Natural sources of vagueness and their implications

José Pedro Correia Michael Franke

Abstract

A vexing puzzle about vagueness, rationality and evolution runs, in crude abbreviation, as follows: vague language use is demonstrably suboptimal if the goal is efficient precise and cooperative information transmission; hence rational deliberation or evolutionary selection should, under this assumed goal, eradicate vagueness from language use. Since vagueness is persistent in all human languages, something has to give. In this paper, we investigate a number of reasons why and mechanisms how vagueness may come into the picture in formal models of rational or evolutionary optimal signaling. We show how uncertainty about not only the linguistic practices of others, but also about the world itself can lead to vagueness. We explore the consequences of these reasons and mechanisms for a notion of meaning and a notion of rationality.

Contents

1 Vagueness and the classical picture

The classical philosophical problem of vagueness is most starkly embodied by the sorites paradox. The original formulation is attributed to Eubulides, an ancient Megarian philosopher (Sorensen, 2009), and uses the example of a heap of sand: if no removal of one grain of sand can make a heap into a non-heap, we can repeatedly remove all but one grain of sand of something that is clearly a heap and be forced to acknowledge that the remaining single grain of sand is still a heap; otherwise, it seems, we would have to accept that there is a determinate number of grains that forms a heap, and anything under it is not a heap. Neither choice is, however, intuitively satisfying. The paradox is interesting because it can be made general and re-applied to many other words besides 'heap'. Predicates for which one can find a suitable instance of the general formulation of the sorites paradox are called *vague*. Paradigmatic examples besides 'heap' include 'tall', 'red', 'bald', 'tadpole', and 'child' (Keefe and Smith, 1997). How widespread is the problem? It is easy to find more examples of predicates based on more finely grained properties—as 'tall' is intuitively based on height—for which constructing a sorites paradox would be easy. Mereological nihilists argue that instances of the sorites can be designed for any material object that can be decomposed into small enough parts. If we subscribe to the scientific picture of matter as composed of molecules and atoms, this applies to tables and chairs, cats and mats, and any other ordinary thing (Unger, 1979). Bertrand Russell famously argued (1923) that all words, including "the words of pure logic", are vague when used by human beings. The problem thus seems to be a wide-ranging one, but it does not obviously undermine our everyday linguistic practices. Most of the time we seem to communicate just fine. Therefore, the issue must lie with the conceptions we have of the forces or mechanisms that underlie those practices.

Vagueness is typically seen as a challenge to a classical conception of language that relies on reference and truth as two core notions for understanding meaning. One problem is that vague predicates seem to lack precise boundaries; if that is the case, how can words like 'tall' stand in correspondence to something like a well defined set of tall people? And if we cannot know exactly whether a certain person is 'tall' or not, as with borderline cases, how are we to determine the truth value of sentences that involve statements of tallness regarding these borderline cases? Supervaluationism, many-valued logics, and degree theories all propose changes to the classical picture in order to accommodate for vague predicates. They are, however, still strongly committed to its two core notions. As argued by Mark Sainsbury (1999), because of that they fail to address an important characteristic of vague predicates: *higher order vagueness*. All the aforementioned proposals end up being committed to new artificial demarcating boundaries (*e.g.* true-under-all-precisifications versus neither true nor false versus false-under-all-precisifications, true versus indefinite versus false, true to degree 1 versus true to degree 0 versus the rest). But a vague predicate not only fails to demarcate between the cases where it clearly applies and the ones where it clearly doesn't, it also fails to establish a boundary between the cases where it clearly applies and the borderline cases, as well as between the borderline cases and the cases where it clearly doesn't apply. Further introducing borderline borderline cases would lead to an infinite regress. Because of their attachment to the classical picture, the standard approaches to vagueness also fail to draw an important lesson from the sorites paradox, namely that "we do not know, cannot know, and do not need to know these supposed boundaries to use language correctly" (1999, p. 256). By trying to cling as much as possible to the classical picture of logic and semantics, these standard approaches are ignoring a simple observation: natural language users are sensitive to the sorites paradox, *i.e.* they are able to recognize the logical inconsistency but do not have a good answer as to how to overcome it. Even more importantly, they apparently do not need to solve the inconsistency in order to continue using natural language to achieve their daily goals. Nobody ever stopped using the word 'tall' after being confronted with a sorites series. Why should we develop theories of meaning that are impervious to the paradox?

The reluctance to give up truth and logic as valuable notions to explain meaning is perhaps associated with the fear of what would also need to be abandoned as a consequence. One notion that seems to quickly be in peril is that of rationality. In reference to philosophers who defend the desirability of a classical notion of truth, Richard Rorty says:

> In the past, such philosophers have typically conjoined the claim that there is universal human agreement on the supreme desirability of truth with two further premises: that truth is correspondence to reality, and that reality has an intrinsic nature (that there is, in Nelson Goodman's terms, a Way the World Is). [...] The rise of relatively democratic, relatively tolerant, societies in the last few hundred years is said to be due to the increased rationality of modern times, where 'rationality' denotes the employment of an innate truth-oriented faculty. (2000, p. 1)

By tying all of these notions tightly together, the fear could be that by dropping the picture of meaning as intimately tied to truth and logic we necessarily lose the ground on which rationality stands, and the whole edifice could collapse. Rationality, in this picture, has truth as its guiding light and logic as the means to attain it. Being rational is about following universally valid rules of reasoning that, given the correct inputs, necessarily lead us to the correct conclusions. Although humans are not necessarily consciously aware of the rules, they are considered to be nevertheless unconsciously following them when they act rationally. This type of rationality is one that supposedly demarcates humans from other animals.

We want to argue here that to say that language does not follow strict logical rules, or that truth is not a useful notion to guide our inquiries into meaning, is not to say that language is unstructured, meaningless, or unusable; neither is it to say that we, language users, are therefore irrational. Giving up the ideal of truth and logic as relevant explanatory notions for understanding meaning in natural language does not imply giving up on rationality altogether. It also does not necessarily eradicate all hope for a (formal) explication of meaning. In order to make this argument, we need a different paradigm of language and meaning. But what could such a paradigm even be?

We suggest that inspiration can be drawn from the later work of philosopher Ludwig Wittgenstein[1]. His *Philosophical Investigations* (1953/2009) adumbrate a characterization of language that departs from the classical picture by highlighting its diversity, heterogeneity, and dynamism. Language is seen as a patchwork of various language-games, with new ones continuously being added and old ones falling out of use. These language-games, in turn, can be thought of as the use of signs in the context of and intermingled with a practice. There is no meaning in an abstract atemporal world, it is rather created dynamically and in a way that is contingent to the practices we engage in, to the language-games we play. The notion of language-game is also used to characterize one of Wittgenstein's methodological tools. Following the idea that "[i]t disperses the fog if we study the phenomena of language in primitive kinds of use in which one can clearly survey the purpose and functioning of the words" (1953/2009, §5), the method is to set up a language-game as a hypothetical scenario, or thought experiment, where language is used in a certain type of activity, and then reflect on the assumptions that underlie our interpretation of this set-up, as well as on how the scenario would play out according to those assumptions. It has been argued (Correia, 2013) that the framework of *signaling games*, introduced to the study of meaning by David Lewis (1969/2011) and later naturalized by Brian Skyrms (1996), permits us to do exactly that while reaping the benefits of a formalism that enables the in-depth exploration of the implications of complex hypothetical language-games via mathematical analysis or computer simulation. From this perspective, creating a signaling game model is like setting up a language-game, only using a formulation that improves perspicuity, and conducting mathematical analysis or computer simulations is like contemplating how the game would play out according to the assumptions built into the model.

Following this idea, we here survey proposals for addressing vagueness in the context of signaling games and reflect on their implications for how one thinks of rationality in the context of language use. Before we start, we want to establish some vocabulary that will be useful later on. Rationality is an elusive notion which is frequently debated in areas like philosophy, psychology, and economics. Discussions around it, although touching on different aspects of the concept, often fail to clearly demarcate them. We can start by considering what is deemed to be most important when characterizing the rationality of a given choice. The aforementioned classical picture of language and meaning is focused on dichotomies like true and false, meaningful and meaningless, correct and incorrect. A sentence can be true or false only if it is meaningful, and it is meaningful if it is constructed according to correct rules. Rationality is intimately connected with the ability to follow good procedures, not only of sentence production, but ultimately of sentence combination and reasoning (think of what is required for making a logically valid deductive inference). A *procedural* account of rationality thus focuses on the means, rather than the ends. One can also do the opposite and focus on the consequences instead. *Instrumental* rationality, as it is typically called, characterizes an agent's choice as rational if it maximizes the possibility of achieving a desired goal, regardless of the means. The notion is linked to David Hume (1738/2005), epitomized in his assertion that "Reason is, and ought to only be the slave of the passions." Instrumental rationality is close to a notion of rational choice that is used in economics and game theory[2]: agents are rational if and only if they make decisions that maximize their expected utility. A more in-depth discussion of the opposition in the context of theoretical economics can be found, for example, in the work of Herbert A. Simon (*e.g.* 1986)[3].

[1]To the extent that substantial views can be said to be defended by the author. Skipping over the debate (but see Kahane, Kanterian, and Kuusela, 2007/2013 for more details), we are here assuming an interpretation of Wittgenstein as a kind of pragmatist, along the lines of the readings of Hilary Putnam (1994) and Richard Rorty (2007).

[2]In fact, it has been argued (Vanderschraaf, 1998) that Hume's whole account of convention is very closely in line with modern game theory.

[3]Simon uses the term substantive instead of instrumental rationality, but the characterization is basically the

When developing models where rationality is relevant, be it constructing a logical system or setting up a formal game, the assumed epistemic relation between agents and their environment can come to bear on considerations of rationality. Depending on how accurate and complete an agent's knowledge of its environment, the goals to be achieved, the choices or rules available, the relation between those choices and the objectives, etc, the verdict over the rationality of certain choices can vary. An *omniscient* agent is one that is in possession of the same information as the modeler, whereas an agent with less than that is said to only have *limited* awareness of the relevant aspects of the model. Models working within the classical picture typically do not make a distinction between modeler and agent, thus there is no room to express any epistemic gaps. It follows that agents can only be assumed to be omniscient of the details involved. By abstracting away from language users, these models cannot represent potential interactions between them, let alone repeated interaction and language change. In other types of models, however, a further aspect of rational belief formation needs to be considered, namely the ability of agents to make accurate predictions about how other agents behave. We can say that an agent is more or less *strategic* depending on the extent to which he is able to anticipate the actions and beliefs of other agents, and to predict medium/long term gains from repeated play. Lack of perfect knowledge of the situation or lack of strategicness can be caused by many possible factors. These include, but are not limited to, limitations in handling information (receiving, storing, retrieving, transmitting) and limited computational resources to solve complex problems. We can talk about *bounded* rationality to characterize the choices of such agents.

Where the above definition of instrumental rationality is usually understood as requiring a single choice of act to maximize the expected utility in a single concrete decision situation, we might also be interested in more general choice mechanisms (Fawcett, Hamblin, and Giraldeaub, 2013; Galeazzi and Franke, to appear; Hagen et al., 2012; Zollman and Smead, 2010, *e.g.*). A choice mechanism is a general way of behaving for an agent involved in a variety of decision situations. When considering only a single situation, rationality can only be *local*. If we take into account the possibility of the agent's choice or choice mechanism hinging on multiple situations, we can also talk about *global* or *ecological* rationality. This can be important because, hypothetically, there could be choice mechanisms that are sub-optimal at a local level (for each situation) but are actually perfectly rational at a global level.

Note that we consider that, in theory, most combinations of these aspects are possible. Although we have been pinning logic to the classical picture, this is only with the most traditional logical systems in mind. We are not denying that advances in dynamic, epistemic, fuzzy, paraconsistent, and other types of logic could potentially enable one to capture procedural rationality with different characteristics. Game-theoretical models of language can, on the other hand, also assume local instrumental rationality with omniscient highly strategic agents. Our objective here is not to survey all the possibilities. We want to focus on signaling games as a framework for the study of language use and meaning. The vocabulary just introduced will, we hope, help inform the discussion that follows. We proceed by introducing the framework of signaling games in Section 2. In Section 3 we look into explanations of vagueness in a particular kind of signaling game. Section 4 reflects on the implications of these explanations for considerations of rationality. Section 5 generalizes our discussion to the natural case where agents play more than one game.

2 Signaling games

Signaling games were first introduced as models of communication by David Lewis (1969/2011). In order to support the idea that linguistic conventions can arise without any prior conventional activity, Lewis considers situations where agents' choices involve sending and receiving signals or messages. Thus, we could think of two players with different roles. The first player, the sender, has knowledge about which of a number of possible states of affairs obtains and, depending on this information, chooses a signal to send. The second player, the receiver, has knowledge about which signal the sender chose and, based on this information, chooses one of several possible responses. A

same (Simon, 1986, pp. 210-212).

preference relation exists between responses and states of affairs, and a payoff is attributed to each player based on the choices of both. Note that Lewis assumes that no player has any preference regarding the particular signal that is used, provided that it enables coordination. Formally, in order to describe the setup all we need is to specify a set of possible states of affairs T, a probability measure P such that $P(t)$ is the probability of frequency with which $t \in T$ occurs, a set of available signals or messages M, a set of responses or actions A, and a pair of utility functions $U_{S,R} : T \times A \to \mathbb{R}$, one for the sender and one for the receiver, each of which yields a payoff value for each possible pairing of state and action. These so-called signaling problems can be seen as particular cases of coordination problems if we consider the players' choices to be of contingency plans or strategies. A sender strategy is a specification of a choice of message for each possible state of affairs. It thus describes the sender's behavior conditional on the state of affairs that obtains. A receiver strategy analogously specifies a choice of action for each possible message. Thus, formally, what the sender chooses is a function $\sigma : T \to M$ and the receiver a function $\rho : M \to A$. The expected utility EU of a strategy can be calculated using the utility function and aggregating payoffs for all pairings of states of affairs and actions, weighted by the probability of each state. Concretely, the expected utility of σ given ρ is $EU_S(\sigma \mid \rho) = \sum_{t \in T} P(t)\ U_S(t, \rho(\sigma(t)))$ and the expected utility of ρ given σ is $EU_S(\rho \mid \sigma) = \sum_{t \in T} P(t)\ U_R(t, \rho(\sigma(t)))$. The expected u As an example, consider a game with $T = \{t_1, t_2\}$, $M = \{m_1, m_2\}$, $A = \{a_1, a_2\}$, and the following utility matrix:

	a_1	a_2
t_1	1, 1	0, 0
t_2	0, 0	1, 1

Possible sender and receiver strategies are, for example, $\sigma = \{t_1 \mapsto m_2, t_2 \mapsto m_1\}$ and $\rho = \{m_1 \mapsto a_2, m_2 \mapsto a_1\}$. These would have an expected utility of 1 for both sender and receiver, since when t_1 obtains with probability $P(t_1)$ the sender will use m_2 and to this message the receiver will respond with a_1 which achieves a payoff of 1, and when t_2 obtains with probability $P(t_2) = 1 - P(t_1)$ the sender will use m_1 and to this message the receiver will respond with a_2 which also achieves a payoff of 1. They also represent one of the two stable conventions in this game, the other being the pair of strategies $\sigma = \{t_1 \mapsto m_1, t_2 \mapsto m_2\}$ and $\rho = \{m_1 \mapsto a_1, m_2 \mapsto a_2\}$. Conventions of this kind in a signaling problem are what Lewis calls *signaling systems*. An example of complete miscoordination would be $\sigma = \{t_1 \mapsto m_1, t_2 \mapsto m_2\}$ and $\rho = \{m_1 \mapsto a_2, m_2 \mapsto a_1\}$. Partial coordination is achieved, for example, by $\sigma = \{t_1 \mapsto m_1, t_2 \mapsto m_1\}$ and $\rho = \{m_1 \mapsto a_1, m_2 \mapsto a_2\}$.

The approach adumbrated so far is not, nor does it attempt to be, a full-blown theory of meaning like one would have in the classical picture. However, we can already see how it attempts to address the study of meaning from a very different angle. There is a focus, not on meaning as a kind of correspondence, but rather on the use of signals for specific goals (but see Franke, 2013; Harms, 2010; Huttegger, 2007; Skyrms, 2010). There is also no appeal to truth as a guide for our inquiry. One advantage of such a paradigm shift is that we are no longer tied to the need of explaining how language hooks on to the world; this question is no longer of primary relevance. We can merely focus on trying to see how agents can use signals to cope with the world and achieve their goals. On the way, we will better understand meaning, not by trying to say what meaning is, but by trying to better understand how communication works. In the examples discussed so far, agents are assumed to possess and exercise rationality. Unlike in the classical picture, it is an instrumental, rather than procedural, rationality, but one where agents are still omniscient and highly strategic.

Lewis' account of the stability of conventions rests on what could be considered strong demands for there to be a certain degree of required common knowledge between the players. Namely, there needs to be a state of affairs that indicates to everyone involved that a certain regularity will hold, as well as "mutual ascription of some common inductive standards and background information, strategic rationality, mutual ascription of strategic rationality, and so on" (1969/2011, pp. 56–57). Agents are thus envisioned as omniscient of the game and highly strategic. These requirements can seem excessive, even more so if we consider how simple signaling systems are when compared to human languages. The models were introduced in order to help explain how language could get off the ground as a conventional system without any sort of prior agreement. However, if we consider the origins of language from a historical perspective, it seems implausible that the agents that started making use of primordial signaling systems which (hypothetically) evolved into languages possessed such advanced rationality. Furthermore, communication through simple message exchange is something that almost all animals do: monkeys use calls, birds use singing, bees use dances, ants use pheromone trails, and so on. A plausible account of the origin of language should first explain how signaling systems like those could get started, without requiring high standards of rationality of the agents involved.

In order to address this problem, Brian Skyrms (1996) proposes to study signaling problems in evolutionary terms. Rather than imagining, as Lewis does, rational agents making conscious decisions in possession of knowledge of the game and expectations of the behavior of other agents, we can imagine a simpler scenario inspired by biological evolution: there is a population of agents with biologically hardwired behaviors for engaging in interactions characteristic of a signaling problem; utility does not represent preference, but rather fitness for survival and reproduction; the make-up of the population evolves based on the relative fitness of the strategies represented in the population. Such a setup attempts to capture the main features of natural selection: in a diverse population, agents with more successful strategies thrive, while agents with less fit strategies die off. Although the inspiration for this scenario is biological evolution, similar things could be said about how ideas spread in a population of agents who can adopt or abandon them depending on how successful they prove to be (*e.g.* Benz, Jäger, and van Rooij, 2006; Pagel, 2009; Thompson, Kirby, and Smith, 2016), *i.e.* we can interpret these notions in terms of cultural evolution (Boyd and Richerson, 1985; Dawkins, 1978). The principles can be captured in formal models that abstract away from details of single interactions and behavior of individual agents, for example in the replicator dynamics (Taylor and Jonker, 1978). The only thing relevant to the replicator dynamics are the relative proportions of strategies in a given population and the utility function. Using it, one can compute which strategies evolve under which conditions.

Skyrms' evolutionary game theory approach to signaling games not only gives more plausible grounds to support Lewis' discussion of convention, but it also accomplishes an important conceptual change: it moves most of the theory and mathematical formalism to the descriptive side of the investigation. Utility represents how the modeler views the signaling problem and understands the relative advantages or disadvantages of different possible strategy combinations. Dynamics describe how strategies can evolve when driven by mechanisms of utility maximization. The shift in perspective allows interpretations that accommodate for limited non-strategic agents. While the general framework manages to abstract quite some details away from the formalization, it nevertheless leaves room for them, especially when it comes to the dynamics. We already mentioned the replicator equation that can be seen as representing biological or cultural evolution, but one can also use dynamics inspired by learning mechanisms (*e.g.* Roth and Erev, 1995), or even ones assuming a high degree of knowledge of the game and other players (*e.g.* Gilboa and Matsui, 1991; Mühlenbernd, 2011; Spike et al., 2016). This range of options goes hand in hand with a range of pictures of rationality, from nothing more than survival of the fittest in a biologically-inspired setting, to a certain degree of instrumental but limited and non-strategic rationality in the case of learning dynamics, to higher levels of rationality and even recursive strategic reasoning about the co-players' beliefs and choices. Each of these can be utilized depending on the problem that one is interested in characterizing. Thus, although Skyrms shows that high requirements of rationality are not necessary for signaling conventions to evolve, the framework does leave room for the study of linguistic interactions between highly strategic agents.

The characterization of signaling problems in terms of evolutionary game theory allow us to explain why certain equilibria come to be and how. A core notion in this context is that of an *evolutionary stable state* (Maynard Smith, 1982): an equilibrium situation that a population tends to under certain initial conditions and standard evolutionary pressures, and where it remains after small disturbances. With these tools, we can better understand why signaling systems are stable even without any strong assumptions of rationality. We can also map out which initial conditions drive the system towards which equilibria and which do not. In a simple case like the example discussed above, an evolutionary process of the kind described always drives the population into a state where one signaling system takes over completely. More complex signaling problems may have different evolutionary outcomes, sometimes unexpected ones. Skyrms (2010) gives an overview of different topics studied using signaling games, including expansions of the framework itself (for example, considering other dynamics beyond the replicator equation), exploration of other factors that impact the evolution of signaling (for example, how agents are interconnected), or variations on the signaling problem and its basic assumptions (for example, loosening the alignment of interests in order to provide accounts of deceptive signal use). Other uses of signaling games include discussions of categorization (*e.g.* Jäger and van Rooij, 2007), compositionality (*e.g.* Barrett, 2009), incommensurability (*e.g.* Barrett, 2010), just to name a few. More recent overviews are given by Huttegger, (2014), Huttegger et al., (2014), and Franke and Wagner, (2014).

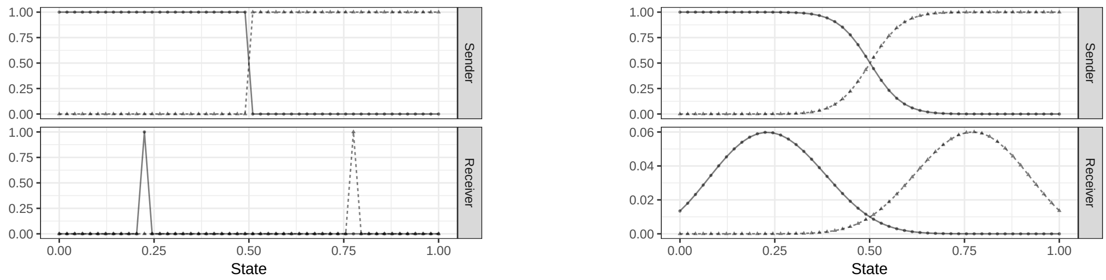
In the context of game-theoretic approaches to language, vagueness presents a new challenge. Barton Lipman (2009) gives a detailed account of the problem. Very succinctly, it can be put as follows. In standard game-theoretic models of communication, vague signal use yields a lower expected utility than crisp use. Therefore, given that the dynamics (be it natural selection, cultural evolution, or rational choice) maximize utility, vagueness should be weeded out by these forces, giving rise to only precise languages. However, vagueness is pervasive in our natural languages, and there is no reason to believe it is going away. Lipman concludes that "we cannot explain the prevalence of vague terms in natural language without a model of bounded rationality which is significantly different from anything in the existing literature" (2009, p. 1). What proposals have been put forward to address this issue?

3 Vagueness in sim-max games

In order to work, the sorites paradox requires us to assume a relation between the vague terms and a more precise underlying dimension (height for tallness, number of hairs for baldness, number of grains of sand for "heapness", and so on). Not only does this property need to be much more fine-grained than the vague term, but it also needs to have some structure, in the sense that there is at least an order between the elements in it (thinking of height in centimeters, $180 > 179 > \ldots > 120$), but usually even a degree of how far apart these elements are from each other. In terms of signaling games, we can take the state space to be constituted of values of the underlying dimension, and the message space as constituted by the terms in question. Because of the difference in granularity, we will typically be interested in cases where the state space is much larger than the message space. We can model the structure of the state space by defining a distance or similarity function between every value, effectively making it a metric space. Another important ingredient of the paradox is the acknowledgment of a certain degree of tolerance with respect to whether a certain term applies or not. This tolerance decreases with distance in state space: assuming a 180cm person is tall, one would easily tolerate the use of the term for a person measuring 179cm, less so for someone who is 170cm, and much less so for 160cm. This can be modeled by using a utility function that is continuous rather than discrete and that is monotonously decreasing with distance, *i.e.* success is not a matter of black and white, right or wrong, but a matter of degree, of how close the receiver got to the optimal response to the sender's perceived state.

The simplest type of game to study in this scenario is one where the state space and the action space are the same. We can imagine this as a game of guessing the state of affairs: the sender has knowledge of it, sends a message to the receiver, who in turn has to guess it; their payoff, as discussed above, is proportional to how close he got. These games, called similarity-maximization or sim-max games for short, were first introduced by Gerhard Jäger and Robert van Rooij (Jäger, 2007; Jäger and van Rooij, 2007) and further studied by Jäger, Metzger, and Riedel, (2011). What these authors find about this setup is that the evolutionary stable states are what they call Voronoi languages. Roughly, these are situations where the sender uses messages in a way that can be seen as partitioning the state space into convex regions and the receiver responds with the central element of those regions. For example, imagine a state space consisting of values of height in centimeters ranging from 40 to 280.[4] Given two possible messages, the optimal sender strategy

[4]Giving some slack to the values of the shortest and tallest people ever as recorded by Guinness World Records,



(a) Crisp (b) Vague

Figure 1: Example strategies for a state space with 50 states. Each line corresponds to a message. For the sender, it plots for each state the probability that the message is used. For the receiver, it plots for each state the probability that the response is that state, given the message.

could be to use one message for all values from 40 up to 160, and the other for all values from 160 up to 280; the optimal receiver strategy could be to guess 100 if given the first message and 220 given the second message. These precise numerical values give mutually optimal behavior for a case where the prior probability is the same for each state and utility is a linear or quadratic function of the distance between the actual state and the receiver's guess. In general, at which point the sender switches the use of messages and which guesses of the receiver are optimal or rational critically hinges on the priors over states and the utility function. This will be important later. Still, confirming Lipman's argument, there is no vagueness in any such optimal language: when given a height of 159 the sender will always use one message and given a height of 161 will always use the other; correspondingly, the receiver's response is also crisp, univocally associating one guess with each message.

In an abstract setup, using 50 states uniformly distributed over the unit interval and two possible messages, such an optimal language looks like what we see in Figure 1a: the probability that the sender uses one message decays sharply from 1 to 0, and increases sharply from 0 to 1 for the other message; the response of the receiver for each message is a degenerate distribution over the state space. The sender/receiver strategy pairs that we are looking for look more like Figure 1b, where use of one message over the other by the sender is characterized by a smooth transition, and the receiver strategy assigns for each message a positive probability to more than one state. The particular shapes of these strategies are not important to characterize a vague signal use. What is important is that the transition is smooth and monotonous for both players. This means that, for some states in the middle of the state space, there is uncertainty as to which message will be used, whereas for states in the extremes this is as good as certain.

The interpretation of this uncertainty will be different depending on the interpretation of the model. If we see it as an explanatory model of how two agents play the game, we can see it as randomization. If we interpret the model as descriptive, it simply represents expected behavior in a manner agnostic to the underlying mechanisms. A third option is to see probabilities as capturing relative numbers in a population of agents. For example, if the sender strategy assigns a probability of 0.4 of message m being sent for state t, this would mean that 40% of the population uses that message for that state, while 60% uses the other. This latter option leaves open the possible interpretation that each agent commands a crisp language and vagueness is only a population-level effect. However, given the level of abstraction of the description so far, none of this is necessarily implied by the model. Our question is thus what additional modifications to sim-max games are sufficient for the optimal languages to be more like Figure 1b, rather than like Figure 1a.

Franke, Jäger, and Rooij, (2011) make two suggestions of how vagueness in a signaling game can be boundedly rational, *i.e.* how vagueness could arise as a consequence of cost-saving limitations in the cognitive capacities of instrumentally rational agents. The first proposal is called limited

respectively at 54.6 and 272.

memory fictitious play (LMF) and models agents playing a sim-max game where their ability to recall past interactions with others is limited to a certain number. For a given interaction, each agent uses his limited memory of the other agent's past behavior to estimate the other's strategy. Each agent plays an instrumentally rational best response (given the utility function) to their estimate of the other's strategy. In order to study the evolution of strategies in repeated interaction, the authors model several individual agents in actual play. What they observe is the emergence of vague signaling at the level of the population, *i.e.* population averages of individual strategies exhibit the characteristics of a vague language as characterized before. However, each agent still commands a crisp language, which is inadequate if the intention is to capture how vagueness presents itself in human languages.

In order to overcome this limitation, Franke, Jäger, and Rooij make another proposal using the notion of a quantal response equilibrium (QRE). The idea, inspired by experimental psychology, is to model the choice of best response as stochastic rather than deterministic.[5] A prominent explanation for such soft-max or quantal choice behavior is that agents make small mistakes in the calculation of their expected utilities (Train, 2009). They still choose the option with the highest expected utility, but each assessment of the expected utilities is noise perturbed. This, in turn, may actually be boundedly rational since the calculation of expected utilities relies on assessing stochastic uncertainty, which in turn may be costly to calculate precisely. Choice based on a few samples from a distribution can be optimal if taking more samples or other means of better approximating degrees of uncertainty is resource costly (*e.g.* Sanborn and Chater, 2016; Vul et al., 2014). The degree to which agents tremble in the calculation of expected utilities and therefore deviate from the instrumentally rational behavior can be characterized by a parameter. Franke, Jäger, and Rooij find that for low values of this parameter, only babbling equilibria are possible, where sender and receiver simply randomize, respectively, message and interpretation choice uniformly. Above a certain value of the parameter, other equilibria of the kind described in the beginning of this section arise, where agents communicate successfully, though not perfectly, using fuzzy strategies similar to the ones depicted in Figure 1b. However, it is not clear whether soft-max choices capture the right stochastic trembles in decision making as they would arise under natural sources of uncertainty about the context (see Franke and Correia, 2017).

Cailin O'Connor (2014) proposes a way in which vagueness could be expected to evolve as a side-effect of a particular type of learning process. She studies sim-max games driven not by rational choice dynamics, but by generalized reinforcement learning (GRL), a variant of Herrnstein reinforcement learning (HRL) (Roth and Erev, 1995). In HRL, agents learn to play a signaling game by strengthening particular choices (of messages for the sender, of responses for the receiver) proportionally to how successful that choice proves to be in an interaction. O'Connor's proposal is to model generalization as the propagation of reinforcement to nearby states, where "nearby" is defined in terms of distance in state space. For example, if a sender was successful in using message m for state t, he will not only positively reinforce that choice of message for t, but also for states similar to t. This is done to a degree that is proportional to the similarity between t and other states. The dynamics also gives rise to vague signaling of the kind we are looking for.

But why should agents evolve to generalize? O'Connor suggests that, despite a vague language having lower expected utility than a precise one, the learning mechanism that induces vagueness does have evolutionary advantages, though less obvious ones (see O'Connor, (2015) for the detailed argument). Considerations of optimality of strategies, for example by considering the evolutionary stable states, are typically made in terms of hypothetical atemporal comparisons of expected utility. However, GRL has an interesting property in comparison with more strict learning mechanisms: it achieves higher payoffs in a shorter period of time. In a naturalized evolutionary setting, where we care about bounded rationality in a multi-context environment, learning speed is an advantage which should also be taken into account. Imagine an initial population of agents with random strategies, some using GRL and others using classical HRL to adapt to each other. Although the latter type of agent can hypothetically develop a precise and more efficient signaling system,

[5] Probabilistic choice rules are also the source of vagueness in recent accounts by Lassiter and Goodman, (online first) and Qing and Franke, (2014).

agents using GRL would coordinate on vague signaling strategies with high (though not optimal) expected utility sooner than agents using HRL. In such a scenario, they could drive the other agents to extinction before the latter had time to achieve coordination and reap the benefits of a more precise signaling system.

But would such a language be stable? In terms of evolutionary game theory, one could argue that a small enough group of mutants playing a precise signaling system could later invade a population of vague signalers. There is, however, an additional advantage to the generalized learning that is relevant here, namely the ability to adapt to a changing environment. Analysis of a game in terms of evolutionary stable states works well if we assume that the environment where the agents interact remains static. However, in a more realistic setting this is not the case. In a dynamic game with a variable environment, an ability to converge faster could give an evolutionary advantage over strict learning rules: as soon as the state space or the utility function change, GRL would quickly adapt and again potentially drive more strict learning mechanisms to extinction because of its short-term gains.

Franke and Correia, (2017) study a variant of the replicator dynamics that is motivated by perceptual limitations. Assuming that agents do not have perfect perception, there will always be a possibility that senders confuse states and receivers mix up responses. Furthermore, it seems reasonable to assume that this state confusability is proportional to state similarity, *i.e.* that the more similar two states are, the more likely it is that they will be mistaken for each other. Incorporating these considerations into a derivation of the replicator dynamics based on imitation processes, they develop a variant of that dynamic that also induces vague signal use of the kind we expect here. The consequence is very similar to that of the GRL model discussed above, in that the way the behavior for a given state is updated takes into account the behavior of similar states, proportionally to their similarity. Given the known relation between reinforcement learning and the replicator dynamics (Beggs, 2005), it is actually quite plausible that the two are tightly related (although this would need to be formally demonstrated). The account is, furthermore, interpretable at a lower level of rationality, given that the replicator equation is suitable to represent biological processes of natural selection.

The motivation underlying this model of vague signaling is still one of inevitability. A vague strategy is not claimed to have higher expected utility than a crisp one. However, the authors observe an effect similar to that pointed out by O'Connor: signaling converges faster in scenarios where there is some degree of state confusability. Furthermore, they observe one additional potentially beneficial property. By running several rounds of simulation for each parameter set, they can measure for each group of results how close resulting strategies are to each other, and how they would fare playing against one another. The results show that the within group distance between strategies becomes smaller with growing confusability, and the within group expected utility is actually higher for strategies evolving under a certain degree of state confusion. Thus, some amount of uncertainty seems to promote more homogeneous populations of signalers that are better at achieving cooperation within a group. In this picture, vagueness is the natural and unavoidable consequence of perceptual limitations, but may still contribute to the overall boundedly rational character of language if we take multiple contexts of learning into account.

Lawry and James, (2017) explore the potential benefits of vagueness in variations of a sim-max game with multiple senders and one receiver. Although the models considered are not explicitly identified as sim-max games, they are equivalent in the relevant aspects. They are signaling games where senders have private knowledge of a given state x, each sends a message to the receiver, which then selects an estimate y of the original state. Although the authors do not define a utility function, their analysis evaluates the models in terms of the expected error between the estimate and the original state, defined as $(x-y)^2$. Better models have lower expected error, thus one could say that the objective is to minimize this function, which would be equivalent to maximizing its inverse. The game could then be seen as a sim-max game with multiple senders and negative squared difference as utility function.

Unlike the models discussed so far that introduce minimal constraints to a dynamic and study the strategies that emerge out of an evolutionary process, Lawry and James' approach is mostly static and analytical. They consider several possible sender strategies, with and without uncertainty, and calculate if and when sender and receiver achieve better transmission accuracy, *i.e.* smaller expected error between original state and receiver estimate. Senders without uncertainty have a strategy that boils down to what is represented in Figure 1a: they split the state space according to a fixed threshold, switching from using one message to another abruptly. The model for the other type of sender is similar, but assumes that there is uncertainty regarding the value of the threshold, representing it as a uniformly distributed random variable. What results from this in terms of sender strategy is also a smooth transition between the use of two messages.[6]The receiver strategy is not inherently vague for either type of sender, it merely aggregates the messages received to provide an estimate of the original state using a simple arithmetic mean.

The authors analyze a number of variations of this model. For a model with one sender, results are consistent with Lipman's (Lipman, 2009) in that strategies without uncertainty lead to more accurate communication. For models with two messages and more than one sender, the expected error is a strictly decreasing function of the number of senders, and above a certain number of senders, strategies with uncertainty actually achieve lower expected error than crisp ones. This minimal number of senders from which vagueness surpasses sharpness is a decreasing function of the number of messages, *i.e.* whereas at least 8 senders are needed if the number of messages available is 2, this number lowers to 5 for 3 messages, 4 for and 5 messages, and 3 for more than 6 messages. If random errors are introduced between the choice of the sender and the message that reaches the receiver, increasing the number of senders can compensate for the noise, to a certain extent. Results are mixed when dropping the assumption of uniform priors for the state space.

These results are interesting since they reveal scenarios where vague strategies have an advantage over strict ones, even in a setting of game-omniscient strategic agents. The approaches discussed so far argue that vagueness is the inevitable byproduct of either cognitive limitations or a particular learning mechanism in the context of other learning mechanisms, which still leaves the idealist[7] some room for speculation. In the models of Franke, Jäger, and Rooij, the idealist could argue that higher rationality (better memory, or more reliable computation of expected utilities) would lead to more precision. O'Connor's learning strategy could theoretically be dominated by a process with higher strategic rationality, like best response dynamics (Gilboa and Matsui, 1991). A combination of both would hypothetically lead to a better outcome in the scenario of Franke and Correia. However, Lawry and James give us something that holds even if the idealist is unwilling to acknowledge the constraints that pervade our finite existence. Whether the scenario of multiple senders is representative of the language games where vagueness comes about, and whether the strategies proposed by the authors would be the ones to evolve[8], is something that is, however, open for debate.

4 How could language not be vague?

Arguments of the kind presented by Lipman (2009), that a vague language is suboptimal compared to a crisp one, work under a number of various assumptions. Together, these assumptions form a picture of the context where agents learn and use language that is highly idealized and probably derives, via game theory, from the conception of rationality of traditional theoretical economics. Herbert A. Simon describes this picture as follows:

> Traditional economic theory postulates an "economic man," who, in the course of being "economic" is also "rational." This man is assumed to have knowledge of the relevant aspects of his environment which, if not absolutely complete, is at least impressively clear and voluminous. He is assumed also to have a well-organized and stable system of

[6]Unlike what we see in Figure 1b, it is a linear transition. This is certainly connected with the choice of modeling the uncertain threshold as uniformly distributed. We speculate that another choice of distribution, such as for example a normal distribution, could yield something similar to Figure 1b.

[7]In an ordinary, non-philosophical, sense.

[8]The authors' analysis is a static one. Despite the advantage of the proposed vague strategies against strict ones in certain settings, one cannot know whether other strategies of higher accuracy would not be possible in those same settings without studying some dynamics in an evolutionary context.

> preferences, and a skill in computation that enables him to calculate, for the alternative courses of action that are available to him, which of these will permit him to reach the highest attainable point on his preference scale. (Simon, 1955, p. 99)

This "economic man", a game-omniscient and highly strategic, instrumentally rational agent, seems to have survived almost unscathed to now populate most game theoretic models of language and signaling. Both Simon and Lipman call for this picture to be revised.

In order to account for vagueness in natural language in the context of these models, the proposals surveyed here all need to peel away from this idealized picture and bring some of these assumptions down to earth. In the process, they point to ways in which we, as language learners and language users, are finite beings finding ways to cope with a highly complex and dynamic world:

1. Our existence is temporally finite; language does not have an infinite amount of time to evolve, nor can it take an infinite time to be learned. The faster a language can start being useful, the better.

2. Language learning through experience has to rely on a limited number of observations. Not only is the state space typically much larger than what one can survey in sufficient time, it is even potentially infinite and constantly changing.

3. A corollary of the former is that there will always be heterogeneity in a population of language learners, at the very least in their prior experience, since each agent will have relied on a different set of observations. Furthermore, this information is not directly or fully accessible to others.

4. Given that an agent is almost always learning and using language in a population of other agents, there are also various potential sources of linguistic input the agent is constantly integrating in his practice.

All of these observations support the weakening of the modeling assumptions. The research surveyed here shows us some examples of assumptions which, when weakened, make vague signal use a natural outcome of certain evolutionary dynamics. But it gives us even more. It suggests ways in which the mechanisms that lead to vagueness can have positive effects that are extremely important in the context of the points just enumerated. We learned from O'Connor and Franke and Correia that vague languages are quicker to converge and adapt, which is valuable given the finite and dynamic character of our experience (point 1). O'Connor also showed how generalization, an invaluable feature of any procedure for learning from a limited number of observations (point 2), leads to vagueness. We also learned from Franke and Correia that state confusability, a mechanism thats leads to vague signal use, can have a homogenizing effect on vocabularies, potentially compensating for the heterogeneity of agents' experiences (point 3). Lawry and James demonstrated the benefits of strategies that incorporate uncertainty, leading to vagueness, in language games where agents need to aggregate information from various sources (point 4).

What do these observations tell us about rationality? Generalized reinforcement learning (O'Connor, 2014), and the work of Franke and Correia, (2017), both paint a picture of agents with a basic level of instrumental rationality, possibly limited awareness of the game and a lack of strategic capabilities, adapting their behavior with only short-term gains in sight. Under the standard assumptions, even agents with such limited rationality would evolve crisp signaling in sim-max games. All approaches introduce constraints on agent behavior or information processing that prevent the evolution of crisp signal use. But a crisp language would still have a higher expected utility than the evolved strategies. Agents in those models seem to be rational only to the extent of their capabilities. Despite the plausibility of the mechanisms proposed (limited memory, imprecise calculation of expected utilities, generalization, state confusability), this feels like a bittersweet conclusion because of the hypothetical possibility of an ideal strategy that seems to be only artificially barred from the agents. But this is so only if rationality is equated with maximization of expected utility on a pure local level. The mechanisms that lead to vague signal use, as O'Connor,

(2014) and Franke and Correia, (2017) stress, have the aforementioned important advantages of faster speed of convergence, higher flexibility, and homogenization. These advantages matter, especially in the complex, dynamic context of real language use. Limits on memory and precision in the calculation of complex probabilistic beliefs may have a reduced metabolic processing cost. One would thus like to say that using a strategy that promotes those characteristics without a significant loss of expected utility is certainly a rational choice. But in order to be able to say that in the context of the model, we need either a notion of rationality that goes beyond maximization of expected utility, or a notion of expected utility that incorporates other dimensions. It is in this sense that these explanations tie in with a notion of bounded ecological rationality.

For more complex agents, the story is potentially more nuanced. Consider an agent that has awareness that something like a game is being played, some estimate of the potential payoff involved, ability to gauge other players' expected behaviors, and the capacity to think strategically. Is it rational for a sender to use a message in a crisp way, splitting the state space as if a strict threshold existed between the use of one term and the other? In a scenario where the agent's ability to make the optimal choice is bounded, such as in the models of Franke, Jäger, and Rooij, (2011), the agent is again doing the best he can given the constraints imposed on him. Lawry and James, (2017) draw attention to another option we need to consider: such agents could opt to ignore their limitations and external sources of uncertainty and use a crisp strategy regardless. The authors' analysis shows that such an approach would be irrational in a number of different scenarios, even in the narrow interpretation of rationality as local utility maximization. But even in other scenarios, we could again argue that, if we broaden our notion of rationality, it would be less rational to use a crisp strategy rather than one that incorporates at least some source of uncertainty. A crisp strategy for a particular game would, in most realistic scenarios, take more time to achieve convergence, be less adaptable to changes, less easy to reuse in other similar games, and less able to foster cooperation with a bigger group of agents. Though potentially a locally rational choice, it might thus not be globally optimal.

We believe these considerations are, again, very much in line with Wittgenstein's later philosophy:

> The more closely we examine actual language, the greater becomes the conflict between it and our requirement. (For the crystalline purity of logic was, of course, not something I had discovered: it was a requirement.) The conflict becomes intolerable; the requirement is now in danger of becoming vacuous. We have got on to slippery ice where there is no friction, and so, in a certain sense, the conditions are ideal; but also, just because of that, we are unable to walk. We want to walk: so we need friction. Back to the rough ground! (1953/2009, §107)

This seems like a very clear call for what we are arguing here. But are we exposing ourselves to "philosophical disease" by following a "one-sided diet" (1953/2009, §593) when looking only at sim-max games? In the spirit of Wittgenstein's pluralistic picture of language, we would like to reflect on vagueness in one more type of signaling game and speculate on the impact of plurality for vagueness.

5 Vagueness in multiple reference games

The type of linguistic interaction captured by a sim-max game has its own particular scope. There is a fixed prior, a fixed utility function and a fixed similarity measure between states. If the sender observes a state t and the receiver guesses state t' the utility of that interaction is proportional to the similarity between t and t'. Here is a natural example of a situation captured by a sim-max model: Kiki wants Bubu to draw her dissertation cover in a particular shade of blue; the closer Bubu's color choice matches Kiki's idea, the better. Here Kiki wants to *describe* the color for Bubu to pick out or realize from the whole set of possibilities. It's a natural thing to want to do with language. But other types of communication with vague and contiguously variable properties are possible. A *reference game* is a model of a situation in which the speaker has a particular object in mind and uses a description in order for the receiver to pick this object and no other object. One crucial difference with sim-max games is that close will not be close enough: if Kiki needs a long nail, a nail a little longer or shorter may be good enough; but if Bubu needs the long, slightly bend nail with the flat top (which she previously disassembled from her art installation), no other nail will do.

Interestingly, referential communication can take place under a plethora of different epistemic conditions. A lot of modeling work considers *reference games without displacement* where sender and receiver communicate about a context, i.e. a set of objects, which is immediately accessible in their direct perceptual environment (*e.g.* Baronchelli et al., 2010; Franke, 2012a,b). Agents may make mistakes about observing exactly how blue or long etc. a particular object is, but there is no further fundamental uncertainty about the context. This, of course, need not be so. Lipman, (2009) and van Deemter, (2009), for example, consider *reference games with displacement*. Kiki asks Bubu to fetch the red book from the library. No other book than the one Kiki needs will serve. When Kiki describes the book, the perceptual context in which Bubu must make her choice is not co-present to Kiki at the time of speaking. Yet other situations are imaginable. In *reference games with partial uncertainty*, for instance, interlocutors may know that they have shared and immediate access only to a subset of the objects in the context, but be uncertain about whether the co-player sees any other objects. This is the set-up for the widely-used director task in psycholinguistics (*e.g.* Keysar et al., 2000; Keysar, Lin, and Barr, 2003; Krauss and Glucksberg, 1977).

What would a rational agent do in repeated reference games? Take a universe in which agents play repeated reference games without displacement. A single game has sender and receiver both flawlessly perceive a context of k objects (with k possibly randomly drawn for each encounter); each object has its individual features, e.g., drawn from a rich multidimensional feature space of continuously variable properties (*e.g.* Franke, 2012a,b). The sender chooses a message m for the target object t, but she can condition her choice on the current context c. Likewise, the receiver chooses a referent t' from the available objects in c, again conditional on c. But if agents play many such reference games with different contexts c and c', strategic and instrumental rationality of local choices in c and c' is not enough to guarantee that agents' beliefs and behavior in c have any relation to their beliefs and behavior in c'. An agent could use a certain signal to denote, say, the tallest object in every context with between five and twelve objects or when the number of objects is a prime number, but use the same signal for the darkest object in all other contexts. In other words, neither rationality of choice in a fixed context nor evolutionary adaptation to each single context entails a natural, globally consistent behavior across multiple contexts. The situation only gets worse if we want to have agents use the same expressions in a consistent way also across different types of reference games, with possibly different epistemic conditions.

The picture just sketched neglects learning from observation, as well as generalization across contexts. For strategic and instrumentally rational agents, we need to consider also the ability to learn from previous experience and to generalize from the past to the future. For evolutionary adaptation, the story so far misses the key fact that languages and practices have to be learnable in appropriate time from the actual observables, lest what is learned and further transmitted will simply be a different language or practice (*e.g.* Kirby, Griffith, and Smith, 2014). But even if every language learner in history had been perfectly rational, it is extremely unlikely, as was discussed in the previous section, that anyone ever succeeded in reducing all uncertainty about language use and say with confidence: this is the threshold that truly separates heaps from non-heaps no matter what. For one, there may simply not be enough time in the lifespan of a rational learner to observe enough instances of language use for certainty about crisp denotations. For another, what needs to be learned likely changes over time. The optimal use for gradable properties hinges crucially on the prior over states (*e.g.* Franke, 2012a,b; Lassiter and Goodman, online first; Qing and Franke, 2014). If the prior captures real world attributes, such as the probability of an individual's body height, these attributes will not only be hard to know for certain, but they will also vary over time. They may also depend on location: do we use our language exclusively in this village (where all heaps are extraordinarily small and people are extraordinarily tall) or do we travel beyond? If so, the priors that determine the goodness of a convention for one agent need not be the same as those for another. Going beyond the priors for property degrees, the probability of seeing a particular context c in some kind of reference game, is also crucial for determining globally efficient linguistic behavior. But here it is even more inconceivable how a rational learner could acquire a fairly confident estimate over how often the upcoming reference games will have a particular context or another. The list of sources of natural and seemingly insurmountable uncertainty about the parameters necessary to determine locally or globally optimal strategies is much longer than hinted at here. Still, it should be clear enough that a finite existence, a sparse learning input and the inherent fuzziness of the linguistic context itself (its temporal, local and social variability) is not a fault of rationality on the part of the agent, or the mechanism of selection of successful linguistic practices.

What about meaning and vagueness in a universe of multiple language games and heavy uncertainty? We do not play only one language game in our existence; there is a plurality of them and which one an agent is engaged in at a particular moment is never clearly identified, neither are the exact benefits one might take out of it by choosing a certain behavior over another. These are furthermore not fixed in time; old language games fall out of fashion or stop being useful, and new ones emerge all the time (Wittgenstein, 1953/2009, §23). We could look for meaning at several levels in this pluralistic picture. Firstly, as before, there is the actual behavior of a single agent in each actualized language game. As mentioned above in connection the soft-max choice function used by Franke, Jäger, and Rooij, (2011), behavior that strictly maximizes expected utility under uncertainty may be resource heavy, so it might be compatible with local strategic rationality that agents' production choices are stochastic. Secondly, if we look at behavior across many game types and contexts, there is also the level of an agent's internal theory of how words and phrases are likely to be used (or even normatively: how messages should be used), conditional on a given context. Notice that a single agent's rational beliefs about linguistic practices or linguistic meaning may well have to reflect the actual stochasticity: under natural assumptions about information loss, the best belief for prediction matches the actual distribution in the real world (*e.g.* Vehtari and Ojanen, 2012). In sum, both at the level of behavior and at the level of beliefs about use or meaning, we should expect to find vagueness. Still, despite the natural vagueness, there does not seem to be anything fundamentally missing or conceptually incoherent in a naturalistic, rationality- or optimality-driven explanation of what each agent is doing or what each agent beliefs about language, use and meaning.

Taken together, when we look at the picture of multiple language games side by side, the question of what rational or optimal behavior is in any one of them, or the question of why human language users do not more closely approximate instrumentally rational or optimal local behavior more closely seem entirely misplaced. What seems more pressing is to answer how the interaction of individual learning, individual local optimization and general pressure for successful communication at a cross-situational level lead to the shaping of linguistic practices in an open-ended landscape of language games. How to attain locally sub-optimal play in a single language game is not a theoretical puzzle we should take serious, if what really matters is a global multi-game perspective. Adopting this perspective does not *solve* the problem of how vagueness is compatible with rationality. But it suggest that it might be an inflated issue that we need not or should not worry about, because it will likely not matter in a more realistic and encompassing picture of language, its use and evolution as driven by an appropriate notion of rational choice and evolutionary selection.

6 Conclusion

We tried to show that, if we pragmatically acknowledge the complex and dynamic environment that surrounds language learning and use, we find many plausible sources of uncertainty that can lead to vague signal use. We supported this claim by giving an overview of recent research that tries to accommodate for vagueness in the context of sim-max games, and by discussing, in more general terms, how these considerations are relevant for reference games, and in the broader context of language as a multiplicity of language-games. Furthermore, we argued that if we broaden our notion of rationality beyond mere maximization of expected utility, we also find reasons to welcome the positive aspects of these processes and to see their acceptance as a perfectly rational choice. Irrational would be to ignore the finite nature of our existence and the constraints that condition language learning and use. Our challenge as researchers is of course not only to draw attention to the possibilities, but also to characterize which particular constraints give rise to vague signal use and how.

The investigation doesn't end here, then. The research discussed here suggests that although vagueness can be beneficial, that is the case only in moderation: too much can lead to babbling equilibria where no information is transmitted. Thus, for agents capable of adjusting the level of vagueness in their strategies, questions remain regarding how this can be modeled and what factors influence the amount of vagueness that is optimal for a given situation. Other challenges include how to incorporate into our models the intuitions about the importance of speed of convergence, adaptability, and the promotion of group homogeneity. Finally, perhaps the biggest of them all is how to be able to focus on one language-game without losing sight of the ecosystem of games that need to be taken into account as well. Whatever solutions we find to these challenges, we believe that a notion of rationality cannot be strongly tied to mere utility maximization in a particular game, but has to consider an agents' need to optimize his behavior to navigate the plurality of uncertainties that besiege him. We need a clear notion of ecological rationality in order to see vagueness in the right light.

References

Baronchelli, Andrea et al. (2010). "Modeling the Emergence of Universality in Color Naming Patterns". In: *PNAS* 107.6, pp. 2403–2407.

Barrett, Jeffrey A. (2009). "The Evolution of Coding in Signaling Games". In: *Theory and Decision* 67.2, pp. 223–237.

— (2010). "Faithful Description and the Incommensurability of Evolved Languages". In: *Philosophical Studies* 147.1, pp. 123–137.

Beggs, Alan (2005). "On the convergence of reinforcement learning". In: *Journal of Economic Theory* 122.1, pp. 1–36.

Benz, Anton, Gerhard Jäger, and Robert van Rooij, eds. (2006). *Game Theory and Pragmatics.* Hampshire: Palgrave MacMillan.

Boyd, Robert and Peter J. Richerson (1985). *Culture and the Evolutionary Process.* University of Chicago Press.

Correia, José Pedro (2013). "The Bivalent Trap: Vagueness, Theories of Meaning, and Identity". MA thesis. Universiteit van Amsterdam.

Dawkins, Richard (1978). *The Selfish Gene.* New edition edition. London: Flamingo.

van Deemter, Kees (2009). "Utility and Language Generation: The Case of Vagueness". In: *Journal of Philosophical Logic* 38.6, pp. 607–632.

Fawcett, Tim W., Steven Hamblin, and Luc-Alain Giraldeaub (2013). "Exposing the behavioral gambit: the evolution of learning and decision rules". In: *Behavioral Ecology* 24.1, pp. 2–11.

Franke, Michael (2012a). "On Scales, Salience & Referential Language Use". In: *Amsterdam Colloquium 2011.* Ed. by Maria Aloni, Floris Roelofsen, and Katrin Schulz. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 311–320.

— (2012b). "Scales, Salience and Referential Safety: The Benefit of Communicating the Extreme". In: *The Evolution of Language: Proceedings of 9^{th} International Conference (EvoLang 9).* Ed. by Thomas C. Scott-Phillips et al. Singapore: World Scientific, pp. 118–124.

— (2013). "An adaptationist criterion for signal meaning". In: *The dynamic, inquisitive, and visionary life of ϕ, $?\phi$, and $\diamond\phi$: A festschrift for Jeroen Groenendijk, Martin Stokhof, and Frank Veltman.* Ed. by Marial Aloni, Michael Franke, and Floris Roelofsen, pp. 96–104.

Franke, Michael and José Pedro Correia (2017). "Vagueness and Imprecise Imitation in Signaling Games". In: *The British Journal for the Philosophy of Science.*

Franke, Michael, Gerhard Jäger, and Robert van Rooij (2011). "Vagueness, Signaling and Bounded Rationality". In: *New Frontiers in Artificial Intelligence*. Springer, Berlin, Heidelberg, pp. 45–59.

Franke, Michael and Elliott Wagner (2014). "Game Theory and the Evolution of Meaning". In: *Language and Linguistics Compass* 8/9, pp. 359–372.

Galeazzi, Paolo and Michael Franke (to appear). "Smart Representations: Rationality and Evolution in a Richer Environment". In: *Philosophy of Science*.

Gilboa, Itzhak and Akihiko Matsui (1991). "Social Stability and Equilibrium". In: *Econometrica: Journal of the Econometric Society*, pp. 859–867.

Hagen, Edward H. et al. (2012). "Decision Making: What Can Evolution Do for Us?" In: *Evolution and the Mechanisms of Decision Making*. Ed. by Peter Hammerstein and Jeffrey R. Stevens. Cambridge, MA: MIT Press, pp. 97–126.

Harms, William F. (2010). "Determining truth conditions in signaling games". In: *Philosophical Studies* 147, pp. 23–35.

Hume, David and Michael P Levine (1738/2005). *A Treatise of Human Nature*. English. OCLC: 962416442. New York: Barnes & Noble.

Huttegger, Simon et al. (2014). "Some Dynamics of Signaling Games". In: *Proceedings of the National Academy of Sciences of the United States of America* 111.Suppl 3, pp. 10873–10880.

Huttegger, Simon M. (2007). "Evolutionary Explanations of Indicatives and Imperatives". In: *Erkenntnis* 66, pp. 409–436.

— (2014). "How Much Rationality Do We Need to Explain Conventions?" en. In: *Philosophy Compass* 9.1, pp. 11–21.

Jäger, Gerhard (2007). "The Evolution of Convex Categories". In: *Linguistics and Philosophy* 30.5, pp. 551–564.

Jäger, Gerhard, Lars P. Metzger, and Frank Riedel (2011). "Voronoi Languages: Equilibria in Cheap-Talk Games with High-Dimensional Types and Few Signals". In: *Games and Economic Behavior* 73.2, pp. 517–537.

Jäger, Gerhard and Robert van Rooij (2007). "Language Structure: Psychological and Social Constraints". In: *Synthese* 159.1, pp. 99–130.

Kahane, Guy, Edward Kanterian, and Oskari Kuusela, eds. (2007/2013). *Wittgenstein and His Interpreters: Essays in Memory of Gordon Baker*. Reprint edition. Chichester: Wiley-Blackwell.

Keefe, Rosanna and Peter Smith (1997). "Introduction: theories of vagueness". In: *Vagueness: A Reader*. Ed. by Rosanna Keefe and Peter Smith. MIT Press, pp. 1–57.

Keysar, B. et al. (2000). "Taking Perspective in Conversation : The Role of Mutual Knowledge in Comprehension". In: *Psychological Science* 11.1, pp. 32–38.

Keysar, Boaz, Shuhong Lin, and Dale J. Barr (2003). "Limits on Theory of Mind Use in Adults". In: *Cognition* 89.1, pp. 25–41.

Kirby, Simon, Tom Griffith, and Kenny Smith (2014). "Iterated Learning and the Evolution of Language". In: *Current Opinion in Neurobiology* 28, pp. 108–114.

Krauss, R. M. and S. Glucksberg (1977). "Social and nonsocial speech". In: *Scientific American* 236, pp. 100–105.

Lassiter, Daniel and Noah D. Goodman (online first). "Adjectival vagueness in a Bayesian model of interpretation". In: *Synthese*.

Lawry, Jonathan and Oliver James (2017). "Vagueness and Aggregation in Multiple Sender Channels". en. In: *Erkenntnis*, pp. 1–38.

Lewis, David (1969/2011). *Convention: A Philosophical Study*. Oxford: Blackwell.

Lipman, Barton L. (2009). "Why Is Language Vague?"

Maynard Smith, John (1982). *Evolution and the Theory of Games*. Cambridge; New York: Cambridge University Press.

Mühlenbernd, Roland (2011). "Learning with Neighbors: Emergence of Convention in a Society of Learning Agents". In: *Synthese* 183, pp. 87–109.

O'Connor, Cailin (2014). "The Evolution of Vagueness". In: *Erkenntnis* 79.4, pp. 707–727.

— (2015). "Evolving to Generalize: Trading Precision for Speed". In: *The British Journal for the Philosophy of Science*.

Pagel, Mark (2009). "Human Language as a Culturally Transmitted Replicator". In: *Nature Reviews Genetics* 10, pp. 405–415.
Putnam, Hilary (1994). *Pragmatism: An Open Question*. Wiley-Blackwell.
Qing, Ciyang and Michael Franke (2014). "Gradable Adjectives, Vagueness, and Optimal Language Use: A Speaker-Oriented Model". In: *Proceedings of SALT 44*. Ed. by Jessi Grieser et al. elanguage.net, pp. 23–41.
Rorty, Richard (2000). "Universality and Truth". In: *Rorty and His Critics*. Ed. by Robert Brandom. Blackwell Publishers, pp. 1–30.
— (2007). "Wittgenstein and the Linguistic Turn". In: *Philosophy as Cultural Politics*. Philosophical Papers 4. Cambridge: Cambridge University Press, pp. 160–175.
Roth, Alvin E. and Ido Erev (1995). "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term". In: *Games and economic behavior* 8.1, pp. 164–212.
Russell, Bertrand (1923). "Vagueness". In: *Australasian Journal of Psychology and Philosophy* 1.2, pp. 84–92.
Sainsbury, R. Mark (1999). "Concepts without Boundaries". In: *Vagueness: A Reader*. Ed. by Rosanna Keefe and Peter Smith. Paperback edition. Cambridge, Massachusetts: MIT Press.
Sanborn, Adam N. and Nick Chater (2016). "Bayesian Brains without Probabilities". In: *Trends in Cognitive Sciences* 20.12, pp. 883–893.
Simon, Herbert A. (1955). "A Behavioral Model of Rational Choice". In: *The Quarterly Journal of Economics* 69.1, pp. 99–118.
— (1986). "Rationality in Psychology and Economics". In: *Journal of Business*, S209–S224.
Skyrms, Brian (1996). *Evolution of the Social Contract*. Cambridge University Press.
— (2010). *Signals: Evolution, Learning, and Information*. Oxford University Press.
Sorensen, Roy (2009). "Sorites Arguments". In: *A Companion to Metaphysics*. Ed. by Jaegwon Kim, Ernest Sosa, and Gary S. Rosenkrantz. Second. John Wiley & Sons, pp. 565–566.
Spike, Matthew et al. (2016). "Minimal Requirements for the Emergence of Learned Signaling". In: *Cognitive Science*.
Taylor, Peter D. and Leo B. Jonker (1978). "Evolutionary Stable Strategies and Game Dynamics". In: *Mathematical Bioscience* 40.1–2, pp. 145–156.
Thompson, Bill, Simon Kirby, and Kenny Smith (2016). "Culture Shapes the Evolution of Cognition". In: *Proceedings of the National Academy of Sciences of the United States of America* 113.16, pp. 4530–4535.
Train, Kenneth E. (2009). *Discrete Choice Methods with Simulation*. Cambridge, MA: Cambridge University Press.
Unger, Peter (1979). "There Are No Ordinary Things". In: *Synthese* 41.2, pp. 117–154.
Vanderschraaf, Peter (1998). "The Informal Game Theory in Hume's Account of Convention". In: *Economics and Philosophy* 14.02, pp. 215–247.
Vehtari, Aki and Janne Ojanen (2012). "A survey of Bayesian predictive methods for model assessment, selection and comparison". In: *Statistics Surveys* 6, pp. 142–228.
Vul, Edward et al. (2014). "One and Done? Optimal Decisions From Very Few Samples". In: *Cognitive Science* 38.4, pp. 599–637.
Wittgenstein, Ludwig (1953/2009). *Philosophical Investigations*. Ed. by P. M. S. Hacker and Joachim Schulte. Revised 4th edition. Chichester, U.K.: Wiley-Blackwell.
Zollman, Kevin J. S. and Rory Smead (2010). "Plasticity and Language: An Example of the Baldwin Effect?" In: *Philosophical Studies* 147.1, pp. 7–21.