# Imperfect Perception and Vagueness

Giri Parameswaran[*]& Timothy Lambie-Hanson[†]

March 27, 2016

VERY PRELIMINARY & INCOMPLETE

**Abstract**

This paper investigates the epistemic approach to vagueness — that the source of vagueness is the speaker's inability to perfectly perceive the world. We study a standard communication game and show that imperfect perception is insufficient to render vague communication about the world as perceived by the sender. However, if the receiver interprets the sender's message as a true statement about the world, rather than merely how it appears to the sender, then language becomes vague. We show that this vagueness is characterized by probability distributions that describe the degree to which a statement is likely to be true. Hence, we provide micro-foundations for truth-degree functions as an equilibrium consequence of the sender's perception technology and the optimal, non-vague language in the perceived world — thereby unifying the epistemic and truth-degree approaches to vagueness.

---

[*]Department of Economics, Haverford College, Haverford, PA 19041. Email: gparames@haverford.edu
[†]Department of Economics, Haverford College, Haverford, PA 19041. Email: tlambiehan@haverford.edu

# 1 Introduction

Language is vague when the receiver cannot be certain of which 'states' the sender of a message seeks to invoke. For example, although we routinely describe people or things as "tall," "heavy," "fast," and so on, we would, in most cases, struggle to identify the boundary between tall and short, or between fast and slow. Indeed, vagueness is often associated with a 'blurring of the boundaries' between the meanings of words. Many theories seek to explain the nature and source of vagueness. In this paper, we investigate epistemicism — the idea that vagueness arises because agents perceive the world imperfectly, and so cannot describe it in a way that is crisp — as a source of vagueness. We begin with the observation that communication in the presence of imperfect perception is fundamentally different from communication with perfect perception; imperfect perception limits the scope of communication to statements about the perceived world rather than about the actual world. We show that imperfect perception does not cause language about the perceived world to be vague. We additionally show that, if this perceived-world language is extended to the actual (objective) world — i.e. if a claim about what appears to be is interpreted as a claim about what is — then meaning will be vague in the actual world. However, this vagueness is metaphysical (it inheres in the objects being described) rather than epistemic, and is closest in spirit to the truth-degree (continuum-valued logic) approach. We thus provide micro-foundations for truth-degree functions as the consequence of the optimal, non-vague language in the perceived world, and the sender's perception technology, thereby unifying the epistemic and truth-degree approaches to vagueness.

Language has long been recognized (see David (1969)) as a coordination device that facilitates communication. The sender's choice of message depends on his belief about how that message will be interpreted by the receiver, and the receiver's interpretation will depend on how she expects the sender to use the messages available to him. The meaning of words and messages, then, are not exogenous, but arise as a consequence of a communication game between sender and receiver. To capture this dynamic, we develop a formal model of communication in the spirit of Crawford and Sobel (1982). The sender observes an informative, but imperfect, signal about a state of the world. The sender transmits a message to the uninformed receiver, who then takes an action that affects both parties. We make the behavioral assumption that the communicants have a limited vocabulary, and so the sender has fewer messages at his disposal than there are states to describe. To highlight the effect of epistemic uncertainty, we abstract from other frictions (such as preference disagreement or 'type' uncertainty), that may induce vagueness through other channels. Our model characterizes the optimal

*use* of language by the sender, given the anticipated response by the receiver. The *meaning* of language is pinned down in equilibrium, given the sender's use. The optimal language is *imprecise* if the sender associates multiple states with the same message. The language is *vague* if the sender additionally associates multiple messages with the same signal or state. Accordingly, we distinguish vagueness and imprecision, and show that imprecision is a necessary condition for language to be vague.

As we previously noted, the nature of communication fundamentally differs between perfect and imperfect perception environments. Whereas a perfectly perceiving sender can make objective statements of the form 'John is tall', an imperfectly perceiving sender can only claim that 'John *appears* tall (*to me*)'. Imperfect perception relegates communication to the world of subjective claims, even if the sender renders a statement in a seemingly objective way. We show that imperfect perception alone is insufficient to generate vagueness in the realm of statements about apparent truths. Even if the sender is uncertain that what he perceives is true, this should not prevent him from clearly indicating what he has perceived. To do so, we show that the sender will optimally partition the set of perceived states so that each apparent state is associated with precisely one message. For example, there will be a threshold that partitions the set of apparent heights into those that appear tall and those that appear short.

It is common, however, to interpret subjective statements as objective ones — e.g. we often take the statement 'John is tall' to actually be an objective claim about John's height. And we are inclined to do so, even knowing that the sender does not have privileged access to objective truths. Reflecting this tendency, we extend the perceived world language to the actual world, and analyze its meaning in this new space. To be clear, the sender's use cannot be different in the extension, since the sender does not observe the true state when choosing his message. However, if the true state is revealed *ex post*, we may observe the behavior of the sender's language use, after the fact.[1] Given that the sender's perception is imperfect, there will be borderline cases where the sender classifies persons of the same height as tall in some instances, and short in others. In this *ex post* sense, language appears vague, since multiple predicates are associated with the same state of the world. The language that was well defined in the subjective world becomes vague when extended to the objective world. We show that the sender's ideal language (in the subjective space) combined with the technology that governs perception, induces a probability distribution that describes

---

[1]This is consistent with Williamson's (1994) account in his motivating example, where a sender claims that there are 30,000 people at a sporting arena, when in reality the actual number is close to, but not exactly 30,000. At the time of his utterance, Williamson's sender does not know the true number of attendees, although we may be able to determine this *ex post*, by counting ticket sales (for example).

the likelihood that each predicate is ascribed to a given object. As long as perception is imperfect, this probability distribution will be non-degenerate over a range of 'borderline' outcomes.

In its extension to the objective world, our model has strong similarities to the continuum-valued logic approach to vagueness. This approach rejects the principle of bivalence and instead conceives of statements as having 'truth-degrees' that range from zero (definitely false) to one (definitely true). Since use determines meaning, and our model determines the likelihood of using a particular message to describe a given state, we, in effect, provide micro-foundations for the assignment of truth-degrees. Our model, therefore, unites two distinct approaches to explaining vagueness that are predominant in the literature. We use epistemic theory to provide the causal mechanism that generates the descriptive features of the truth-degrees approach. Truth-degrees are determined *in equilibrium*, given the properties of the optimal language in the subjective realm and the properties of the technology that governs perception. Similar to canonical truth-degree models, our induced truth-degree functions respect comparisons over ordered objects — if John is taller than Mary, then the truth-degree assigned to John being tall will be at least as large as the truth-degree for Mary being tall. However, in contrast to many truth-degree theories, our truth-degree functions are not truth-functional. Instead, our truth-degree functions satisfy the axioms of probability (which still permits the assignment of truth-degrees to compound statements if the joint-probability distribution is known) as well as standard rules of logic such as the Law of the Excluded Middle.

# 2 Literature

## 2.1 Philosophical Accounts

There are many accounts of the sources and characteristics of vague language. These can be broadly categorized into three approaches: metaphysical, semantic and epistemic. (See Smith (2008).) Metaphysical accounts attribute vagueness directly to properties of the object being described. For example, whilst it is clear that a sky-scraper is tall, it is unclear whether a ten-story building ought to be described as tall or not. This clarity, or lack thereof, arises directly from the building's height, and is inherent to the object being described. As a matter of logic, metaphysical accounts are typically forced to reject the principle of bivalence — it

may be neither (clearly) true nor (clearly) false that the building is tall — in favor of multi-valued logics (see Halldén (1949)). The continuum-valued logic approach (see Zadeh (1975), Smith (2008) amongst others) is a particular instantiation of this approach, which replaces the binary notions of truth or falsity with 'truth-degrees' which can take any value from 0 (clearly false) to 1 (clearly true). These truth-degrees are typically taken as primitives of the language. A useful property for truth-degrees is truth-functionality — the property that the truth degree of a compound statement can be determined purely from the truth degrees of the constituent simple statements. Indeed, most truth-degree proponents endow truth-degrees with this property (Edgington (1997) being a notable exception). However, as Fine (1975) demonstrates, truth-functionality causes truth-degree to be inconsistent with the laws of probability and standard results of logic, such as the Law of the Excluded Middle. Our model departs from the canonical truth-degree approach in two ways. First, truth-degrees are not primitives in our model. Rather, they are determined in equilibrium by more primitive features, such as the sender's perceptive faculties. Second, we construct truth-degrees as probability measures, making them consistent with standard results in logic. We do so at the cost of truth-functionality, although with a complete specification of the joint probability distribution of events, we can still assign truth degrees to compound statements.

Semantic accounts attribute vagueness to indeterminacy in the way language is used by different speakers. Plurivaluationism (see Smith (2008)), captures this idea that different speakers may describe the same object differently. It is closely related to, although distinct (as Smith (2008) takes pains to argue) from, Supervaluationism (see Dummett (1975), Fine (1975), Keefe (2000)), which posits that language is vague when its extension to indeterminate cases admits multiple interpretations. Under this approach, vagueness arises because of an inability by the community to coordinate on a common use of language.

The epistemic account (see Williamson (2002) and Sorensen (2001)) locates vagueness in the limitations of human perception. Proponents of this approach insist that vagueness is neither metaphysical (it does not inhere in objects) nor semantic (it is not a consequence of who is communicating). Language itself is well-defined and characterized by sharp thresholds, and a perfectly informed speaker would use language in a way that is consistent with its meaning. Vagueness arises because imperfect perception prevents agents from precisely comparing the true state of affairs against these thresholds. A challenge for the epistemic theorist is that, by this account, these thresholds are seemingly determined independently of the speaker's usage, thereby severing the link between use and meaning.[2] Our model lays bare this challenge.

---

[2]Williamson (1994) argues that the mechanism linking use and meaning may be complicated and unknown to the philosopher — but that this in no way refutes that the former determines the latter. We find

The speaker's use of messages is determined by what he perceives, and his expectation of how any given message will be interpreted. The receiver, in turn, interprets the meaning of messages based on her expectation of how the sender uses each message. This implies that, if the parties communicate optimally, language will be characterized by firm thresholds in the subjective world. However, meaning necessarily cannot be governed by firm thresholds when extended to the actual world, since the same actual state may be mapped onto multiple perceived states that are associated with different messages. If thresholds exist that delineate meaning in the actual world, they must be generated by some mechanism other than the sender's use of language.

## 2.2  Economic Accounts

There is a long literature on the economics of communication dating back to the canonical models of persuasion (see Grossman (1981) and Milgrom (1981)) and cheap talk (see Crawford and Sobel (1982)). Crawford and Sobel (1982) study a communication game between an informed sender and an uninformed receiver who must take an action that affects both parties. The paper provides two important insights. First, it demonstrates the (equilibrium) relationship between use and meaning; the sender's use is determined by the meaning ascribed to each message by the receiver, and these ascribed meanings are in turn determined by the sender's anticipated use. Second, differences in preferences between a sender and receiver generate incentives for the sender to not fully reveal his information to the receiver, thereby rationalizing imprecise communication.[3] Qing and Franke (2014), building on the model in Lassiter and Goodman (2014), present a variant of this analysis in which the communicants' strategies are probabilistic, reflecting satisficing rather than perfectly maximizing behavior.

However, until recently, economic models of communication did not feature messages that

this account difficult to sustain. Plainly, if use determines meaning, it cannot be that meaning is determined by factors inaccessible to the speaker when choosing which words to use. The mapping from use to meaning should be readily determined by simply observing how the sender uses his words. We understand the meaning of the word 'tall' by observing all the instances in which we describe an object as tall. To say that use determines meaning isn't to merely suggest that there is *some* mechanism that links the meaning of words to their use. Rather, it is the stronger claim that use is itself that mechanism. Smith (2008) provides a more detailed critique.

[3]Imprecision is to be distinguished from vagueness. A message is imprecise if the sender associates multiple sates with that message, thereby preventing the receiver from exactly learning the true state. By contrast, a message is vague if it is unclear which states are associated with that message. For example, it is imprecise to say that 'John's height is at least 6 feet', since saying provides the receiver with a range of possible heights for John, rather than his actual height. However, the message is not vague — it clearly delineates the set of heights that John may have.

could be construed as vague (Lipman, 2003, 2009). To our knowledge, Blume and Board (2013a), with what they term 'message indeterminacy', are the first to study what we term vagueness. They do so by assuming uncertainty about language competence (which roughly corresponds to the richness of vocabulary). 'Message indeterminacy' arises when the receiver is uncertain about the sender's language competence. Lambie-Hanson and Parameswaran (2016) study a communication game in which language is ideally modified to suit different 'contexts'. (For example, 'tall' would be used differently in the context of the town of Lilliput as compared to the town of Brobdingnag.) In such a game, meaning will be clear as long as the receiver correctly perceives the sender's belief about the context. Vagueness arises when the sender's language use fails to coincide with the receiver's belief about how the sender is using language. As an example, a sophisticated Lilliputian who knows to use the word 'tall' differently when communicating in the Brobdingnag context, will be able to communicate without misunderstanding. Vagueness arises, not because different contexts *per se*, but because of a lack of common knowledge about how each communicant is modifying his language to suit the context at hand. Both these theories are *semantic* in that they locate the source of vagueness in differences in (expected) language use between the communicants. Importantly, in both cases, speakers are not intentionally vague. The sender always transmits a message with well-defined meaning; vagueness arises when the meaning inferred by the receiver and the meaning intended by the sender, diverge.

A different approach locates the source of vagueness in frictions in the communication technology itself, which may cause messages to become 'garbled' during transmission. Blume and Board (2014) demonstrate that garbling may provide an incentive for the sender to be intentionally vague. Rick (2013) similarly shows that there may be deliberate miscommunication in the presence of garbling, and that this may improve outcomes for both parties.

Finally, if communication is costly, vague communication may be optimal, even if messages are transmitted perfectly. For example, Mialon and Mialon (2013) study a model where some messages may be more costly than others and show that, with a finite number of states a figurative (vague) message may be preferable. Lambie-Hanson and Parameswaran (2016) explore the related idea that developing a sizable vocabulary may be costly. They show that the incentive for parsimonious communication causes communicants to use the same word to describe distinct ideas, and that this can render language vague if the set of ideas that are 'pooled' into a given message (rationally) vary across speakers.

# 3    Model

In studying the question of vagueness, the literature has largely focused its attention on gradable adjectives, such as 'tall' or 'heavy' or 'cold'. Adjectives are gradable when they express differences in intensity, quality or 'grade'. For example, we may describe a person as 'short', 'medium-heighted', or 'tall'. These words are clearly ordered in a natural way that expresses differences in the underlying object being graded. The potential for vagueness arises when the underlying property being described (i.e. the height, or weight or temperature) can vary gradually, whilst the adjectives themselves do not. By necessity, the speaker must use the same adjective-grade to describe similar but distinct objects. Vagueness arises when it is unclear which objects ought to be 'pooled' together, or when the speaker is unable (or unwilling) to draw a line demarking the use of different grades of the adjective. The stylized model, below, is designed to capture these essential features.

Let $X = [0, 1]$ be the set of possible states. There is a partially-informed sender (he) who observes a noisy signal $y$ about the true state $x$, and an uninformed receiver (she) who has no information about the state. Both communicants share common prior beliefs over the likely realization of the state, represented by the distribution function $F$, with associated density $f$. The receiver must take an action $a \in X$ that affects both the sender and receiver. We abstract from cases where the sender has incentives to hide information from the receiver (such as the standard models of cheap talk) by assuming that the sender and receiver have identical preferences. This enables us to focus attention on the effect of perceptive limitations on communication. We assume that both agents have state-dependent preferences represented by the utility index: $u(x, a) = -(x-a)^2$. Intuitively, the agents seek to match the action to the realized state of the world, and suffer increasingly larger losses as the action deviates from the true state. Utility maximization is tantamount to maximizing the efficiency of communication between sender and receiver.

The sender observes signal $y \in Y = [0, 1]$. The signal technology is represented by a distribution function $Q(y|x)$ with density $q(y|x)$. The density $q(y|x)$ is the 'likelihood' that the sender observes signal $y$ given that the true state is $x$. We assume that the signal is informative in the sense that a higher signal statistically indicates a higher true state. This property is formalized in the following assumption:

**Assumption.** *Q satisfies the monotone likelihood ratio property. I.e. if $x_1 > x_0$ then $\frac{q(y_0|x_1)}{q(y_0|x_0)} \leq \frac{q(y_1|x_1)}{q(y_1|x_0)}$ whenever $y_1 > y_0$.*

The monotone likelihood ratio property is standard in signaling games. It implies that if the

true state is high, then the sender will be more likely to observe a high signal than a low signal, than if the true state is low. Apart from this restriction, the perception technology is quite general, and there is considerable scope for variation in precision and bias. For example, the technology is consistent with a sender who perceives very tall and very short buildings quite accurately, but is more error prone when observing buildings of intermediate height. Similarly, it is consistent with a sender who systematically misperceives buildings as taller than they are.

After observing a given signal, the sender can make inferences about the true state of the world, according to Bayes' Rule. Let $F(x|y)$ denote the sender's posterior belief about the true state after observing signal $y$. The monotone likelihood ratio property implies that these posterior beliefs respect first-order stochastic dominance.[4] After observing a higher signal, the sender rationally infers that the true state is more likely high than low.

Upon observing the signal, the sender can send a message $m \in M = \{m_1, ..., m_K\}$ to the receiver. The set of messages is finite, capturing the idea that the communicants' share a limited vocabulary, and ordered, capturing the grades of an adjective. We can think of each message as behaving similarly to a first-order predicate. To transmit message $m_k$ is to ascribe to a subject the $k^{th}$ degree of a gradable adjective (with $K$ possible degrees). In this section's motivating example, we had three possible messages, with $m_1$ indicating that the subject is 'short', and $m_3$ indicating that the subject is 'tall'.

A strategy $\mu : Y \to M$ for the sender assigns a message $\mu(y)$ to each signal $y \in Y$. A strategy $\alpha : M \to X$ for the receiver assigns an action $\alpha(m)$ to each message received. Let $F(x|y)$ be the sender's posterior belief about the true state after receiving signal $y$, and let $G(x|m)$ be the receiver's posterior belief about the true state after observing message $m$. Let $f(x|y)$ and $g(x|m)$ be the associated densities. A *sequential equilibrium* is a strategy $\mu$ for the sender, a strategy $\alpha$ for the receiver, and a pair $(f(\cdot|y), g(\cdot|m))$ of belief functions which satisfy:

1. For each signal $y \in Y$, the sender chooses the message which maximizes his expected utility, given his posterior beliefs and the equilibrium strategy of the receiver:

$$\mu(y) = \arg\max_{m \in M} \left\{ -\int_X (x - \alpha(m))^2 f(x|y) \, dx \right\}$$

___
[4]I.e. $y_1 > y_0$ implies $F(x|y_1) \leq F(x|y_0)$ for every $x \in X$.

2. For each message $m \in M$, the receiver chooses the action which maximizes her expected utility, given her posterior beliefs:

$$\alpha\left(m\right) = \arg\max_{a \in X}\left\{\int_X \left(x - a\right)^2 g\left(x|m\right) dx\right\}$$

3. The communicants' posterior beliefs are determined according to Bayes' Rule, given their common prior beliefs and the equilibrium strategies:

$$
\begin{aligned}
f\left(x|y\right) &= \frac{f\left(x\right)q\left(y|x\right)}{\int_{z \in X} f\left(z\right)q\left(y|z\right)dz} \\
g\left(x|m\right) &= \frac{f\left(x\right)\left[\int_{y \in Y} q\left(y|x\right)\mathbf{1}_{\mu(y)=m}\left[y\right]dy\right]}{\int_{z \in X} f\left(z\right)\left[\int_{y \in Y} q\left(y|z\right)\mathbf{1}_{\mu(y)=m}\left[y\right]dy\right]dz}
\end{aligned}
$$

It is well known that communication games of this sort typically admit multiple equilibria. We alert the reader to two important sources of multiplicity. First, there are equilibria in which the sender transmits fewer messages than are available, thereby forgoing opportunities to send more informative messages. A particularly stark example is the 'babbling equilibrium' in which the sender transmits a randomly selected message, and the receiver implements $E\left[\theta\right]$ regardless of which message is received. Second, there are equilibria in which the sender uses all available messages, but the receiver interprets these in unusual ways. For example, the sender may use the word 'tall' to describe a person who appears small-heighted and use the word 'short' to describe a person who appears large-heighted. This is an equilibrium provided that the receiver understands the sender's usage. The first type of problem involved inefficient under-utilization of messages sustained by beliefs by both communicants that messages would be uninformative. Since the communicants are incentivized to communicate as efficiently as possible, such pessimistic beliefs are quite strange. The second type of problem involved an unnatural mapping between messages and meaning. At its core, an equilibrium language is simply a code that is commonly understood by the communicants. There is no requirement that the code conform to the natural ordering over messages. Nevertheless, there is a clear sense in which the equilibrium in which 'tall' refers to large-heighted people and 'short' refers to short-heighted people is more natural and hence focal.

In this paper, we assume that the communicants are able to coordinate on the efficient, natural, focal equilibrium. Formally, following Chen (2011), we refine the set of equilibria to focus on *message monotone sequential equilibria*.

**Definition.** A *message monotone equilibrium* is a sequential equilibrium which satisfies the following: If $\mu(y) = m_i$ and $\mu(y') = m_j$ with $i < j$, then $y < y'$.

Message monotonicity insists that the sender does not use higher-ordered messages to describe lower-ordered states. As we demonstrate in the proof of Proposition 1, it is sufficient to select the most efficient, naturally ordered equilibrium from amongst the multiple possible equilibria.

This paper uses a formal model to analyze the effect of imperfect perception on the nature of language. As is standard in all formal models, we make several assumptions that keep the model simple and tractable. Our goal is to focus on factors relevant to the issue of interest; namely the effect of imperfect perception on language and communication. To this extent, we abstract from other factors that may be salient in their own right, but are not crucial to the epistemic story. Before proceeding to the equilibrium analysis, we briefly comment on some of our modeling choices to provide assurances that our assumptions are benign.

First, and in common with most 'economic' models, we locate decision making within the rational choice framework. We acknowledge that agents are not always (or even ever) perfectly rational. Nevertheless, this assumption plays an important role in determining how the communicants would ideally use and assign meaning to language, even if they occasionally make mistakes. We think it is a useful fiction that genuinely captures the interaction between the communicants. Moreover, we note that many features of our equilibrium are robust to error-prone decision making by the agents. For example, suppose the a utility maximizing sender occasionally 'trembles' and sends the wrong message, that a utility maximizing receiver similarly occasionally 'trembles' and chooses the wrong action, and that both agents are aware of the fallibility of the other. Solving for a 'trembling-hand perfect equilibrium' (see Selten (1975)), we note that the optimal communication strategies remain qualitatively unchanged, and that in particular, the sender continues to choose messages according to a threshold strategy.

Another simplifying feature of our model is that both communicants share a common prior over the likely state of the world, and that all aspects of the model (other than the true state) are commonly known by the players. In particular, we assume that the receiver understands the nature of the sender's perception technology. (Since the sensory abilities of humans are roughly similar, we think it is not unreasonable to assume that the receiver can predict how the sender may misperceive the world.) Again, we acknowledge the strength of these assumptions, and the reality that the communicants' beliefs about these objects may not

perfectly align. However, whilst such differences may affect the nature of communication, they are not intrinsically linked to the problem of imperfect perception. To the extent that these features induce vagueness, they do so through the channel of interpersonal differences between the communicants, and therefore more properly represent a *semantic* source of vagueness rather than an epistemic one. (See Blume and Board (2013b), Lambie-Hanson and Parameswaran (2016) and Körner (1962).) Since these forces would continue to operate even if the sender's perception were perfect, their abstraction does not pose a threat to understanding the epistemic account of vagueness.

# 4  Analysis

Recall that to transmit message $m_k$ is to ascribe to the subject the $k^{th}$ degree of the gradable adjective. But what precisely is the subject? We previously distinguished between subjective messages that described what the sender perceives from objective messages which describe what is. Subjective message are of the form: 'The state appears to have property $m_k$'. Since such statements are conditioned upon the signal received by the sender, we say that they live in $Y$-space. By contrast, objective statements are of the form: 'The state actually has property $m_k$'. Since these statements are about the true state, we say that they live in $X$-space. From herein, we use $X$-space and $Y$-space as a shorthand for indicating objective and subjective claims, respectively.

We these distinctions in mind, we turn to solving the model. Our analysis is in two parts. First, we characterize the equilibrium in the communication game between the imperfectly informed sender and the uninformed receiver. The equilibrium determines how language will be used by the sender to describe the world as it appears to him. We then analyze the extension of this language to claims about the objective world.

## 4.1  Equilibrium and Properties of 'Apparent' Statements

We being by characterizing the equilibrium of this game:

**Proposition 1.** *There exists a vector* $(s_0, ..., s_K) \in Y^{K+1}$ *with* $0 = s_0 < ... < s_K = 1$, *that defines a message monotone sequential equilibrium in which:*

1. *The sender transmits message* $\mu(y) = m_k$ *whenever* $y \in (s_{k-1}, s_k]$;

2. *The receiver takes action $\alpha\left(m_k\right) = \int_X x g\left(x|m_k\right) dx$ after receiving message $m_k$; and*

3. *The communicants' belief functions satisfy:*

$$
\begin{aligned}
f\left(x|y\right) &= \frac{f\left(x\right) q\left(y|x\right)}{\int_{z \in X} f\left(z\right) q\left(y|z\right) dz} \\
g\left(x|m_k\right) &= \begin{cases} \frac{f(x) \int_{s_{k-1}}^{s_k} q(y|x) dy}{\int_{z \in X} f(z)\left[\int_{s_{k-1}}^{s_k} q(y|z) dy\right] dz} & y \in \left(s_{k-1}, s_k\right] \\ 0 & otherwise \end{cases}
\end{aligned}
$$

The proposition states that sender partitions the signal space into $K$ disjoint intervals, such that each interval is associated with a given message. The sender transmits message $m_k$ whenever the received signal is contained within the $k^{th}$ interval. Several properties of the equilibrium are worth noting.

First, given the signal technology, both communicants form beliefs about the likely true state according to Bayes' Rule. For example, after observing signal $y$, the sender's beliefs about the true underlying state $x$ are given by the density $f\left(x|y\right)$. Let $S_S\left(y\right) = \{x \in X | f\left(x|y\right) > 0\}$ be the support of the sender's conditional beliefs, which is the set of possible true states that the sender cannot rule out. This corresponds to the *'margin for error'* in Williamson (1994). The receiver similarly forms beliefs about the likely true state. Although she doesn't observe the sender's signal, she can make such inferences given the message she receives, and her knowledge of the sender's optimal communication strategy. Upon receiving message $m$, the receiver's belief that the true state is $x$ is given by the density $g\left(x|m\right)$. In determining their optimal choices, both sender and receiver use these updated beliefs about the true state; i.e. both players take into account the possibility that the sender misperceives when making their choices.

Second, upon receiving message $m_k$, the receiver's choice of optimal action $\alpha\left(m_k\right)$ simply reflects her best guess about the true state, given her information. If the receiver knew the state perfectly, she would choose the action that precisely matched the state. Since she does not, she chooses the action that matches the state in expectation, given her updated beliefs. It should thus be clear that the modeling fiction of the receiver taking an action simply serves to capture the process of information transmission between sender and receiver.

Third, the sender's optimal strategy associates a unique message with every possible signal. To see why, note that the receiver chooses a different action for each different message received. If so, the sender will generically not be indifferent between transmitting each of

the available messages, but rather has a strict incentive to send the message that induces the action that is closest to the expected true state. Hence, there will generically be a unique message associated with each signal. In spite of the sender's uncertain perception, he will typically be certain about the message he wishes to send, given what he perceives and given the receiver's anticipated response.

Since the sender's optimal strategy partitions the signal-space, we can find thresholds $\{s_0, ..., s_K\}$, which delineate the intervals and determine which message is sent. For example, there will be some threshold perceived height, such that the sender will report that the 'building appears tall' whenever his signal of the building's height exceeds this threshold. Two comments about this threshold strategy are worth noting. First, we stress that the threshold strategy arises as an equilibrium result, rather than as an assumption of the model. Nothing in our model compels the sender to use a threshold strategy. Instead, we demonstrate that efficient communication between sender and receiver requires that the sender partition the signal-space using clearly delineated thresholds.

Second, we address the common objection[5] that the location of thresholds is arbitrary and so threshold strategies ought to be impermissible. It is indisputable that in drawing thresholds, we distinguish seemingly similar states which just happen to fall on opposite sides of the threshold. Taken in isolation, such distinctions do indeed appear arbitrary. Nevertheless, when considered globally, these thresholds are in fact located optimally when considered globally. Threshold strategies are a consequence of the agents' limited vocabulary. If there were no limit on the number of degrees that we could express, we would associate a separate message with every possible signal, thereby appropriately acknowledging every nuance and distinction between signals (or states). Since we make the reasonable assumption that our vocabulary is limited, we are forced to 'pool' several states into the same message. An unavoidable consequence of pooling is that some pairs of states will be treated identically when pooled together even though they are distinct, whilst other seemingly similar pairs of states will be treated differently by virtue of not being pooled together. The more states that are pooled into the same message, the less informative that message will be. The challenge for optimal communication is to pool states together in the way that facilitates efficient information transfer. In our model, the location of the thresholds $\{s_0, ..., s_K\}$ have the property of minimizing the expected (square) deviation between the true state and the receiver's expectation of the state. It should be clear that the location of these thresholds

---

[5]For example, the Sorites Paradox, which is commonly associated with the problem of vagueness arises precisely because of a rejection of threshold behavior. With threshold behavior, the induction argument that generates the Sorites series would not hold globally.

depends on the global properties of the system, anticipating the agents' likely communication needs.

An important consequence of the above proposition is that language is not vague in the $Y$-space. The sender's communication strategy is characterized by an unambiguous mapping from signals to messages, and this is understood by the receiver. Given the perceived height of any building, the receiver knows whether the sender will describe it as tall or not. Although imperfect perception may leave the sender with some doubt about the true state, it does not prevent him from clearly indicating the signal that he has perceived. This is consistent with the critique in Lipman (2009), that a speaker should not intentionally use language in a manner that is vague.

To demonstrate the features of the equilibrium language, we construct the following stylized example:

**Example.** Suppose the state $x$ is drawn from a uniform distribution on $[0, 1]$, and that, conditional upon the realized state $x$, the sender observes a signal $y$, which is itself drawn from a uniform distribution on $[x - \varepsilon, x + \varepsilon]$.[6] The signal precision, or 'margin for error' is parametrized by $\varepsilon > 0$, where a larger $\varepsilon$ implies less perfect perception. (We assume $\varepsilon < \frac{3}{8}$ for technical convenience.) For any true state, the sender's signal is contained within a band of uniform width. The size of this band indicates how accurately the sender perceives the world.

Suppose $K = 3$, so that the sender has access to three messages (e.g. small, medium and large). Then, the equilibrium is characterized by thresholds: $s_1(\varepsilon) = \frac{1 + \sqrt{1 + 4\varepsilon^2}}{6}$ and $s_2(\varepsilon) = \frac{5 - \sqrt{1 + 4\varepsilon^2}}{6}$. The sender transmits message 1 whenever he observes a signal $y$ in the interval $[-\varepsilon, s_1(\varepsilon)]$, he transmits message 2 whenever he observes a signal in the interval $(s_1(\varepsilon), s_2(\varepsilon)]$, and he transmits message 3 whenever he observes a signal in the interval $(s_2(\varepsilon), 1 + \varepsilon]$. The receiver's optimal action after each message are: $a_1(\varepsilon) = \frac{\sqrt{1 + 4\varepsilon^2}}{3} - \frac{1}{6}$, $a_2(\varepsilon) = \frac{1}{2}$ and $a_3(\varepsilon) = \frac{7}{6} - \frac{\sqrt{1 + 4\varepsilon^2}}{3}$. We provide a full characterization of the equilibrium, including the equilibrium belief functions, in the appendix.

As a benchmark, note that if $\varepsilon = 0$, so that the sender perfectly perceives the world, then $(s_1, s_2) = \left(\frac{1}{3}, \frac{2}{3}\right)$ and $(a_1, a_2, a_3) = \left(\frac{1}{6}, \frac{1}{2}, \frac{5}{6}\right)$. The sender partitions the state-space into three equally sized intervals, and the receiver implements the action which corresponds to the expected state in each interval. Partitioning the state space into equally sized intervals ensures

---

[6]We note that the signal space is now $Y = [-\varepsilon, 1 + \varepsilon]$, which no longer coincides with the state space $X = [0, 1]$. If desired, we can easily find an order preserving transformation that maps the signals back onto $[0, 1]$. For example, it suffices to take $y' = \frac{y + \varepsilon}{1 + 2\varepsilon}$.

that message sent is equally informative, no matter which state of the world is realized. With a uniform prior, this ensures that information transfer is efficient. We notice that the senders' thresholds and the receiver's optimal actions, are responsive to the signal precision $\varepsilon$ — both communicants are aware that the sender imperfectly perceives the world, and they adjust their use and understanding of language accordingly. Importantly, an imperfect sender's use may systematically vary from the perfect-perception benchmark; the imperfectly informed sender doesn't simply implement the perfectly-perceiving sender's strategy 'on average'.[7]

In summary, we have shown that the optimal communication strategy for an imperfectly perceiving sender is to partition the signal space into disjoint intervals (characterized by thresholds), and to associate a distinct message with each interval. Accordingly, we have shown that imperfect perception alone is not sufficient to cause language to be vague. As we noted above, vagueness may still arise in this environment if we introduce additional frictions to communication. For example, we could introduce multiple 'types' of receivers, where a 'type' may capture differences in the receiver's beliefs about the underlying distribution of the state, the perception technology, the size of the vocabulary, and so on. Blume and Board (2013a) and Lambie-Hanson and Parameswaran (2016) demonstrate that introducing higher order uncertainty about these types is sufficient to introduce semantic vagueness into the equilibrium language. However, since our focus in this paper is on the effect of imperfect perception, and these effects operate independently, it does not pose a threat to our analysis to ignore those other channels.

## 4.2 Properties of Statements about Actualities

We now turn our attention to statements about actualities, rather than appearances. We begin by noting that, if the sender could directly observe the true state $x$, we could simply repeat the above exercise and characterize the optimal language in $X$-space. This ideal language over $X$-space would retain all of the characteristics of the optimal $Y$-space language, including that the use of messages is delineated by sharp thresholds. But for the sender's fallible perception, these would be the (epistemic theorist's ideal) thresholds that governed our use of language. We caution, however, that, if we take impediments to perfect

---

[7]Indeed, as the signal imprecision $\varepsilon$ increases, the sender will be more likely to transmit messages $m_1$ and $m_3$, and less likely to transmit $m_2$. With greater imprecision, the receiver recognizes that a given signal is consistent with a larger range of true states. Then, if the thresholds did not change, the average state which generated message $m_1$ would be higher, and the average state which generated $m_3$ would be lower — causing the receiver to choose higher $\alpha_1$ and lower $\alpha_3$, respectively. But, this feeds back into the sender's choice, making him less inclined to transmit $m_2$.

perception seriously, such an exercise is no more than a thought experiment. Whilst we can conceive of such a language, it is not actually available for use by an imperfectly-perceiving an sender. Moreover, whilst we can insist on the meaning generated by such a language as being 'correct' or 'ideal', doing so necessarily severs the relationship between actual use (by imperfectly perceiving senders) and meaning. If use is to determine meaning, then use cannot be conditioned upon information to which the sender lacks access.

Instead, we take the following approach: We retain the assumption that the sender imperfectly perceives the world, and instead ask how to give meaning to messages if they are to be interpreted as being about the actual state. Recall that, when the sender transmits message $m_k$, this has the unambiguous meaning in $Y$-space, that the sender's signal is contained in the interval $[s_{k-1}, s_k]$. The receiver is able to determine this meaning because she understands which signals are associated with which messages. Similarly, to determine meaning in $X$-space, the receiver must understand the mapping from states to message. Of course, we have just argued that the sender can only condition his message on his signal, and not the true state. Since a given state can generate multiple signals, a state may be associated with multiple messages. The mapping from states to messages need not be unique. This causes language to be vague.

We formalize this idea. Suppose the true state is $x$, and that the sender receives signal $y$ which is consistent with the signal technology $Q$. Given the above discussion, we know that the sender will transmit a unique message $\mu(y)$ which depends on the signal $y$. The communication strategy is not random or probabilistic. However, from the perspective of an observer who can perfectly perceive the state, but not the sender's signal, the sender's communication strategy may appear random — since the sender may send multiple different messages in the same state. (One way to conceive of such an observer is to suppose that the true state is revealed, *ex post*, and agents keep track of the frequency with which the sender transmits each message in each state.) Let $\phi_k(x)$ denote the probability that state $x$ is characterized as having property $m_k$. We know that:

$$
\begin{aligned}
\phi_k(x) &= \int_{\{y \in Y \,|\, \mu(y) = m_k\}} q(y|x)\, dy \\
&= \int_{s_{k-1}}^{s_k} q(y|x)\, dy
\end{aligned}
$$

The function $\phi_k(x)$ indicates the probability that state $x$ has property $m_k$. By the laws of probability, we know that $\sum_k \phi_k(m) = 1$. We can think of $\phi_k(x)$ as the membership function which determines the application of predicate $m_k$ to state $x$. Since use determines

meaning, these probabilities have the natural interpretation as the *degree of truth* that $x$ has property $m_k$. If the sender is more likely to associate state $x$ with property $m_k$ than $m_l$, then it is natural (in the sense of use determining meaning) to assign a higher degree of truth to $x$ having property $m_k$ than $m_l$. We stress that these truth degree functions are not primitives of the model (i.e. they are not taken as exogenous facts). Rather, they are determined in equilibrium by the sender's optimal $Y$-space language use and the signal technology which determines the sender's perception.

Before characterizing the properties of truth-degrees, we briefly digress to make the following observation: The process of extending the optimal $Y$-space language to $X$-space is analogous to the procedure that determines truth (or super-truth) under the supervaluationist approach. For each actual state $x$, we consider every signal $y$ that could potentially be observed by the sender (given the perception technology $Q$), and ask which message the sender would transmit, given his observed signal. The state $x$ is then *definitely* characterized by property $m_k$ if, under every possible signal-realization, the sender would transmit message $m_k$. By contrast, if different signal realizations result in different messages being transmitted, then the property associated with state $x$ is indefinite. We note that our model directly determines which extensions are admissible in generating the supervaluation, and that, in particular, admissibility is governed by the perception technology $Q$ and the optimal $Y$-space language. Our model also demonstrates a connection between the supervaluation and truth-degrees approaches to vagueness. Whereas the supervaluation approach enumerates the possible interpretations under admissible extensions, truth-degrees describe how likely these interpretations are.

Following the literature on fuzzy sets (see Zadeh (1975)), we define the *support* of message $m_k$, $S(m_k) = \{x \in X \,|\, \phi_k(m) > 0\}$, as the set of states that are associated with message $m_k$ with positive probability. Similarly, the *core* of message $m_k$, $C(m_k) = \{x \in X \,|\, \phi_k(m) = 1\}$, is the set of states that are definitely associated with message $m_k$. Naturally $C(m_k) \subset S(m_k)$. If $x \in C(m_k)$, then the sender's use will be definite and non-random in state $x$. If $x \in S(m_k)$ then the receiver understands the meaning of $m_k$ to convey the possibility that the true state is $x$. Language is clear (i.e. not vague) if every state is associated with a unique message. If so, the core and support must coincide for every message. By contrast, if language is vague, there must be some state which is associated with multiple messages, which implies that the supports of at least two messages must overlap. The set of states for which the supports overlap are precisely the 'borderline regions' that characterize vagueness.

We could, equivalently, formalize vagueness in the following way: Let $\mathcal{M} : X \mapsto M$ be the message correspondence, where $\mathcal{M}(x) \subset M$ is the set of possible messages that are associated

with state $x$. Language is clear if the correspondence is singleton valued — i.e. if $\mathcal{M}$ is a function. By contrast, language is vague if $\mathcal{M}(x)$ is set valued for some $x$ — i.e. if $\mathcal{M}$ may be many-to-many.[8] This characterization highlights the essential difference between clarity and precision. Language is precise if $\mathcal{M}^{-1}$ is a function — i.e. if each message is associated with exactly one state, so that transmitting a message exactly reveals the underlying state. By contrast, language is imprecise if $\mathcal{M}^{-1}$ is multi-valued. Language can be imprecise but clear if $\mathcal{M}$ is single-valued whilst $\mathcal{M}^{-1}$ is multi-valued. However, if language is vague (i.e. if both $\mathcal{M}$ and $\mathcal{M}^{-1}$ are multi-valued), then clearly it must also be imprecise.

For concreteness, we return to the example in the previous section, and characterize its extension to $X$-space. As before, the full characterization can be found in the appendix. Let $\varepsilon^* = \frac{3-\sqrt{3}}{8}$. The diagrams below show the truth-degree functions and the support and core sets for the equilibrium language in the above example. There are two cases to consider. First, suppose $\varepsilon < \varepsilon^*$ so that the 'margin for error' is relatively small. We have:



Case 1: $\epsilon < \epsilon^*$

---
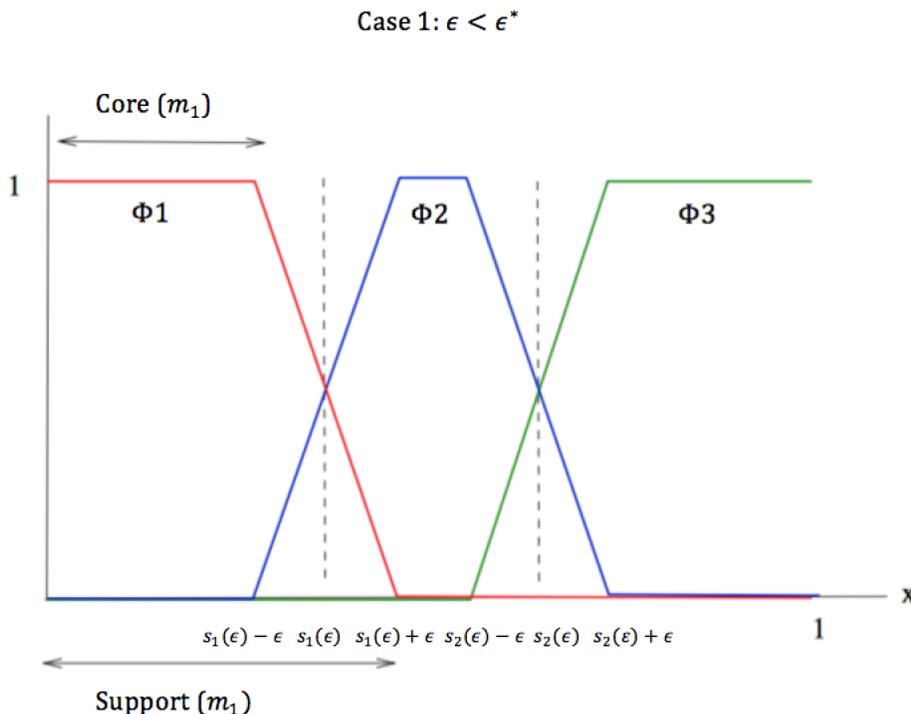
[8]Since there are more states than messages, it is impossible that $\mathcal{M}$ is one-to-many. However, if it were possible that $\mathcal{M}$ were one-to-many, then there is redundancy in language — each message is only associated with one state, but multiple messages can be associated with the same state. Such a language is clear, since every message reveals a unique underlying state.

The red line is the truth-degree function $\phi_1$ for message $m_1$. Similarly, the blue and green lines are the truth-degree functions $\phi_2$ and $\phi_3$ associated with messages $m_2$ and $m_3$, respectively. (The linearity of the truth-degree functions is an artifact of the state and signal both being drawn from a uniform distribution. These functions will generically be 'curved'.) The core and support sets are indicated for message 1 (and can be intuited for the remaining messages). The core of $m_1$ contains states that are sufficiently small, such that even if the sender receives an above-average signal, this signal will still be low enough that he is guaranteed to transmit $m_1$. Likewise, the core of $m_3$ contains states that are sufficiently large that the sender is guaranteed to receive a signal above threshold $s_2$, and so is guaranteed to transmit $m_3$. By contrast, there are a range of states for which the sender reports multiple messages with positive probability. Intuitively, such states will be 'close' to a threshold of the $Y$-space language (in this example, if it lies within $\varepsilon$ of a threshold), so that the induced signal will lie on one side of the threshold in some instances, and on the other side in other instances. In such regions, truth-degrees are positive for multiple messages. (Consistent with the theory, these truth-degrees must always sum to one, since it is certain the sender will transmit *some* message.) These regions cannot be contained in the core of any message — they are the 'borderline cases' that are characteristic of vagueness.

Next, suppose $\varepsilon \in \left( \varepsilon^*, \frac{3}{8} \right)$, so that the 'margin for error' is relatively large. The truth-degree functions can be represented as follows:

Case 2: $\epsilon^* < \epsilon < \frac{3}{8}$
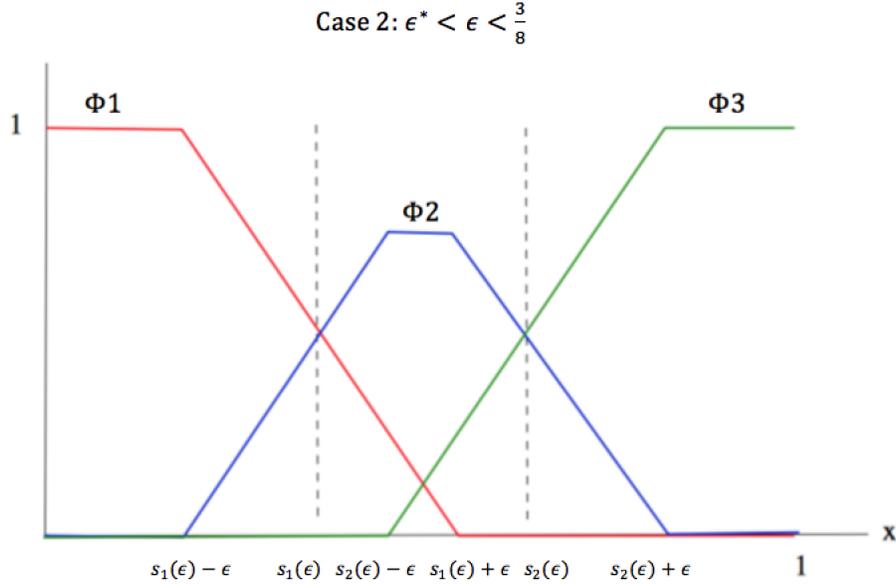
This case is distinguished from the previous case in two ways. Firstly, the core of $m_2$ is empty — there is no state for which we can be certain the sender will transmit $m_2$. The margin for error is sufficiently large that for any intermediate state, there is always the possibility that the sender may occasionally perceive it as being small or large. The core of $m_1$ and $m_3$ remain non-empty, although we note that these sets are smaller than in the previous case. By contrast, the supports of all messages are larger. As the margin for error increases, so does the range of states that may be associated with a given message. Second, there exists a range of states for which the sender reports all three messages with positive probability. The margin for error is sufficiently large that, when the true state takes an intermediate value, the sender will sometimes perceive it as being (sufficiently) small and other times perceives it as being (sufficiently) large.

Finally, as the example makes clear, vagueness is inherent to certain states and not to others. And, this is true even though the perception technology behaved identically across all states. It is not simply the case that language is vague in regions of the state space where the perception technology is particular noisy, and clear in regions where the perception technology is more precise. Instead, language becomes vague in regions of the state-space for which the generated signals will straddle the optimal $Y$-space thresholds. Whilst this does depend on the precision of the perception technology, it is also depends in a far more

21

basic sense on the state itself, and its 'location' relative to the threshold. As such, vagueness is necessarily *metaphysical* — it is inherent in the boundary cases themselves.

We now characterize the properties of truth-degree functions. We demonstrate that truth-degrees satisfy three properties. First, they are continuous. Second, they are monotone. Third, they are not truth-functional, but instead satisfy the axioms of probability.

**Lemma 1.** *Suppose $Q(y|x)$ is continuous in $x$ for every $y \in Y$. Then, for every $k = 1, ..., K$, the truth-degree function $\phi_k(x)$ is continuous in $x$.*

Loosely speaking, continuity is the property that small changes in the inputs of a function cannot cause dramatic changes in outputs. Continuity of the perception technology formalizes the intuitive assumption that small changes in the underlying state (what is being observed) should not dramatically change the sorts of signals that are generated. For example, if the sender systematically misperceives a person whose true height is 1.7 meters as being much shorter than he actually is, the sender should not then systematically misperceive a slightly taller person as being much taller than is actually the case.[9] Lemma 1 shows that a continuous perception technology causes truth-degrees to be continuous.

An important consequence of continuity is that truth-degree functions respect the desideratum that we treat similarly situated states similarly, *ex ante*. Given imperfect perception, the sender cannot easily distinguish between similarly situated states, and so we should not expect the probability which the observer assigns to the sender ascribing property $m_k$ to vary dramatically across those states. And yet this continuity seems to be in stark contrast to the threshold strategy that we derived in the previous section, which necessarily makes stark distinctions between similar objects. Notwithstanding the apparent contradiction, these features are perfectly consistent with one another. To see why, recall that messages are transmitted based on the signal received by the sender, whilst truth-degrees are assigned as if by an external agent who observes the true state, but not the signal. The observer understands that a small change in the signal can dramatically affect the probabilities of different messages being transmitted. However, the observer does not observe the signal, and can only form probabilistic beliefs about the likelihood of the signal being on one side of the threshold of the other. Continuity of the signal technology ensures that a small change in

---

[9]To be clear, we are not saying that whenever two states are similar, the sender will perceive them as being similar. As we have argued repeatedly, the *same* state may be perceived differently in different instances. Rather, our claim is that the nature of the sender's (mis)perception should be similar. In a given instance, it may be that the sender misperceives a person with height 1.7m as being taller than is the case, and misperceives a slightly taller person as being shorter than is the case. But it should not be that he is systematically more likely to perceive the former as being taller and the latter as being shorter.

the state generates only a small change in these probabilities. Hence, the external observer's beliefs will change in a gradual fashion, even though the sender's message choice may change starkly in any given instance.

Another consequence of continuity is that use and meaning of language cannot be characterized by firm thresholds in $X$-space. (If the language exhibited threshold behavior, it must be that at the threshold, the probability of transmitting a particular message increased *discontinuously* from 0 to 1, or vice versa.) Indeed, the fact that use and meaning are optimally characterized by threshold behavior in $Y$-space precludes the possibility that they respect thresholds in $X$-space, since the mapping between the two spaces is stochastic. Things which look determinate in $Y$-space must necessarily seem probabilistic in $X$-space. As we noted at the beginning of this (sub)-section, it is certainly possible to define a language over $X$-space that is characterized by thresholds — however, such a language cannot respect the requirement that use determines meaning, since the sender does not have access to the appropriate information to use this language in the required way. Accordingly, and in contrast to Williamson (1994), we demonstrate that when subject to imperfect perception, the optimal language cannot be characterized by distinct thresholds with respect to statements about actualities. Instead, we show that as the true state increases, there is a gradual and continuous transition regarding which messages are sent — what Williamson (1994) describes as a 'smear' — which renders the language vague.

A second property of truth-degree functions is that they are monotone in the ordering over the message-space. Recall, the set of messages was ordered so that $m_1$ indicated the lowest degree of the gradable adjective and $m_K$ denoted the highest degree. Monotonicity captures the idea that that higher ranked states will be more likely to be described using messages of higher (rather than lower) degree. To make this notion precise, let $\Phi_k(x) = \sum_{j=1}^{k} \phi_j(x)$ denote the truth-degree assigned to state $x$ having property $m_k$ or *lower*. It is easily verified that $\Phi_k(x) = Q(s_k|x)$. We have the following Lemma:

**Lemma 2.** *Suppose $x_0 < x_1$. Then for every $k$, $\Phi_k(x_0) \geq \Phi_l(x_1)$.*

Suppose building $A$ is (actually) taller than building $B$. Since the sender perceives the world imperfectly, it may be that in some instance, he categorizes $B$ as tall and $A$ as not. However, Lemma 2 demonstrates that this cannot be systematically true. Lemma 2 is a consequence of the monotone likelihood ratio property. This implied that when the true state is high, the sender must be more likely to receiver a higher signal, than when the true state is low. Whilst the sender's classification of objects may be imperfect, it must be statistically consistent with

the true grading of objects. Higher graded objects cannot on average be described by lower degrees of the adjective. The assignment of truth-degrees must accord naturally with the use of comparatives. Importantly, we note that monotonicity arises precisely because the mechanism that generates truth-degrees is the sender's uncertainty about the true state. If the source of uncertainty were otherwise, then vagueness may arise, but the truth-degree functions that represent this vagueness may not be monotone. For example, Lambie-Hanson and Parameswaran (2016) show that if the communicants have uncertain beliefs about each others' communication context, then there will be vagueness (of the semantic variety), but that truth-degree functions are constant over the boundary cases. Indeed, in that example, truth degrees are non-responsive to the state of the world, and depend instead on the communicants prior beliefs about each other. Hence, whilst monotonicity need not be property of truth-degree functions in the abstract, it does characterize truth-degree functions generated by the sender's imperfect perception.

We stress that the monotone property is with respect to cumulative truth-degree functions, rather than individual ones. To make this clear, return to the example of building $A$ which is taller than building $B$, and suppose the sender can describe these using one of three terms — short, medium and tall. Although 'medium' expresses a higher degree than 'short', it need not be that the sender is more likely to describe building $A$ as medium-heighted than building $B$. If building $A$ is a sky-scraper, he may be certain to describe it as 'tall', whilst he may well describe building $B$ as medium-heighted in some instances. However, it will be true that the sender is more likely to describe building $A$ as either 'medium heighted' or 'tall', than he is to describe building $B$ as such.

Finally, we note that, in strong contrast to most truth-degree proponents, our equilibrium truth-degrees are not truth-functional.[10] A simple example makes this clear. Suppose an object is equally likely to be ascribed each of the properties 'small', 'medium', 'large', and 'enormous', so that the truth-degree of each is $\frac{1}{4}$. Then the truth degrees associated with the ascriptions 'either small or medium-sized', 'either medium-sized or large', and 'either large or enormous' will each be $\frac{1}{2}$. (This follows from the axioms of probability, which the truth-degree functions obey, since they are probability measures by construction.) We have constructed compound statements using the disjunction, and thus far, the truth-degree of the disjunction appears to simply be the sum of the truth-degrees of the disjuncts. However, now consider the ascriptions 'either small, medium-sized or large' and 'either small, medium-sized, large or enormous'. The former is the disjunction of the first and second compound statements

---

[10]Truth-degrees are *truth-functional* if the truth degree of compound sentences can be determined directly from the truth degree of each component sentence.

above, whilst the latter is the disjunction of the first and third compound statements. If our truth degrees are truth functional, then the truth degrees of these final two sentences must be the same. However, by the laws of probability, the truth-degree of the former is $\frac{3}{4}$, whilst the truth-degree of the latter is 1. Clearly, we cannot universally construct the truth-degrees of compound statements, from the truth-degrees of the constituent statements. (In one of the cases above, we needed some additional information, namely the truth-degree of the conjunction.)

Truth functionality can be a valuable property in a world where truth-degree functions are taken as primitive. Absent truth-functionality, the truth-degree proponent must specify truth-degrees for every conceivable sentence that can be constructed, no matter how long or cumbersome. Such a burden is evidently onerous. Truth-functionality alleviates this need, by reducing all truth-degrees down to the truth-degrees of the underlying simple statements. However, the benefit of truth-functionality is less important in a world where truth-degrees are not primitive, but determined by other known features of the model — in our case, the sender's perception technology and the equilibrium communication strategy. This information (which we used to construct truth-degrees for simple statements in the first place) suffices to construct truth-degrees for any conceivable statement. Truth-functionality provides us no additional benefit. Indeed, since the truth-degree functions in our model are probability measures, we can use the laws of probability to map truth-degrees of simple statements onto truth-degrees of compound statements, and vice versa. The essence of truth-functionality is preserved. Of course, excepting for special cases, this mapping will not be truth functional, reflecting the idea that the joint distribution of random variables cannot generically be constructed from the marginal distributions alone.

Truth functionality is, of course, not without its own problems. For example, Fine (1975), Williamson (1994) and Edgington (1997) (amongst others, although see Smith (2008) for a defense), note that truth-functionality necessitates that truth-degrees violate standard results in classical logic, including the *Law of the Excluded Middle*. Our truth-degrees-as-probability-measures approach avoids these pitfalls, which provides additional support for this approach to measuring degrees of truth. Although truth functionality may generically be desirable, these benefits vanish in the presence of a well-defined probability measure that can consistently assign truth-degrees.

# 5  Conclusion

This paper examined imperfect perception as a source of vagueness. We developed a model of communication in which an imperfectly informed sender may transmit a message to an uninformed receiver, who must take an action that affects both parties. Both agents share identical concave preferences, which incentivizes efficient communication between the parties. To focus attention most cleanly on the effect of imperfect perception, we abstract from other features that may independently cause language to be vague. Although our framework is stylized, we are able to shed light on the several properties of vague language.

Our analysis begins with the recognition that, in a world with imperfect perception, we must distinguish two sorts of statements — subjective statements which convey what the sender perceives, and objective statements which convey what actually is. Since the sender only observes signals about the world — and not the actual world itself — he can only transmit subjective statements, and his messages ought to be interpreted as such. Nevertheless, we are often tempted to interpret messages as though they contained objective truths, notwithstanding our awareness that the sender may be mistaken. Our analysis also recognizes that a sender's optimal choice about which message to transmit will depend on his belief about how the receiver will interpret messages, and that the receiver will optimally interpret messages according to her expectation of when and how the sender transmits each available message. *Use* and *meaning* are, as such, jointly determined in equilibrium, given the agents' communication needs and the sender's perception technology.

We first characterize the equilibrium language when the communication is understood to be about what the agent perceives. We show that, notwithstanding the sender's imperfect information, communication in this world is non-vague and characterized by firm thresholds that demark the use of words. Intuitively, although the sender understands that his perception of the world may not be accurate, this does not prevent him from communicating what he has perceived. Moreover, since the receiver can rationally understand the sender's communication strategy, she can clearly infer what the sender has perceived. Hence, communication is not vague.

We then consider how to assign meaning to statements, if they are to be interpreted as being about actualities. (As we argue, the sender must continue to transmit messages based on what he perceives.) Given the sender's imperfect perception, he may in different instances ascribe different properties to the same state of the world. Although, the sender's use is determined in any given instance by his signal, his use may appear random or probabilistic to

26

an external agent (who observes the state but not the signal). We interpret these probabilities as truth-degrees, since they capture the likelihood that the sender will ascribe a particular property to a given state, *ex ante*. As such, we provided micro-foundations for truth-degrees as a consequence of the equilibrium language (about what is perceived) and the sender's perception technology, thus unifying two distinct theories of vagueness. Indeed, epistemic theory provided the mechanism that enabled us to generate the descriptive features of the truth-degrees approach.

A useful feature of our approach is that, since truth-degrees are determined (in equilibrium) rather than assumed, we can investigate the properties that truth-degree functions are likely to satisfy. We derived three features of truth-degree functions that were predicated upon standard assumptions about the perception technology. First, we showed that truth-degrees are continuous, capturing the natural idea that the senders will, on average, similarly describe similarly situated states. An important consequence of continuity is that language about the objective world cannot simultaneously satisfy the use-determines-meaning criterion and be characterized by thresholds. This result stands in strong contrast to epistemic theory, which insists that the underlying language is not inherently vague. Additionally, we showed that truth-degree functions are monotone, and therefore accord naturally with the use of comparatives. Finally, we demonstrated that our induced truth-degree functions were not truth functional. However, we argued that this was not problematic, since truth-degrees of compound statements could still be discerned from the truth-degrees of simpler statements, given the laws of probability. Hence, the essence of truth-functionality is preserved. Moreover, we showed that truth-degrees-as-probabilities avoid some of the more problematic features inherent to truth-functional truth-degrees, such as inconsistency with classical results in logic, such as the law of the excluded middle.

# 6 Appendix: Proofs

***Proof of Proposition 1.*** Let $\Sigma = \left\{ s \in Y^{K+1} | s = (s_0, ..., s_K) \text{ with } 0 = s_0 \leq ... \leq s_K = 1 \right\}$. Take some $s \in \Sigma$. Suppose the sender uses the strategy: $\mu(y) = m_k$ provided that $y \in (s_{k-1}, s_k]$. Since the belief functions follow from Bayes' Rule, we can easily verify that:

$$f(x|y) = \frac{f(x) q(y|x)}{\int_{z \in X} f(z) q(y|z) dz}$$

and:

$$
\begin{aligned}
g(x|m_k) &= \frac{f(x) \left[ \int_{y \in Y} \mathbf{1}_{[\mu(y)=m_k]} q(y|x) dy \right]}{\int_{z \in X} f(z) \left[ \int_{y \in Y} \mathbf{1}_{[\mu(y)=m_k]} q(y|z) dy \right] dz} \\
&= \frac{f(x) \left[ \int_{s_{k-1}}^{s_k} q(y|x) dy \right]}{\int_{z \in X} f(z) \left[ \int_{s_{k-1}}^{s_k} q(y|z) dy \right] dz}
\end{aligned}
$$

as required. Note that since $\int_{s_{k-1}}^{s_k} q(y|x) dy$ is continuous in the sender's communication strategy $s$, so is $g(x, m_k)$.

We next verify that the receiver's strategy is optimal, given her beliefs. The receiver's expected utility from choosing action $a \in X$ after receiving message $m_k$ is $-\int_{x \in X} (x - a)^2 g(x|m_k)$ Taking first-order conditions, we have:

$$
\begin{aligned}
-2 \int_{x \in X} (x - a) g(x|m_k) dx &= 0 \\
a_k^* &= \int_{x \in X} x g(x|m_k)
\end{aligned}
$$

since $\int g(x|m_k) dx = 1$. For each $k = 1, .., K$, let $A_k(s) = a_k^*$, and note that $A_k(s)$ is continuous in $s$. Furthermore $A_k(s) \in [s_{k-1}, s_k]$. Let $A(s) = (A_1(s), ..., A_K(s))$, and note the $A_1(s) \leq ... \leq A_K(s)$.

Finally, we must verify that the sender's strategy is optimal, given the receiver's action and the sender's beliefs. By construction, we have: $\mu(y) = \arg\max_{m_k \in M} - \int_{x \in X} (x - A_k(s))^2 f(x|y) dx$. Fix some $y \in Y$, and suppose $m_k$ is an op-

timal message. Then for all $k' < k$, we must have:

$$\int_{x \in X} (x - A_k(s))^2 f(x|y) \, dx \leq \int_{x \in X} (x - A_{k'}(s))^2 f(x|y) \, dx$$

$$\int x f(x|y) \, dx \geq \frac{1}{2} (A_k(s) + A_{k'}(s))$$

Since this must be true for every $k' < k$, and $A_1 \leq ... \leq A_K$, we have: $\int x f(x|y) \, dx \geq \frac{1}{2} (A_k(s) + A_{k-1}(s))$. By a similar argument, we must:

$$\int x f(x|y) \, dx \leq \frac{1}{2} (A_k(s) + A_{k+1}(s))$$

Let $\psi(y) = \int x f(x|y) \, dx$. We have $\frac{1}{2} (A_k(s) + A_{k-1}(s)) \leq \psi(y) \leq \frac{1}{2} (A_k(s) + A_{k+1}(s))$. Now, by the monotone likelihood ratio property, we know that $F[x|y_1]$ first-order stochastically dominates $F(x|y_1)$ whenever $y_1 > y_0$.[11] Hence:

$$\int x f(x|y_0) \, dx = [x F(x|y_0)]_0^1 - \int_0^1 F(x|y_0) \, dx$$

$$= 1 - \int_0^1 F(x|y_0) \, dx$$

$$< 1 - \int_0^1 F(x|y_1) \, dx$$

$$= \int x f(x|y_1) \, dx$$

and so $\psi(y)$ is a strictly increasing and continuous function. Hence it is optimal to report $m_k$ if $y \in \left[ \psi^{-1} \left( \frac{1}{2} (A_k(s) + A_{k-1}(s)) \right), \psi^{-1} \left( \frac{1}{2} (A_k(s) + A_{k+1}(s)) \right) \right]$ where the inequality follows since $F(x|y_0) \geq F(x|y_1)$ for all $x$ with strict inequality for some $x$. Then let $S_k(s) = \psi^{-1} \left( \frac{1}{2} (A_k(s) + A_{k+1}(s)) \right)$ for $k = 2, ..., K$, and let $S_0(s) = 0$ and $S_{K+1}(s) = 1$. Note that $S_k$ is continuous in $s$. We have message $m_k$ is optimal if $y \in [S_{k-1}(s), S_k(s)]$. Let $S(s) = (S_0(s), .., S_K(s))$.

Finally, $s$ is an equilibrium threshold strategy if $s = S(s)$ — i.e. if $s$ is a fixed point of the mapping $S : \Sigma \to \Sigma$. Since this mapping is continuous over a compact space, Brouwer's fixed point theorem ensures that it admits a fixed point. We must show that $s_{k-1} < s_k$ in equilibrium. Suppose not. I.e. suppose there is a fixed point for which $s_{k-1} = s_k$. Then

---

[11] We show this explicitly in the proof of Lemma 2, below.

$A_k(s) = \psi(s_k)$. Next note that

$$\int_{x \in X} (x-a)^2 f(x|y)\, dx = (\psi(y) - a)^2 + \int (x - \psi(y))^2 f(x|y)\, dx$$

, which is simply the usual mean-square error decomposition, since $\psi(y) = E_{X|Y=y}[X]$. For message $k$ with $a = A_k(s)$, this expression is minimized when $y = s_k$. Moreover, for any $k'$ s.t. $A_{k'}(s) \neq A_k(s)$, this expression is strictly larger when $\mu = m_{k'}$. Then by continuity $m_k$ must be optimal for $y \in (s_{k-1}, s_{k-1} + \varepsilon)$ where $\varepsilon > 0$ is small enough. Hence $s_k > s_{k-1}$. $\quad\square$

***Example in Detail.*** Suppose $x \sim U[0,1]$ and let the perception technology generate a signal $y \sim U[x - \varepsilon, x + \varepsilon]$, where $\varepsilon < \frac{3}{8}$ denotes the signal precision. Note this implies $Y \in [-\varepsilon, 1 + \varepsilon]$. This implies the following posterior beliefs for the sender, derived according to Bayes Rule. If $y < \varepsilon$, then

$$f(x|y) \;=\; \begin{cases} \frac{1}{y+\varepsilon} & x \in [0, y+\varepsilon] \\ 0 & \text{otherwise} \end{cases}$$

If instead $y \in [\varepsilon, 1 - \varepsilon]$, then:

$$f(x|y) = \begin{cases} \frac{1}{2\varepsilon} & x \in [y - \varepsilon, y + \varepsilon] \\ 0 & \text{otherwise} \end{cases}$$

and finally if $y > 1 - \varepsilon$

$$f(x|y) = \begin{cases} \frac{1}{1+\varepsilon-y} & x \in [y - \varepsilon, 1] \\ 0 & \text{otherwise} \end{cases}$$

Suppose $K = 3$, so that 3 messages are sent. Let $a_i$ be the action chosen by the receiver following the receipt of message $m_i$. Additionally, let $s_1$ and $s_2$ be the thresholds that delineate the use of these messages. Since the problem is perfectly symmetric, we know that $s_2 = 1 - s_1$, that $a_3 = 1 - a_1$ and $a_2 = \frac{1}{2}$. Hence, it suffices to focus on locating $s_1$ and $a_1$. To do so, we must determine the receiver's posterior beliefs. We have:

$$g(x|m_1) \;=\; \frac{\int_{-\varepsilon}^{s_1} q(y|x)\, dy}{\int_{z \in X} \left[ \int_{-\varepsilon}^{s_1} q(y|z)\, dy \right] dz}$$

Now, note that: $\int_{-\varepsilon}^{s_1} q(y|x)\, dy = \frac{\min\{2\varepsilon, s_1 + \varepsilon - x\}}{2\varepsilon}$, and this quantity is 1 provided $x \leq s_1 - \varepsilon$. Moreover, this quantity is 0 whenever $x > s_1 + \varepsilon$. Assume that $s_1 > \varepsilon$. ( Since $\varepsilon < \frac{3}{8}$ and

30

$s_1 < \frac{1}{2}$, we know that $s_1 + \varepsilon < 1$.) This implies that:

$$\int_{z \in X} \left[ \int_{-\varepsilon}^{s_1} q\left(y|z\right) dy \right] dz = \int_0^{s_1 - \varepsilon} dx + \int_{s_1 - \varepsilon}^{s_1 + \varepsilon} \left( \frac{s_1 + \varepsilon - x}{2\varepsilon} \right) dx$$
$$= s_1$$

and so, we have:

$$g\left(x|m_1\right) = \begin{cases} \frac{1}{s_1} & x \in [0, s_1 - \varepsilon] \\ \frac{s_1 + \varepsilon - x}{2\varepsilon s_1} & x \in [s_1 - \varepsilon, s_1 + \varepsilon] \\ 0 & x \in [s_1 + \varepsilon, 1] \end{cases}$$

We can construct $g\left(x|m_2\right)$ and $g\left(x|m_3\right)$ similarly.

Next, we evaluate the receiver's optimal action after receiving message $m_1$. We know that the receiver chooses the action which corresponds to the expected true state conditional upon the message received. We have:

$$\begin{aligned} a_1 &= \int_X x g\left(x|m_1\right) dx \\ &= \int_0^{s_1 - \varepsilon} \frac{x}{s_1} dx + \int_{s_1 - \varepsilon}^{s_1 + \varepsilon} x \frac{s_1 + \varepsilon - x}{2\varepsilon s_1} dx \\ &= \frac{3s_1^2 + \varepsilon^2}{6s_1} \end{aligned}$$

Next, let $\psi\left(y\right) = \int x f\left(x|y\right) dx$. Take $y \in [\varepsilon, 1 - \varepsilon]$. Then we know that $\psi\left(y\right) = \int_{y - \varepsilon}^{y + \varepsilon} \frac{x}{2\varepsilon} dx = y$. Moreover, we know that $s_1$ satisfies $\psi\left(s_1\right) = \frac{1}{2}\left(a_1 + a_2\right)$, which implies:

$$\begin{aligned} s_1 &= \frac{1}{2} \left[ \frac{3s_1^2 + \varepsilon^2}{6s_1} + \frac{1}{2} \right] \\ s_1 &= \frac{1 + \sqrt{1 + 4\varepsilon^2}}{6} \end{aligned}$$

It is easily confirmed that $s_1 > \varepsilon$ provided that $\varepsilon < \frac{3}{8}$. In summary, we have: $\left(s_1^*, s_2^*\right) = \left( \frac{1 + \sqrt{1 + 4\varepsilon^2}}{6}, \frac{5 - \sqrt{1 + 4\varepsilon^2}}{6} \right)$ and $\left(a_1^*, a_2^*, a_3^*\right) = \left( \frac{\sqrt{1 + 4\varepsilon^2}}{3} - \frac{1}{6}, \frac{1}{2}, \frac{7}{6} - \frac{\sqrt{1 + 4\varepsilon^2}}{3} \right)$.

Finally, we characterize the core and support of each message, and the associated truth-

degree functions. The support sets are:

$$
\begin{aligned}
S_R\left(m_1\right) &= \left[0, \frac{1+6\varepsilon+\sqrt{1+4\varepsilon^2}}{6}\right] \\
S_R\left(m_2\right) &= \left[\frac{1-6\varepsilon+\sqrt{1+4\varepsilon^2}}{6}, \frac{5+6\varepsilon-\sqrt{1+4\varepsilon^2}}{6}\right] \\
S_R\left(m_3\right) &= \left[\frac{5-6\varepsilon-\sqrt{1+4\varepsilon^2}}{6}, 1\right]
\end{aligned}
$$

The core sets are:

$$
\begin{aligned}
C_R\left(m_1\right) &= \left[0, \frac{1-6\varepsilon+\sqrt{1+4\varepsilon^2}}{6}\right] \\
C_R\left(m_3\right) &= \left[\frac{5+6\varepsilon-\sqrt{1+4\varepsilon^2}}{6}, 1\right]
\end{aligned}
$$

If $\varepsilon > \frac{3-\sqrt{3}}{8}$ , then $C_R\left(m_2\right) = \emptyset$; else $C_R\left(m_2\right) = \left[\frac{1+6\varepsilon+\sqrt{1+4\varepsilon^2}}{6}, \frac{5-6\varepsilon-\sqrt{1+4\varepsilon^2}}{6}\right]$. Finally, the truth-degree functions are generally given by $\phi_i\left(x\right) = \int_{s_{k-1}}^{s_k} q\left(y|x\right) dx$. Hence, we have:

$$
\phi_1\left(x\right) = \begin{cases}
1 & x \in \left[0, \frac{1-6\varepsilon+\sqrt{1+4\varepsilon^2}}{6}\right] \\
\frac{1+6\varepsilon+\sqrt{1+4\varepsilon^2}-6x}{12\varepsilon} & x \in \left[\frac{1-6\varepsilon+\sqrt{1+4\varepsilon^2}}{6}, \frac{1+6\varepsilon+\sqrt{1+4\varepsilon^2}}{6}\right] \\
0 & x \in \left[\frac{1+6\varepsilon+\sqrt{1+4\varepsilon^2}}{6}, 1\right]
\end{cases}
$$

$$
\phi_3\left(x\right) = \begin{cases}
0 & x \in \left[0, \frac{5-6\varepsilon-\sqrt{1+4\varepsilon^2}}{6}\right] \\
\frac{6x-\left(5-6\varepsilon-\sqrt{1+4\varepsilon^2}\right)}{12\varepsilon} & x \in \left[\frac{5-6\varepsilon-\sqrt{1+4\varepsilon^2}}{6}, \frac{5+6\varepsilon-\sqrt{1+4\varepsilon^2}}{6}\right] \\
1 & x \in \left[\frac{5+6\varepsilon-\sqrt{1+4\varepsilon^2}}{6}, 1\right]
\end{cases}
$$

If $\varepsilon < \frac{3-\sqrt{3}}{8}$, then we have:

$$
\phi_2\left(x\right) = \begin{cases}
0 & x \in \left[0, \frac{1-6\varepsilon+\sqrt{1+4\varepsilon^2}}{6}\right] \\
\frac{6x-\left(1-6\varepsilon-\sqrt{1+4\varepsilon^2}\right)}{12\varepsilon} & x \in \left[\frac{1-6\varepsilon+\sqrt{1+4\varepsilon^2}}{6}, \frac{1+6\varepsilon+\sqrt{1+4\varepsilon^2}}{6}\right] \\
1 & x \in \left[\frac{1+6\varepsilon+\sqrt{1+4\varepsilon^2}}{6}, \frac{5-6\varepsilon-\sqrt{1+4\varepsilon^2}}{6}\right] \\
\frac{\left(5+6\varepsilon-\sqrt{1+4\varepsilon^2}\right)-6x}{12\varepsilon} & x \in \left[\frac{5-6\varepsilon-\sqrt{1+4\varepsilon^2}}{6}, \frac{5+6\varepsilon-\sqrt{1+4\varepsilon^2}}{6}\right] \\
0 & x \in \left[\frac{5+6\varepsilon-\sqrt{1+4\varepsilon^2}}{6}, 1\right]
\end{cases}
$$

By contrast, if $\varepsilon \in \left[\frac{3-\sqrt{3}}{8}, \frac{3}{8}\right]$, then we have:

$$\phi_2(x) = \begin{cases} 0 & x \in \left[0, \frac{1-6\varepsilon+\sqrt{1+4\varepsilon^2}}{6}\right] \\ \frac{6x-\left(1-6\varepsilon-\sqrt{1+4\varepsilon^2}\right)}{12\varepsilon} & x \in \left[\frac{1-6\varepsilon+\sqrt{1+4\varepsilon^2}}{6}, \frac{5-6\varepsilon-\sqrt{1+4\varepsilon^2}}{6}\right] \\ \frac{3-\sqrt{1+4\varepsilon^2}}{6\varepsilon} & x \in \left[\frac{5-6\varepsilon-\sqrt{1+4\varepsilon^2}}{6}, \frac{1+6\varepsilon+\sqrt{1+4\varepsilon^2}}{6}\right] \\ \frac{\left(5+6\varepsilon-\sqrt{1+4\varepsilon^2}\right)-6x}{12\varepsilon} & x \in \left[\frac{1+6\varepsilon+\sqrt{1+4\varepsilon^2}}{6}, \frac{5+6\varepsilon-\sqrt{1+4\varepsilon^2}}{6}\right] \\ 0 & x \in \left[\frac{5+6\varepsilon-\sqrt{1+4\varepsilon^2}}{6}, 1\right] \end{cases}$$

$\square$

**_Proof of Lemma 1._** Follows immediately from the continuity of $Q(y|x)$ in $x$. Recall that $\phi_k(x) = \int_{s_{k-1}}^{s_k} q(y|x)\, dy = Q(s_k|x) - Q(s_{k-1}|x)$. Then since $Q(\cdot|x)$ is continuous in $x$, so is $\phi_k$.

$\square$

**_Proof of Lemma 2._** Follows as a well known consequence of the monotone likelihood ratio property. To see this, first note that $\Phi_k(x) = \sum_{j=1}^{k} \phi_k(x) = \sum_{j=1}^{k} \int_{s_{j-1}}^{s_j} q(y|x)\, dy = Q(s_k|x)$. Take $x_1 > x_0$. It suffices to show that $Q(s_k|x_1) \le Q(s_k|x_0)$ for every $k$.

By the monotone likelihood ratio property, we know that $\frac{q(y_0|x_1)}{q(y_0|x_0)} \le \frac{q(y_1|x_1)}{q(y_1|x_0)}$ whenever $y_1 > y_0$, which we can rewrite as: $q(y_0|x_1)\, q(y_1|x_0) \le q(y_1|x_1)\, q(y_0|x_0)$. This implies:

$$\begin{aligned} \int_0^{y_1} q(y_0|x_1)\, q(y_1|x_0)\, dy_0 &\le \int_0^{y_1} q(y_1|x_1)\, q(y_0|x_0)\, dy_0 \\ Q(y_1|x_1)\, q(y_1|x_0) &\le Q(y_1|x_0)\, q(y_1|x_1) \\ \frac{Q(y|x_1)}{Q(y|x_0)} &\le \frac{q(y|x_1)}{q(y|x_0)} \end{aligned}$$

Similarly, we have:

$$\begin{aligned} \int_{y_0}^1 q(y_0|x_1)\, q(y_1|x_0)\, dy_1 &\le \int_{y_0}^1 q(y_1|x_1)\, q(y_0|x_0)\, dy_1 \\ q(y_0|x_1)\left[1 - Q(y_0|x_0)\right] &\le q(y_0|x_0)\left[1 - Q(y_0|x_1)\right] \\ \frac{q(y|x_1)}{q(y|x_0)} &\le \frac{1 - Q(y|x_1)}{1 - Q(y|x_0)} \end{aligned}$$

Combining these gives:

$$\frac{Q(y|x_1)}{Q(y|x_0)} \le \frac{1 - Q(y|x_1)}{1 - Q(y|x_0)}$$

33

for every $y$. This implies that $Q\left(y|x_1\right) \leq Q\left(y|x_0\right)$ for every $y$, which completes the proof. $\square$

# References

**Blume, Andreas and Oliver Board**, "Language barriers," *Econometrica*, 2013, *81* (2), 781–812.

_ **and** _ , "Intentional vagueness," *Erkenntnis*, 2014, *79* (4), 855–899.

_ **and Oliver J Board**, "Higher-Order Uncertainty About Language," 2013.

**Chen, Ying**, "Perturbed communication games with honest senders and naive receivers," *Journal of Economic Theory*, 2011, *146* (2), 401–424.

**Crawford, Vincent P and Joel Sobel**, "Strategic information transmission," *Econometrica: Journal of the Econometric Society*, 1982, pp. 1431–1451.

**David, LEWIS**, *Convention: a philosophical study*, Cambridge, Harvard university press, 1969.

**Dummett, Michael**, "Wang's paradox," *Synthese*, 1975, *30* (3-4), 301–324.

**Edgington, Barbara**, "Truth by Degrees," in Rosanna Keefe and Peter Smith, eds., *Vagueness: A Reader*, Bradford, 1997.

**Fine, Kit**, "Vagueness, truth and logic," *Synthese*, 1975, *30* (3), 265–300.

**Grossman, Sanford J**, "The informational role of warranties and private disclosure about product quality," *The Journal of Law & Economics*, 1981, *24* (3), 461–483.

**Halldén, Sören**, "The logic of nonsense," 1949.

**Keefe, Rosanna**, *Theories of vagueness*, Cambridge University Press, 2000.

**Körner, Stephan**, "The Philosophy of Mathematics, an Introduction," 1962.

**Lambie-Hanson, Timothy and Giri Parameswaran**, "A Model of Vague Communication," 2016.

**Lassiter, Daniel and Noah D Goodman**, "Context, scale structure, and statistics in the interpretation of positive-form adjectives," in "Semantics and Linguistic Theory" 2014, pp. 587–610.

**Lipman, Barton L**, "Language and economics," in N. Dimitri, M. Basili, and I. Gilboa, eds., *Cognitive Processes and Economic Behaviour*, Routledge, 2003.

_ , "Why is language vague?," *Boston University*, 2009.

**Mialon, Hugo M and Sue H Mialon**, "Go figure: The strategy of nonliteral speech," *American Economic Journal: Microeconomics*, 2013, *5* (2), 186–212.

**Milgrom, Paul R**, "Good news and bad news: Representation theorems and applications," *The Bell Journal of Economics*, 1981, pp. 380–391.

**Qing, Ciyang and Michael Franke**, "Gradable adjectives, vagueness, and optimal language use: A speaker-oriented model," in "Semantics and Linguistic Theory," Vol. 24 2014, pp. 23–41.

**Rick, Armin**, "The Benefits of Miscommunication," 2013.

**Selten, Reinhard**, "Reexamination of the perfectness concept for equilibrium points in extensive games," *International journal of game theory*, 1975, *4* (1), 25–55.

**Smith, Nicholas JJ**, *Vagueness and degrees of truth*, Oxford University Press, 2008.

**Sorensen, Roy**, *Vagueness and contradiction*, Oxford University Press, 2001.

**Williamson, Timothy**, *Vagueness*, Psychology Press, 1994.

_ , *Vagueness*, Routledge, 2002.

**Zadeh, Lotfi A**, "Fuzzy logic and approximate reasoning," *Synthese*, 1975, *30* (3-4), 407–428.