilistic framework. Smith (2008, §5.4) distinguishes between a degree of confidence in asserting ϕ and the acceptability of asserting ϕ. Degrees of confidence are, like my internally accessible reasons for assertion, graded (albeit linked to degrees of truth). Acceptability for assertion is not graded (Smith, 2008, p. 249). For example, if the degree of confidence is asserting ϕ is 0.5, then "This does not mean that an (ordinary, unhesitant) assertion of [ϕ] would be 'half right' [...]; it is pass/fail. Rather it means that an assertion of [ϕ] with a degree of confidence 0.5 would be right [...]." (Smith, 2008, p. 250)

However, Smith's position prompts the asking of Wright's question: What is the smallest degree of confidence for assertion such that, for any lower degree, assertion would not be, on balance, justified? (I.e. What is the cut-off point for acceptable assertion?) Smith's (2008) answer is that we have more than the two options of outright asserting ϕ or not-ϕ "We can apply the predicate with varying degrees of confidence or hesitation." (Smith, 2008, p. 251). However, this does not yield an adequate response to the problem, since now we can ask what the lowest degree of confidence is such that an outright bare assertion of ϕ (with no hedging) is acceptable. If this is n, then any minuscule degree lower than n will result in a sharp jump between acceptability and non-acceptability of such a bare assertion, which, as Smith (2008) stresses, is pass/fail. In other words, Smith's account is in a similar predicament to the probabilistic views in section 4.1.2. What we still need is to distinguish internal and external reasons and to see the former as a heuristic guide for the latter. This relocates the sharp cut-off to a place where we should not be

expected to discover via our semantic competence alone. We still need to embrace a rules-as-reasons approach.

5 Conclusions

Wright concludes his classic paper on tolerance and the governing view in a somewhat pessimistic manner: If the governing view is incoherent, but we want to retain (T1) (the model of language use as a game in which players act in accordance with rules), then it looks as though we may have to abandon (T2) (the idea that we can introspectively discover any features of these rules). However, such a move "would seem to force on us a purely behaviouristic concept of how a theory of language-use should be accomplished, and a corresponding shift in the concept of a semantic rule" (Wright, 1975, p. 363). Yet Wright is sceptical towards the potential for such a move. As a final note, however, he suggests an alternative way out: "If the notion of a semantic rule is not to be abandoned altogether, some more restricted account of the epistemology of semantic rules is required than afforded by the governing view" (Wright, 1975, p. 364).

Rather than only focussing on the epistemology of semantic rules (in terms of internalist and externalist conceptions of reasons for action), I have promoted the idea that we should instead demote the rules we grasp as part of our semantic competency from their place as determiners of truthful applications of expressions down to a role in which the rules provide competent speakers with reasons for belief (as hearers) and reasons for linguistic actions (as speakers).

Crucially, however, our reasons for action (which are accessible via internal reflection as in (T2)) only act as heuristics for what will be successful/correct evaluated in terms of externalist reasons. Yet we are not, in making this distinction, ending up with a form of behaviourism. An analysis of internally accessible reasons is vital for explaining the actions of agents as they communicate. We may still discover much, via internal reflection on language use, to aid and guide our linguistic actions. Often these guides may lead us to success in our endeavours. Yet, sometimes, perhaps especially when our guides conflict, success may finally be determined by something we cannot or should not be expected to know in advance.

By seeing rules as reasons, it looks as though there is an avenue open to us on which the notion of a semantic rule is not abandoned altogether, but neither is the idea that features of such rules are discoverable from more than mere behavioural evidence. Furthermore, this approach provides a response to the relocation problem of vagueness. Sharp cut off points will never entirely go away. However, by relocating sharp boundaries to a place where we could not expect agents to discover them as a part of semantic competence, their presence is not only tolerable, but is actually to be expected.

Notes

[1]The formulation given here was popularised in, for example, van Rooij (2011); Cobreros et al. (2012); Smith (2008).

[2]A fully model theoretic representation could be given. For Models $\mathcal{M} = \langle \mathcal{D}, [\![\cdot]\!] \rangle$ with variable assignment g:

$$\text{For all } d, d' \in \mathcal{D}_{\mathcal{M}}, \text{if } [\![P(x) \wedge x \sim_P y]\!]_{\mathcal{M}}^{g[x:=d, y:=d']} = 1, \text{then } [\![P(y)]\!]_{\mathcal{M}}^{g[y:=d']} = 1 \tag{15}$$

[3]This may, however, implicitly restrict our discussion to declarative sentences or sentences used with assertoric force, given that, on the governing view, it is not clear that imperatives and interrogatives should receive a *truth*-conditional interpretation. However, semantics for interrogatives are frequently proposed in terms of sets of propositions/possible answers (a classic reference for which is (Roberts, 1996/2012)). Starr (2016ms) provides a helpful overview of the imperatives literature and provides a proposal for the semantics of imperatives in terms of a dynamic semantics inspired update on interlocutors' preferences (an update on the order of preference for propositions).

[4]It is not enough for a defender of the governing view to simply claim that one knows the vague truth conditions without providing an analysis of what being a vague truth condition amounts to. In the following, the *inexact knowledge* approach that will be discussed could be read as one way of filling out this claim. However, as we shall see, this approach makes some revisions to the governing view.

[5]The alternative being that we drop the idea that rules determine *correct* usage. However, one could doubt the utility of rules that determine incorrect/not fully correct usage.

[6]There may be more strategies available than the inexact knowledge or rules-as-reasons approaches. However, I will limit my investigations to just these two.

[7]On which there is a highly comprehensive commentary and discussion in Égré & Barberousse (2014).

[8]Both Sutton (2013, 2015) and Lassiter & Goodman (2015) argue that vagueness can be characterised in terms of *metalinguistic* uncertainty.

[9]However, MacFarlane's (2009) account involves uncertainty about degrees of truth.

[10]Although Égré & Barberousse (2014) say that Borel's account bears affinity with a graded version of supervaluationism in that degrees are "parasitic on bivalent verdicts" (p. 1055), the sense in which Borel's probabilities are grounded in use, and/or the judgements of competent speakers lends justification to this association.

[11]That Hampton (2007) calls his model the *Threshold Model* could be confusing in this context. The threshold being referred to is a vaguely defined threshold after which two objects are both members of a class, yet differ with respect to their typicality as members of that class.

[12]As pointed out by an anonymous reviewer, this may be a contingent fact about 'red'. There could be world in which actual borderline cases for vague predicates never arise.

[13]Lassiter & Goodman (2015) are also explicitly quietist on the "metaphysical" questions normally posed regarding vagueness in philosophy: "The present account is intended as an answer to the psychological question of how people understand and use scalar adjectives. We do not propose an answer to the metaphysical questions that have occupied much of the discussion of vagueness, involving when a scalar adjective really is applicable to an object, and to what degree. [...] On some plausible assumptions about the nature of meaning, the latter type of question should be illuminated–perhaps even resolved–by an answer to the former. However, we will not attempt this philosophical project here." One way of understanding the project in this article is as pursuing this philosophical project further.

[14]An alternative would be to use the more sophisticated model in Lassiter & Goodman (2015) in which Soft-max choice rules are incorporated.

[15]The PLK/PJC distinction is not clearly demarcated by situation vs. possible worlds theories. For example, Larsson & Fernández (2014) defend a version of PLK within a situation theoretic formalism, and Hampton (2007) is not formulated in a situation or world theoretic system.

[16]Rather than *Situation Types*, Cooper et al. (2015) use *Record Types*. For a discussion on the connections between records, record types, situations, and situation types, see (Cooper, 2005).

[17]A possible concern is that situation types are not the right kinds of entities to form a probability space, the more traditional entities being possible worlds. However, as Cooper et al argue, although the literature on probability theory tends to adopt possible words talk, it is far from clear that full-blown possible worlds are intended. Take the above example of a dice roll. It is not clear that when probability theorists describe a world in which the dice lands even, that they are describing a fully fledged possible word, rather than a situation type.

[18]As pointed out by an anonymous reviewer, a more sophisticated model could make use of Bayesian confirmation theory. This is an intriguing possibility, which certainly merits future detailed consideration. At least for now, an account based on the simplifying assumption seems sufficient for the limited goals of this paper.

[19]This is itself a drastic oversimplification, not least on the role of comparison classes (see, for example, Kennedy (2007)). Whereas Cooper et al. (2015) focus on nouns, Sutton (2015) gives a semantics for adjectives that incorporates semantic features of adjectives from the linguistics literature. In brief, adjectives are treated as functions on probability distributions, where the distributions they modify are given explicitly by the argument noun, by in implicit comparison class, or from a combination of the two. For example, we have reasonable expectations of heights for basketball players. The use of 'tall' with respect to this comparison class encodes a raising of our expectations as to the heights of basket ball players.

[20]In practice, however, values for priors can be given a more empirical basis.

[21]This does not, however, imply that below the threshold one must always apply the negation of the expression.

[22]With the restriction that all resultant values are in the range $[0, 1]$.

[23]Giving values in the form $p(\phi)$ is a simplification for readability. Such probabilities should, on PJC, be conditional on the situation being described.

[24]For example, Davies (2014) considers cases in which the belief in rape myths of members of a jury can prevent a rape victim from performing the illocutionary act of denying she had given consent.

[25]I use notation from Égré (2011).

References

Alvarez, M., Kinds of Reasons: An Essay in the Philosophy of Action (Oxford: Oxford University Press, 2010)

Austin, J. L., Truth. In J. O. Urmson & G. J. Warnock (Eds.) Philosophical Papers, Third Edition, pp. 117–133 (Oxford: Oxford University Press, 1950/1979). Originally in: Symposium: Truth, Proceedings of the Aristotelian Society, Vol. 24 (1950)

Barker, C. (2002). The dynamics of vagueness. Linguistics and Philosophy, 25, 1–36

Barwise, J. & Perry, J., Situations and Attitudes (Mit Press, 1983)

Black, M. (1937). Vagueness. an exercise in logical analysis. Philosophy of Science, 4, 427–455

Borel, É. (1907/2014). An economic paradox: The sophism of the heap of wheat and statistical truths. Erkenntnis, 79(Suppl. 5), 1081–1088. ISSN 1572-8420

Cobreros, P., Égré, P., Ripley, D. & van Rooij, R. (2012). Tolerant, classical, strict. Journal of Philosophical Logic, 41, 347–385

Cooper, R. (2005). Records and record types in semantic theory. Journal of Logic and Computation, 15, 99–112

Cooper, R., Type Theory and Semantics in Flux. In R. Kempson, T. Fernando & N. Asher (Eds.) Philosophy of Linguistics, Handbook of the Philosophy of Science, pp. 271–323 (Elsevier, 2012)

Cooper, R., Dobnik, S., Lappin, S. & Larsson, S. (2014). A probabilistic rich type theory for semantic interpretation. Proceedings of the EACL 2014 Workshop on Type Theory and Natural Language Semantics

Cooper, R., Dobnik, S., Lappin, S. & Larsson, S. (2015). Probabilistic type theory and natural language semantics. Linguistic issues in Language Technology, 10

Davies, A. (2014). How to silence content with porn, context and loaded questions. European Journal of Philosophy, doi: 10.1111/ejop.12075

Decock, L. & Douven, I. (2014). What is graded membership? Noûs, 48, 653–682

Devlin, K., Situation theory and situation semantics. In D. Gabbay & J. Woods (Eds.) Handbook of the History of Logic Vol. 7., pp. 601–664 (Elsevier, 2006)

Edgington, D. (1992). Validity, uncertainty and vagueness. Analysis, 52, 193–204

Edgington, D., Vagueness by degrees. In R. Keefe & P. Smith (Eds.) Vagueness: A Reader, pp. 294–316 (Cambridge, MA: MIT Press, 1997)

Égré, P., Soritical series and Fisher series. In H. Leitgeb & A. Hieke (Eds.) Reduction. Between Mind and the Brian, pp. 91–115 (Ontos-Verlag, 2009)

Égré, P., Perceptual ambiguity and the sorites. In R. Nouven, R. van Rooij & H.-C. Schmitz (Eds.) Vagueness in Communication, Lectures Notes in Artificial Intelligence 6517, pp. 64–90 (Springer-Verlag, 2011)

Égré, P. & Barberousse, A. (2014). Borel on the heap. Erkenntnis, 79, 1043–1079

van Eijck, J. & Lappin, S., Probabilistic semantics for natural language. In Z. Christoff, P. Galeazzi, N. Gierasimszuk, A. Marcoci & S. Smets (Eds.) Logic and Interactive Rationality (LIRA) 2012, Volume 2, pp. 17–35 (ILLC, University of Amsterdams, 2012)

Franke, M. (2009). Signal to Act: Game Theory in Pragmatics. Ph.D. thesis, Universiteit van Amsterdam

Frazee, J. & Beaver, D., Vagueness is rational under uncertainty. In M. Aloni, H. Bastiaanse, T. de Jager & K. Schulz (Eds.) Logic, Language and Meaning: 17th Amsterdam Colloquium, Amsterdam, The Netherlands, December 16-18, 2009, Revised Selected Papers, Lecture Notes in Artificial Intelligence, pp. 153–162 (Springer, 2010)

Groenendijk, J. & Stokhof, M. (1991). Dynamic predicate logic. Linguistics and Philosophy, 14, 39–100

Hampton, J. A. (2007). Typicality, graded membership, and vagueness. Cognitive Science, 31, 355–384

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. Linguistics and Philosophy, 30, 1–45

Larsson, S. & Fernández, R. (2014). Vagueness and learning: A type-theoretic approach. Proceedings of the Third Joint Conference on Lexical and Computational Semantics (*SEM 2014), p. 151159

Lassiter, D., Vagueness as probabilistic linguistic knowledge. In R. Nouwen, U. Sauerland, H. Schmitz & R. van Rooij (Eds.) Vagueness in Communication (Springer, 2011)

Lassiter, D. & Goodman, N. (2015). Adjectival vagueness in a bayesian model of interpretation. Synthese, pp. 1–36

Lassiter, D. & Goodman, N. D., Context, scale structure, and statistics in the interpretation of positive-form adjectives. In Proceedings of Semantics and Linguistic Theory 23, pp. 587–610 (2014)

MacFarlane, J., Fuzzy epistemicism. In R. Dietz & S. Moruzzi (Eds.) Cuts and Clouds, pp. 438–463 (Oxford University Press, 2009)

Raffman, D., Unruly words: a study of vague language (Oxford University Press, 2014)

Rayo, A. (2008). Vague representation. Mind, 117, 329–373

Roberts, C. (1996/2012). Information structure in discourse: Towards an integrated formal theory of pragmatics. Semantics and Pragmatics, 5, 1–69

van Rooij, R., Vagueness and linguistics. In G. Ronzitti (Ed.) Vagueness: A guide, pp. 123–170 (Springer Netherlands, 2011)

Rosch, E. & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. Cognitive Psychology, 7, 573–605

Smith, N. J. J., Vagueness and Degrees of Truth (Oxford University Press, 2008)

Sorensen, R., Blindspots (Oxford: Clarendon Press, 1988)

Stalnaker, R., Assertion. In Context and Content: Essays on Intentionality in Speech and Thought, pp. 78–95 (Oxford University Press, 1978)

Starr, W. B. (2016ms). A preference semantics for imperatives. Manuscript

Sutton, P. (2013). Vagueness, Communication, and Semantic Information. Ph.D. thesis, King's College London

Sutton, P. R., Towards a probabilistic semantics for vague adjectives. In H. Zeevat & H.-C. Schmitz (Eds.) Language, Cognition, and Mind (Springer, 2015)

Travis, C. (1985). Vagueness, observation and sorites. Mind, 94, 345–366

Williamson, T., Vagueness (Abingdon: Routledge, 1994)

Wright, C. (1975). On the coherence of vague predicates. Synthese, 30, 325–365