I outline an interpretive theory of vagueness, emphasizing its advantages as a theory of language in use, compatibility with the data of descriptive linguistics, and its explanation for the interest-relativity of vagueness. Central to the theory is the notion that a ‘language’ in the usual sense is not a well-defined object but a cluster of more or less similar sound-meaning mappings which speakers and interpreters choose among in concrete circumstances to achieve particular communicative goals. Vagueness is argued to be located not in the expressions of any particular sound-meaning mapping, but in the range of such mappings available as interpretive theories in a particular communicative situation. This approach has several advantages: it is derived from an independently motivated theory of interpretation; it explains the interest-relativity of vagueness noted by Graff (2000) without stipulation; and it provides an account not only of how but also of why some predicates are vague and others are not.

Williamson (1994) explains vagueness as an epistemic phenomenon: speakers have only inexact knowledge of their language. Williamson assumes a strong form of semantic externalism according to which a language is an external object of knowledge for speakers, so that, for instance, “[t]he thoughts and practices of speakers of English establish that the truth condition for ‘Harry is bald’ is that Harry is bald…” (1997, p. 217). Thus the character of a speaker’s language is fully determined by facts about his linguistic community, which is assumed to be a well-defined object.

Lassiter (2008) argues that this type of strong externalism is incompatible with the findings of descriptive linguistics that language variation is typically gradual, that linguistic communities cannot be defined with any precision, and that normal usage is subject to considerable variation. In other words, a ‘language’ in Williamson’s sense is a prototypically vague object, unable to fulfill the demands he places on it. As an alternative, Lassiter proposes a theory of interpretation motivated by facts about code-switching and sociolinguistic accommodation. He argues that a ‘language’ is a cluster of sound-meaning mappings whose precise composition can vary considerably from situation to situation, and that partial and shifting overlap between a speaker’s intentions and a listener’s interpretive theory is the norm.

The account of vagueness that I will explore is an epistemic theory of sorts, beginning with this hypothesis: vagueness is not incomplete knowledge of a precisely defined common language, but incomplete knowledge of the intended language of communication, and therefore a natural byproduct of the process of interpretation. Since an interpreter can never have complete knowledge of a speaker’s intended meaning, there is always a range of interpretive theories which, given the linguistic and non-linguistic context, could be used to interpret a speaker’s utterance in a plausible fashion. Vagueness exists because an interpreter can never know exactly which of these theories are intended.

According to Stalnaker (1978), the function of an assertion is to eliminate certain possibilities from the common ground, construed as a set of worlds considered by the conversational participants as live possibilities for how the actual world might be. Stalnaker’s model of assertion can be extended to metalinguistic assertions (cf. Stalnaker (2001)). Among the assumptions that speakers bring to a conversation are facts about the linguistic context, notably a prior theory about what sound-meaning mappings are likely to be useful in communicating information given facts about the situation (the identity of the speaker, the location of the conversation, etc.; see Davidson (1986)). Thus assertions can also serve to eliminate interpretive theories from the common ground. For instance, suppose someone asks, ‘What is an optometrist?’ Given the speaker’s ignorance, there might be two live possibilities: in L₁, ‘optometrist’ and ‘eye doctor’ are mapped to the same concept or set of individuals; in L₂, ‘optometrist’ and ‘plumber’ are mapped to the same concept or set of individuals. In this context, the reply ‘An optometrist is an eye doctor’ serves to eliminate L₁ from the common ground (if we take possible languages as a separate dimension within a multidimensional semantics as in Stalnaker (2001); alternatively the utterance could serve to eliminate possible worlds in which the language of communication in the ongoing conversation is L₂, cf. Barker (2002)).
One consequence of Lassiter’s (2008) notion of a cluster of interpretive theories is that there is a very large number of possible languages, many of which are minimally different from each other. This, then, is a further similarity between languages and possible worlds in the Stalnakerian theory of assertion described above. In principle, any string of sounds could be used to designate any arbitrary set of objects or properties. Suppose that a language L contains only non-vague expressions. If a string of sounds U is taken by an interpreter to express some predicate φ in L which divides an ontologically continuous object or property (e.g. height or the color spectrum) into two or more sections, U could in principle express an unlimited number of distinct predicates φ₁ ... φₙ corresponding to distinct divisions of the continuum, each of which is embedded in a distinct language L₁ ... Lₙ. For example, there are possible languages in which the dividing line between ‘tall’ and ‘not tall’ is 5’11’’, 6’0’’, 6’0.1’’’, 6’0.2’’’, 6’0.21’’’, 6’0.22’’’, and so on.

Suppose that a speaker A utters a string of sounds U directed at a listener B. B’s initial task as an interpreter is to determine, using whatever prior knowledge and contextual clues she can assemble, the intended meaning of U. In normal situations B will have a reasonably clear idea in advance of what kinds of utterances A is likely to use to convey what meanings even for potentially vague utterances. For instance, if B knows that A is a speaker of English and A utters the sentence ‘My new car is red’, B’s prior theory will exclude most possibilities for the reference of ‘red’, e.g. that it refers to old objects or the portion of the color spectrum normally designated by ‘blue’. Thus B, using features of the context (e.g. knowledge of the interests and linguistic dispositions of A or people like A) immediately eliminates the vast majority of potential candidates for the reference of ‘red’, but many candidates remain. Each of these candidates can be taken as a predicate of any of a class of languages each of which contains only non-vague predicates. Even with most possible languages excluded from consideration by a prior theory, then, there is a large number of possible languages which are potentially useful for interpreting A’s utterance.

By this reasoning, the appearance of vagueness boils down to this: the applicability of a vague predicate appears to drop off gradually as we move along a continuum because the interpreter has gradually diminishing justification for taking the speaker’s utterance U to express just this predicate. This theory predicts that if there were for some reason a sharp cut-off in the plausibility of an interpretive theory at a single point in a continuum, vagueness would disappear. This prediction seems to be borne out: a sharp decline in the plausibility of neighboring interpretive theories is precisely what characterizes non-vague predicates. The reason is that the justification for an interpretive theory is intimately tied up with reasoning about a speaker’s interests, values, and communicative purposes. A human speaker is likely to use terms that apply to groupings of objects or properties that serve some particular human purpose; so, for example, a term covering tables, chairs, and lamps (‘furniture’) is useful not because these items form a natural class ontologically, but because they are useful to humans in particular ways. Similar groupings make less plausible linguistic items not because they are ontologically more diverse than furniture, but because no obvious human purpose would be served in talking about them.

I suggest that the reason that some predicates are non-vague is that there is a sharp cut-off in the plausibility of any interpretive theory that extends them further. For example, ‘human’ is relatively clear: fetuses excepted, there is little room for borderline cases. An interpretive theory taking various human-like things (e.g. chimps and cardboard cut-outs of people) to be borderline cases of ‘human’ would be wildly implausible given known human interests (e.g., talking to other people) and the fact that there are sharp boundaries in the things in the world that can fulfill these interests. The theory proposed here, like Graff (2000), makes the prediction that vagueness should be predictable from human interests. Graff explains interest-relativity by enriching the logic with a NORM operator; in contrast, I propose that the interest-relativity of vagueness can be derived from an independently motivated theory of interpretation.