Language and Linguistics Compass 8/9 (2014): 359–372, 10.1111/lnc3.12086

Game Theory and the Evolution of Meaning

Michael Franke* and Elliott O. Wagner
Institute for Logic, Language and Information, Universiteit van Amsterdam

Abstract

Evolutionary game theory is a general, but mathematically precise framework for modeling the competition between and fitness-based selection of different types of behavior. We review recent applications of this framework to account for the evolution of behavior that lends meaning to ostensible acts and signs.

Many consider it natural to think that linguistic expressions receive their meaning from the way they are used. This view contrasts, for instance, with seeing meaningfulness as a user-independent relation between expressions and some kind of abstract entities, like thoughts or Platonic ideas. The latter view is useful to describe the systematicity that we see in human language when conceived as an abstract system in its own right. But when the goal is to ponder on how meaningfulness could have developed along an evolutionary path from simpler forms of animal communication systems to full human languages, the notion that meaning arises from use in context provides much more explanatory grip.

But this is at best a starting point for further exploration, not an explanation yet. Filling in the necessary details raises complex issues: how does a particular behavioral pattern in a particular context give rise to meaning? Which features of the context of interaction do we need to take into account to explain meaningfulness? When would meaning–constituting behavior arise? How could meaningfulness found in animal communication systems develop into the more flexible, systematic and productive system that human language arguably is?

Answers to these questions are building blocks for a theory of language evolution. But theories of language evolution face a problem of uneven distribution of resources. Relevant data is sparse, but intuitions as fuel for speculation abound. To deal with data sparsity, it is clear what needs to be done (although that does not mean that unearthing new relevant data is easy). To check the quality of intuitions, formalization helps. This is where evolutionary game theory (EGT) comes in handy (although, again, with no promise of an easy academic life).

A game, in the relevant sense, is an abstract description of a situation in which several players make choices. The outcome for each player depends on what others choose. EGT is a mathematical approach to studying how different types of behavior in such games would increase, decrease, die out or prevail in populations of agents who are repeatedly playing the game in question. An explanation of the evolution of meaning would then consist of two steps. First, we fix a game that captures enough relevant features of a context of interaction in which meaningful communicative behavior is possible. Second, we consult a suitable solution concept of EGT to inform us whether or when a population of agents playing the game in question can be expected to evolve towards meaningful behavior.

The remainder of this paper is aimed at adding detail to this sketch of an approach to meaning evolution. Section 1 introduces signaling games, our basic game model to work with. Section 2 describes in what sense certain behavior in these games can be meaningful. Section 3 considers two solution concepts: evolutionary stability and the replicator dynamic. We argue that the latter type of solution is usually more insightful. Section 4 discusses extensions and generalizations of the basic game model with an eye towards current issues in the field.

1. Signaling Games – The Basics

Signaling games were originally introduced by Lewis (1969) to formalize interactions in which actors can use arbitrary signs to transmit information. There are two players, sender and receiver. The sender privately observes the state of the world. Each state occurs with a certain probability. The sender selects a signal, which is just some initially meaningless but publicly observable act. The receiver, who observes the signal, but not the state, then selects an act. Players obtain numerical payoffs depending on the realized state, the message used and the action performed. These payoffs encode how good an outcome is to a player and thus form the basis of defining the fitness of a particular type of behavior. The dynamic structure of the simplest non-trivial signaling game is given in Figure 1.

Consider a classic example drawn from biology. Domestic chickens have two main types of predators, aerial and terrestrial, and there is strong evidence that they have evolved to use two distinct vocalizations to accurately transmit information about which type of predator is present (Searcy & Nowicki, 2005, Chapter 2). To explain how chickens could have evolved this behavior, we model the underlying situation, with crude simplification, as a Lewis signaling game like that in Figure 1. The two types of predators are the states of the game. We assume for simplicity that Nature flips a fair coin to determine the predator type. The sender chicken observes the realized type and produces one out of two available signals. The receiver chicken observes the call and acts in an optimal way for either predator type, namely by either running for cover or standing tall and scanning the horizon. In evolutionary models, it is typical to identify payoffs with fitness. Running for cover is fitness enhancing if there is an aerial predator above, but bad if the predator is terrestrial. Similarly, standing tall and scanning the horizon is a good response to a terrestrial predator, but makes a chicken easy picking for an aerial predator. A stylized

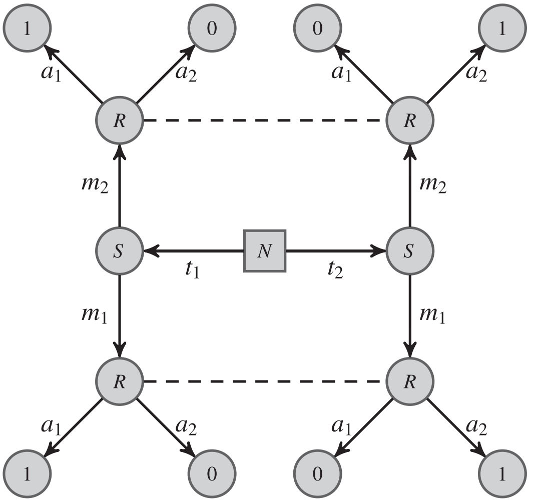

Fig. 1. Process graph of a Lewis game with two states, messages and acts. The game starts in the middle with nature choosing a state. Next, the sender chooses a message conditional on the observed state. The receiver does not know the actual state. Whence the dotted line between histories of play that the receiver cannot distinguish. The receiver selects an act conditional on the signal only. Players receive identical payoffs. If the act matches the state, the game is a success (payoff 1), otherwise a failure (payoff 0).

idealization of the fitness of these responses assumes that the receiver's fitness is 1 if she performs the action that is best for the realized predator type and 0 otherwise. For the time being (but see Section 4.3), we assume that the sender shares the receiver's payoffs and that the choice of signal does not impact payoffs.

Beyond chickens, signaling games have useful applications throughout economics and biology. Exchanges between potential employees and employers in job markets are often characterized as signaling games (following Spence, 1973). Signaling games also help explain why firms pay dividends even though recipients are subject to double taxation (Bhattacharya, 1979), how price setting and advertising can signal product quality (Nelson, 1974). Moreover, signaling games can be found throughout the literature of animal behavior as models of, to take just some examples, begging behavior (Smith, 1991), mating signals (Grafen, 1990) and threat displays (Enquist, 1985).[1]

2. Meaning In Signaling Behavior

One of Lewis' (1969) insights was that if players behave in certain complementary ways, it is as though messages have acquired conventional meanings (cf Nowak & Krakauer, 1999). Strategies are rules telling a player what to do in every relevant situation. Figure 2 shows all non-probabilistic strategies in the game from Figure 1.

Lewis observed that if players use their respective first strategy in Figure 2, then it is as though m_1 means 't_1 has occurred' or 'perform action a_1'. This meaning is conventional because there is another equally good communication scheme that the players could use, namely their respective second strategy. Other combinations do not give rise to meaningfulness in this sense.

This rudimentary type of meaning, which is simultaneously descriptive and directive, has been called a pushmi-pullyu representation (Millikan, 1984; Millikan, 1995) and primitive content (Harms, 2004). At first glance, it might seem as though this variety of meaning is so primitive as to be completely remote from human behavior (indeed, Section 4.4.1 asks how to differentiate these two characterizations), but Millikan (1995), for example, has argued that many utterances in human language have a pushmi-pullyu character, e.g., 'we do not eat peas with our hands' or 'the meeting is adjourned'. Lewis (1969) stressed that this type of meaning is fundamental not only to language but also to human communication more generally. To take just one of his examples, imagine a helper's gestures to a friend trying to park a car in a tight spot. These gestures are perhaps best thought of as conveying pushmi-pullyu content because each gesture seems to carry both descriptive information about the car's position and also directive information about what the driver ought to do.

An alternative characterization of meaning in signaling games draws on information theory. If $P(t_i)$ and $P(m_j \mid t_i) > 0$ are, respectively, the prior probabilities that t_i occurs and that if it does, m_j is sent, then

Strategy 1	Strategy 3
$t_1 \mapsto m_1$	$t_1 \mapsto m_1$
$t_2 \mapsto m_2$	$t_2 \mapsto m_1$
Strategy 2	**Strategy 4**
$t_1 \mapsto m_2$	$t_1 \mapsto m_2$
$t_2 \mapsto m_1$	$t_2 \mapsto m_2$

(a) sender

Strategy 1	Strategy 3
$m_1 \mapsto a_1$	$m_1 \mapsto a_1$
$m_2 \mapsto a_2$	$m_2 \mapsto a_1$
Strategy 2	**Strategy 4**
$m_1 \mapsto a_2$	$m_1 \mapsto a_2$
$m_2 \mapsto a_1$	$m_2 \mapsto a_2$

(b) receiver

Fig. 2. Pure strategies in the Lewis game.

$$P(t_i|m_j) = \frac{P(t_i) \times P(m_j|t_i)}{\sum_t P(t) \times P(m_j|t)}$$

is the likelihood of t_i given m_j by Bayes' theorem. Skyrms (2010a; 2010b) suggested to identify the informational content of m_j via a measure of differences between $P(t)$ and $P(t|m_j)$ for all t. More concretely, Skyrms (2010a) suggested to use the Kullback–Leibler divergence, which is a canonical measure of difference between two probability distributions (Kullback & Leibler, 1951). Doing so, the information of m_j about which state is actual is given by, intuitively speaking, how much information would be lost if the sender's way of using m_j was ignored. A complementary approach tracks the information a signal provides about the acts, so that the information theoretic approach also captures the pushmi-pullyu character of message use in these games. An advantage of this characterization is that in contrast to the informal theories developed by Lewis and others, it yields a quantitative measure of a signal's content that is applicable even when the players are not using deterministic strategies. But it remains to be seen how far these information theoretic ideas can be pushed in service of a real theory of meaning (Godfrey-Smith, 2011; Franke, 2013a).

3. *Evolutionary Stability And Dynamics*

But back to the chickens, how could the chickens evolve to use their available signals to communicate about predators in a meaningful way? There are two main approaches for answering this question: a static and a dynamic approach (cf Huttegger & Zollman, 2013). We will consider each in turn.

3.1. STATICS

The classical solution concept in evolutionary game theory is the static notion of evolutionary stability (Smith, 1974). We assume that there is a large population of agents who are randomly paired to play the game we are interested in. Each player is hardwired to play a strategy of the game. Payoffs received from the game are interpreted as indications of relative fitness, so that higher payoffs indicate higher chances of survival. A strategy is evolutionarily stable if when the whole population plays it, any small number of mutants playing another strategy will be driven to extinction.

To make this more concrete, let us suppose that the payoff for a player endowed with strategy s_i against a player of strategy s_j is $u(s_i, s_j)$. The idea that a huge monomorphic population of s_i players is impenetrable by mutant strategy s_j translates into the requirement that the fitness of the incumbent strategy s_i is strictly bigger than that of the mutant strategy s_j. If the proportion of mutants is ϵ, this requirement comes down to

$$\text{fitness of incumbent} > \text{fitness of mutant}$$
$$(1-\epsilon)u(s_i, s_i) + \epsilon u(s_i, s_j) > (1-\epsilon)u(s_j, s_i) + \epsilon u(s_j, s_j)$$

This is so, because under random encounters, incumbents play against themselves with average probability $(1 - \epsilon)$ and acquire a payoff of $u(s_i, s_i)$. With probability ϵ, they encounter a mutant and receive $u(s_i, s_j)$. Similarly, for the mutants, if we further assume that the proportion of mutants ϵ is infinitesimally small, the above inequality holds when $u(s_i, s_i) > u(s_j, s_i)$. Additionally, if $u(s_i, s_i) = u(s_j, s_i)$, then it also holds when $u(s_i, s_j) > u(s_j, s_j)$. Taken together, this lets us define that strategy s_i is an evolutionarily stable strategy (ESS) if for all s_j (i) $u(s_i, s_i) \geq u(s_j, s_i)$ and (ii) if $u(s_i, s_i) = u(s_j, s_i)$, then $u(s_i, s_j) > u(s_j, s_j)$.

To apply this definition to signaling games, we first need to transform the game into an adequate form. Since we assume that each agent in the population is endowed with both a sender and a receiver strategy, we look at a game between pairs of sender and receiver strategies with payoffs that are averages obtained from playing either role. This is known as symmeterizing the game (see, e.g., Cressman, 2003, Section 3.4). The Lewis signaling game in Figure 1 has four different sender and receiver strategies each (Figure 2). Because each agent has a strategy for each role, there are 16 different behavioral types, which we represent as a quadruple $\langle m_k, m_l, a_i, a_j \rangle$ where messages m_k and m_l are, respectively, the agent's choices for states t_1 and t_2 in his sender role, and a_i and a_j are the acts chosen in his receiver role after messages m_1 and m_2, respectively. Table 1 gives the symmetrized payoffs $u(s_i, s_j)$ of each pair of types playing against each other.[2]

The question to be asked is then when is a monomorphic population of individuals all playing a uniform strategy s_i evolutionarily stable, i.e., impenetrable by mutant invaders? This answer was first given by Wärneryd (1993) who proved that the game's only ESSs are the two strategies s_6 and s_{11} (see also Trapa & Nowak, 2000). These are exactly the strategies in which messages can be said to carry meaning. This result can be read off in Table 1. Strategy s_6, for instance, is an ESS because $u(s_6, s_6)$ is strictly bigger than $u(s_6, s_j)$ for all $j \neq 6$. The same applies to s_{11}. Some other strategies satisfy the first condition of the definition of ESS, but none of these satisfies the second as well. For example, $u(s_4, s_4)$ is among the maxima in the fourth column, but a monomorphic population of s_4 players could be invaded, for instance, by s_1.

Consequently, the static approach answers our motivating question how behavior constitutive of meaningfulness can arise in a resounding way. Only those strategy pairs that are constitutive of meaning in Lewis' sense are evolutionarily stable. This suggests that a population that is evolving according to natural selection and is occasionally subject to small mutations will inevitably end up developing meaning because such strategies are the game's only ESSs. However, as we will see next, this latter conclusion is not warranted in general, and we are well advised to look at explicit models of evolutionary dynamics.

Table 1. Payoffs of the symmetrized version of the signaling game from Figure 1.

		s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}	s_{12}	s_{13}	s_{14}	s_{15}	s_{16}
s_1	$<m_1,m_1,a_1,a_1>$	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5
s_2	$<m_1,m_1,a_1,a_2>$	.5	.5	.5	.5	.75	.75	.75	.75	.25	.25	.25	.25	.5	.5	.5	.5
s_3	$<m_1,m_1,a_2,a_1>$	.5	.5	.5	.5	.25	.25	.25	.25	.75	.75	.75	.75	.5	.5	.5	.5
s_4	$<m_1,m_1,a_2,a_2>$	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5
s_5	$<m_1,m_2,a_1,a_1>$	.5	.75	.25	.5	.5	.75	.25	.5	.5	.75	.25	.5	.5	.75	.25	.5
s_6	$<m_1,m_2,a_1,a_2>$	.5	.75	.25	.5	.75	1	.5	.75	.25	.5	0	.25	.5	.75	.25	.5
s_7	$<m_1,m_2,a_2,a_1>$	.5	.75	.25	.5	.25	.5	0	.25	.75	1	.5	.75	.5	.75	.25	.5
s_8	$<m_1,m_2,a_2,a_2>$	.5	.75	.25	.5	.5	.75	.25	.5	.5	.75	.25	.5	.5	.75	.25	.5
s_9	$<m_2,m_1,a_1,a_1>$	.5	.25	.75	.5	.5	.25	.75	.5	.5	.25	.75	.5	.5	.25	.75	.5
s_{10}	$<m_2,m_1,a_1,a_2>$	.5	.25	.75	.5	.75	.5	1	.75	.25	0	.5	.25	.5	.25	.75	.5
s_{11}	$<m_2,m_1,a_2,a_1>$	.5	.25	.75	.5	.25	0	.5	.25	.75	.5	1	.75	.5	.25	.75	.5
s_{12}	$<m_2,m_1,a_2,a_2>$	.5	.25	.75	.5	.5	.25	.75	.5	.5	.25	.75	.5	.5	.25	.75	.5
s_{13}	$<m_2,m_2,a_1,a_1>$	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5
s_{14}	$<m_2,m_2,a_1,a_2>$|	.5	.5	.5	.5	.75	.75	.75	.75	.25	.25	.25	.5	.5	.5	.5	.5
s_{15}	$<m_2,m_2,a_2,a_1>$	.5	.5	.5	.5	.25	.25	.25	.25	.75	.75	.75	.75	.5	.5	.5	.5
s_{16}	$<m_2,m_2,a_2,a_2>$	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5	.5

3.2. DYNAMICS

Rather than relying on an implicit picture of the evolutionary process and then identifying the game's evolutionarily stable states, the dynamic approach applies an explicit model of evolution and then looks at that model's temporal behavior and eventual outcomes. The most prominent evolutionary dynamic used in game theory is the replicator dynamic (see Weibull, 1995; Hofbauer & Sigmund, 1998; Cressman, 2003; Sandholm, 2010). It was originally introduced to model asexual reproduction in a large, well-mixed population (Taylor & Jonker, 1978). But since it can also be derived, e.g., as a model of social learning based on imitation of successful strategies (see Sandholm, 2010), the replicator dynamic is a general and serviceable model of both biological evolution and cultural change via social learning.

Imagine, once more, that we have a large population of individuals playing the symmetrized game in Table 1. Suppose individuals reproduce in quantities equal to their expected payoffs, i.e., the average payoff earned when randomly matched with another population member. Then the rate of change of the proportion of each strategy s_i is tracked by the differential equation:

$$\dot{x}_i = x_i \left(u(s_i, x) - \sum_k x_k\, u(s_k, x) \right)$$

where x is a vector with x_i the proportion of s_i players in the population and $u(s_i, x) = \sum_j x_j\, u(s_i, s_j)$ the expected payoff for s_i. The rate of change of the frequency of each strategy is equal to its relative abundance times the difference between its expected payoff and the average expected payoff. Accordingly, strategies that perform better than average grow in frequency and strategies that underperform decay.

There are two main techniques for determining the predictions of the replicator dynamic. In some comparatively simple systems, it may be possible to use tools from the theory of non-linear differential equations in order to prove theorems about which states will be the endpoints of evolution. When this approach fails, we can use computer simulations or numerical integration in order to get a feel for the system's behavior.

How does the dynamic approach play out for our simple chicken signaling game embedded in the replicator dynamic? Suppose we are at a random population state x, conceived of as a probability distribution over all 16 possible symmetrized strategies given in Table 1. The average payoff of all of these strategies depends on the relative prevalence of other strategies: how successful a particular chicken can signal and avoid predators given its sending and receiving behavior depends on the types of signaling behaviors that it would be, on average, confronted with. But which strategies proliferate and which decrease in frequency is not easy to see in general. Using numerical simulations, Skyrms (1996) observed that random initial population states of this system always converge to monomorphic populations of s_6 or s_{11} players. Huttegger (2007a) and Pawlowitsch (2007) proved that the probability of selecting an initial population composition that does not evolve a convention of semantic meaning under the replicator dynamic in this game is zero; meaning is guaranteed to evolve.

3.3. COMPARING THE STATIC AND DYNAMIC APPROACHES

A clear virtue of the static approach is its simplicity. Finding a game's ESSs is simple compared to analyzing a system of non-linear differential equations such as the replicator dynamic. Additionally, a well-known set of theorems provides important links between the two approaches (e.g., Sandholm, 2010, Theorems 8.4.1–8.4.7). For instance, if x is an ESS and if the population has a composition near to x, i.e., it is like x but with a few mutants added, then the population will stay near to x and in the limit converge to x under the replicator dynamic. Because similar theorems exist for many other evolutionary dynamics, the static approach is dynamic-neutral in some sense. By finding a game's ESSs, one gains considerable information about the possible evolutionary outcomes of a wide range of actual dynamics.

Unfortunately, the static approach does not provide a full evolutionary picture. It tells us something about equilibrium stability and maintenance, but not necessarily about how likely an evolutionary outcome is. There are, for instance, games in which, although a state is an ESS, the probability of choosing a random initial population configuration that evolves towards the given ESS under the replicator dynamic can be quite small (e.g., the famous stag hunt game, see Skyrms, 2004). Indeed, it can be the case that a dynamic does not lead to an ESS. It could, for instance, cycle infinitely or show another form of out-of-equilibrium behavior. Any such behavior will be missed by the static approach.

Another virtue of the dynamic approach is its flexibility. The static point of view is tied to the notion of an ESS, which is based on an informal and intuitive picture of evolution via natural selection. This picture is serviceable as a first-pass idealization, but is unrealistic in many ways. For example, it models the population as so large as to be effectively infinite and assumes that every agent is equally likely to meet every other. The replicator dynamic makes these idealizations as well, but the dynamic approach has further options. The replicator–mutator dynamic, for example, generalizes the replicator dynamic by including fixed mutation rates from one strategy to another (for overview, see Page and Nowak, 2002; Hofbauer and Sigmund, 2003). The Moran process (Moran, 1962), for instance, is a version of the replicator dynamics for finite populations. Pawlowitsch (2008) showed that it always leads to the evolution of meaning in Lewis signaling games. To relax the assumption that all encounters are equally likely, it would be natural to imagine agents who play the signaling game with their neighbors in a social network. In this setting, it has been shown that the agents develop regional systems of meaning (Zollman, 2005; Wagner, 2009; Mühlenbernd, 2011). Different meaning conventions arise in different areas in the network. The dynamic approach is flexible in that it gives the inquirer the liberty to choose an evolutionary dynamic appropriate for the system to be modeled.

So which approach to use? We agree with Huttegger & Zollman (2013) that it is best to approach the evolution of meaning with a somewhat pluralistic methodology. The dynamic approach provides information about the system that the purely static approach is blind to, and it can be tailored to many different situations for which the background assumptions of the static methodology are inappropriate (cf Franke, 2013b, for similar arguments). But the static approach can supplement dynamic analysis by providing a quick way to identify the states that may be stable across a wide range of dynamics.

4. *Variations of the Standard Model*

So far, we have only considered a very simple signaling game. We turn now to some extensions and generalizations that relate to current and emerging issues in the field.

4.1. MORE STATES, MESSAGES AND ACTIONS

Results about the evolutionary certainty of meaning evolution stated so far do not hold unconditionally for Lewis signaling games with more than two states, more than two messages and more than two actions. For instance, games with three states, messages and actions have a family of evolutionarily significant equilibria in which messages do not fully reveal the world's state, but do transmit partial information. These new equilibria are known as partial pooling equilibria (Figure 3).

Such equilibria are not ESSs, but neutrally stable. Whereas in an ESS, any small group of mutants is driven to extinction, neutral stability only requires that small groups of mutants do not proliferate (Smith, 1982). Pawlowitsch (2007) gave a complete characterization of the partial pooling equilibria of Lewis signaling games and showed that under the replicator dynamic, these equilibria have basins of attraction with positive measure (see also Huttegger, 2007a). This means that, contrary to what one might assume, populations are not guaranteed to evolve a system of maximally informative signaling even if such a system would be to the advantage of both the sender and receiver.

There has been considerable research into how populations can overcome the hurdles posed by partial pooling equilibria and thereby evolve a fully informative language. Pawlowitsch (2007) proved that drift prevents finite populations from becoming trapped in partial pooling equilibria (see also Fox & Shamma, 2011). This picture has been elaborated in Pawlowitsch et al. (2011) to explain language diversification. Huttegger et al. (2010) used numerical simulations to argue that partial pooling equilibria are destabilized by a constant but small rate of mutation, and Wagner (2009) showed that populations who interact with their neighbors in a social network are similarly unlikely to evolve to partial pooling strategies.

Another mechanism that avoids partial pooling is message innovation. Instead of assuming a fixed repertoire of signals, senders may spontaneously innovate new signals (cf McKenzie Alexander et al., 2012). If the receiver's initially random response to the new signal is successful, the innovation is kept. Signals that are not used for a while might be forgotten again (cf Barrett & Zollman, 2009). McKenzie Alexander et al. (2012) showed that innovation and forgetting can be very conducive of communication, often avoiding the pitfalls of partial pooling.

4.2. SIM-MAX GAMES

Sim-max games generalize Lewis signaling games in two ways (cf Jäger, 2007; Jäger & van Rooij, 2007). First, while there are a small number of signals, there is a very large number of states, possibly infinitely many. Second, there is an exogenously given similarity relation between states that inform the game's payoff structure. Optimal signaling strategies bestow meaning on signals, and concomitantly also yield a rudimentary form of categorization.

More concretely, sim-max games can be thought of as implementing a context in which agents communicate about the degree to which a given object x instantiates a continuously variable perceptual property P, like hue, weight, or length. The sender knows the degree t to which x has P, but the receiver does not. After observing the sender's signal, the receiver chooses act a which we imagine as the receiver's interpretation of the degree to which x has property P. Payoffs are proportional to the similarity between t and a, because, so a sketchy motivation runs, the better a matches t, the greater the receiver's chance of identifying x by property P among other objects having property P. For instance, imagine that the sender classifies a particular tree where a food source is as 'tall'. The receiver forms an expectation of the tallness of the tree in question. This interpretation is good depending on the receiver's chance of finding the tree,

Sender			Receiver		
t_1	$\mapsto$	m_1, m_2	m_1	$\mapsto$	a_1
t_2	$\mapsto$	m_3	m_2	$\mapsto$	a_1
t_3	$\mapsto$	m_3	m_3	$\mapsto$	a_2, a_3

Fig. 3. Example of a partial pooling equilibrium in a Lewis signaling game with three states, three messages and three actions. The picture represents a family of mixed populations with some agents sending m_1, others m_2 in state t_1. Similarly, some receivers respond to m_3 with a_2, others with a_3.

which, in very crude approximation, is proportional to the similarity between the actual height of the tree and the receiver's expectation.[3]

Jäger et al. (2011) proved that ESSs of sim-max games with a continuous state space and a similarity measure that satisfies certain natural conditions all give rise to so-called Voronoi languages, where, loosely speaking, the senders' use of signals 'partitions' the state space into convex categories and the receivers choose 'prototype interpretations' for each signal (see Figure 4). Sim-max games thus seem to provide a handle on the formation of categories in perceptual space, driven by the affordances of linguistic interaction. They have been used to study the organization of color terms (Jäger & van Rooij, 2007) and the evolution of vagueness (Franke et al., 2011; O'Connor, 2014). Sim-max games are an object of active current investigation, as, for example, details about their dynamical properties are missing.

4.3. BEYOND COMMON INTEREST

Up to this point, we assumed aligned interests of sender and receiver. Few if any real-world interactions are like this. Consider once more chicken alarm calls. Why does the sender benefit if the receiver takes the best evasive action? In models of alarm calls, this issue is often sidestepped by assuming that players are genetically related so that the sender's inclusive fitness increases if the receiver's evasion is successful (Searcy & Nowicki, 2005, Chapter 6). But the sender might risk drawing the attention of the predator, and this may outweigh the benefit of alerting the receiver to danger. How does meaning evolve when preferences do not match?[4]

This question is largely open. Many of the applications mentioned in Section , for instance, are games where the players' preferences do not completely align, and most of these games have not been studied using evolutionary dynamics.[5] Two notable exceptions include Spence's (1973) job market signaling game and the Sir Philip Sidney game (Smith, 1991). The latter is a game meant to model begging behavior. The sender is either healthy or needy and can either remain silent or produce a costly vocalization. After the vocalization (or lack thereof), the receiver can either donate or withhold food. If players are related, then, in terms of inclusive fitness, there are equilibria in which the costly signal is used to mean 'I am needy' or 'give me food.' These equilibria are evolutionarily stable, and they are the game's only ESSs. But how likely is it that a population evolves to communicate in this way? Huttegger & Zollman, (2010) showed by numerical simulation that unless the degree of relatedness is extremely high, the replicator dynamic only rarely leads to meaningful signaling. The picture is similar in Spence's game: although costly signaling is an ESS, it is an unlikely evolutionary outcome (Wagner, 2013).

What about extreme forms of payoff divergence? If any gain by the sender is a loss to the receiver and vice versa, meaningful signaling is not an equilibrium because any information transmitted will be used by the receiver to the sender's detriment. Based on this fact, it has often been claimed that information transfer is impossible in these games. Wagner (2012) has shown that this intuition is not entirely correct. A population that evolves according to the replicator dynamic, for example, will use signals in meaningful ways. This is because the population does

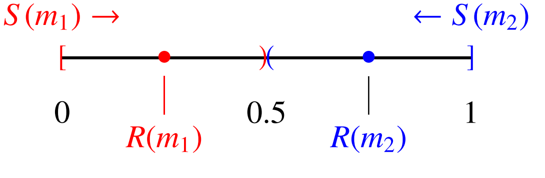

Fig. 4. Example of a Voronoi language for a sim-max game with two messages and T = A = [0; 1] when each state/degree is equally likely. The sender uses m_1 for degrees <.5, the receiver's interpretation of m_1 is .25.

not evolve to an equilibrium but instead leads to persistent out-of-equilibrium behavior in the form of deterministic chaos. Messages always have meaning, but the meaning of each message is constantly in flux.

4.4. MORE PLAYERS

Another way to extend signaling games is to add more players. Adding receivers brings in a notion of context-dependence. Adding senders yields structured signals. The former leads to a richer notion of behaviorally constituted meaningfulness. The latter helps explain rudimentary forms of compositionality and logical inference.

4.4.1. More Receivers

As discussed in Section 2, meaning in ESSs of the Lewis signaling game is functionally ambiguous. Each signal's use in an ESS is perfectly correlated with exactly one state and exactly one action. Lewis (1969) saw this ambiguity as one between description and prescription. To obtain functionally less ambiguous signal meaning, we would ideally like to see some signals' use decoupled from the occurrence of states and other signals' use decoupled from the occurrence of acts. This would constitute a first step towards a functional disambiguation and an evolutionary explanation of basic speech-act distinctions.

Zollman (2011) introduced a signaling game with two receivers. The sender can send one of two messages to each receiver independently. There are two states, two messages and two acts for each receiver. States and acts are perfectly correlated, as before, but additionally, the receivers should perform different acts for maximal payoff (Table 2).

Additional receivers bring in a notion of context-dependence. The distinction between receivers is essentially a distinction between what is the best act in a given state. The sender can condition her choice on which receiver she is facing, i.e., which context the conversation takes place in. Consequently, one type of ESS has the sender ignore the state component and tailor her choice to the receiver's type (Figure 5a). Another ESS-type has the sender ignore the receiver's type and condition her choice only on the state (Figure 5b). In the first type, the use of signals carries information exclusively about the act to be performed, whereas in the second type it carries information exclusively about the state that obtains. The former is prescriptive, almost like an imperative, the latter descriptive, similar to a stereotypical use of an indicative sentence.

Whether this is a good model of the evolution of illocutionary meaning distinctions in the sense of speech-act theory (Austin, 1962; Searle, 1969; Searle & Vanderveken, 1985) is debatable (see Huttegger, 2007b; Franke, 2012, for alternative accounts). Still, we see how the addition of receivers can extend our basic model to include context-dependent use of signals in a manageable and yet potentially insightful way.

Table 2. State-act payoff table for the game discussed by Zollman (2011). The columns list the combined choices of both receivers, i.e., $<i,j>$ is a situation in which the first receiver chooses *i* and the second chooses *j*. Rows give the actual state. All players have identical payoffs.

	$\langle a_1, a_1 \rangle$	$\langle a_1, a_2 \rangle$	$\langle a_2, a_1 \rangle$	$\langle a_2, a_2 \rangle$
t_1	0	1	0	0
t_2	0	0	1	0

Sender			Receiver 1		
t_1, R_1	$\mapsto$	m_1	m_1	$\mapsto$	a_1
t_1, R_2	$\mapsto$	m_2	m_2	$\mapsto$	a_2
t_2, R_1	$\mapsto$	m_2	**Receiver 2**		
t_2, R_2	$\mapsto$	m_1	m_1	$\mapsto$	a_1
			m_2	$\mapsto$	a_2

(a) prescriptive ESS

Sender			Receiver 1		
t_1, R_1	$\mapsto$	m_1	m_1	$\mapsto$	a_1
t_1, R_2	$\mapsto$	m_1	m_2	$\mapsto$	a_2
t_2, R_1	$\mapsto$	m_2	**Receiver 2**		
t_2, R_2	$\mapsto$	m_2	m_1	$\mapsto$	a_2
			m_2	$\mapsto$	a_1

(b) descriptive ESS

Fig. 5. Examples of prescriptive and descriptive ESSs of Zollman's (2011) game.

4.4.2. More Senders

So far, signals were simple and unrelated. Human language, on the other hand, has a rich syntax that interfaces with a compositional semantics to achieve great expressivity. Even many animal signals appear to be structured. Bird song comes to mind, but also some monkey alarm calls consist of sequences of independently meaningful calls (cf Arnold & Zuberbühler, 2006; Ouattara et al., 2009b; Ouattara et al., 2009a). Sequences of signals can be studied by assuming that there are several speakers, or, equivalently, that a single speaker independently sends one of several basic signals repeatedly in sequence. The multi-speaker model was introduced by (Barrett, 2007; Barrett, 2009) to account for conventionality of concepts (see also Section 4.2); the one-speaker reformulation was proposed by (Skyrms, 2010b, Chapter 12) to account for the beginnings of compositionality and logical inference.

Suppose there are four states and four acts. Payoffs are as usual. Suppose further that there are two basic messages, but that the sender can send a message twice in sequence. The receiver, then, responds to the pair of messages received. Figure 6 gives an example of an ESS where the meaning of the structured signals can be broken down into a conjunction of the meaning of its parts. In the given example, the first message reveals whether the state is in $\{t_1, t_2\}$ or $\{t_3, t_4\}$, the second whether the state is in $\{t_1, t_3\}$ or $\{t_2, t_4\}$. Skyrms (2010b) proposed to see the beginnings of compositional meaning and logical inference here (cf Nowak et al., 2000, 2001, for different approaches). This suggestion might be controversial (cf Franke, 2014, for criticism), but the simple example shows how signaling games can in principle be used to investigate the bearing of sequential structure in signals on their, allegedly, compositional meaning.

5. *Conclusion*

We suggest that EGT is a good tool for studying the evolution of language in general and of meaning in particular. The unrivaled appeal of EGT is its generality. EGT provides a mathematical framework for studying precisely the competition among alternative ways of relevant behavior and the gradual increase of 'better' or 'fitter' types of behavior without necessarily

Sender 1			Sender 2			Receiver		
t_1	$\mapsto$	m_1	t_1	$\mapsto$	m_1	m_1, m_1	$\mapsto$	a_1
t_2	$\mapsto$	m_1	t_2	$\mapsto$	m_2	m_1, m_2	$\mapsto$	a_2
t_3	$\mapsto$	m_2	t_3	$\mapsto$	m_1	m_2, m_1	$\mapsto$	a_3
t_4	$\mapsto$	m_2	t_4	$\mapsto$	m_2	m_2, m_2	$\mapsto$	a_4

Fig. 6. Example of an ESS in a game with two senders/sequential signaling.

stipulating any concrete and application-specific details about how individuals interact, learn or adapt. It is due to the combination of its generality and mathematical precision that EGT is an optimal framework for probing the numerous pre-theoretic intuitions we may have about the origin and development of language.

Short Biographies

Michael Franke's work lies at the junction between theoretical linguistics, philosophy, logic and the cognitive sciences. He holds a BS in Cognitive Science from Osnabrück University, an MS in Logic and a PhD in Philosophy from the University of Amsterdam. Most of his previous work was on applications of formal models from game and decision theory to linguistics. After spending 2 years as a researcher at the Department of Linguistics of the University of Tübingen, he is currently a researcher at the Institute for Logic, Language and Computation in Amsterdam.

Elliott Wagner's research focuses on the application of evolutionary dynamics to philosophical questions, including those surrounding the origins of morality and semantic meaning. He completed his BS in Computer Science at Columbia University and his MA in Mathematical Behavioral Science and PhD in Philosophy at the University of California, Irvine. After that, he has been a postdoctoral researcher at the Institute for Logic, Language and Computation at the University of Amsterdam. He is now an assistant professor in the Department of Philosophy at Kansas State University.

Notes

*Correspondence address: Michael Franke, Institute for Logic, Language and Information, Universiteit van Amsterdam, PO Box 94242, 1098 XH Amsterdam, The Netherlands. E-mail: mchfranke@gmail.com

[1] Within linguistics, signaling games have also been used to explain a number of pragmatic phenomena, mostly following the work of Grice, (1975), such as the notion of speaker meaning (Parikh, 2000), the contextual resolution of underspecification (e.g., Parikh, 1992; Parikh, 2001) or conversational implicatures (e.g., van Rooij, 2004; Benz, 2006; Benz & van Rooij, 2007; Franke, 2011; Jäger, 2014). However, this line of research usually assumes that there is already a pre-given semantic meaning (see Jäger, 2008; Franke, 2013b, for overview). This, then, is complementary to the approach outlined here.

[2] To obtain $u(s_i, s_j)$, consider each state t of the game. Look at the message m that s_i sends in t and then the act a that s_j chooses after m. Note the payoff that results for t and a. Do the same for s_i and s_j exchanged. Take the average of all four noted payoffs.

[3] Actually, this motivation is unduly hand-wavy and possibly ill-conceived, since much depends on how the receiver selects referents based on his degree expectation, as well as the distribution of objects from which to pick from. Unfortunately, we have to gloss over the (open) issue of a precise evolutionary interpretation of sim-max payoffs for reasons of space.

[4] A seemingly related question arises for actual language use under conflicting interests (cf Franke et al., 2012; de Jaegher and van Rooij, 2014, for game-theoretic treatments). Again, the difference is whether we assume that a meaningful language is already in place or whether we are interested in whether a certain level of cooperativity is essential for the establishment of a meaning convention. It is the latter we are interested in here. Yet another issue is whether small conflicts of interest between speakers and hearers might be a motor of language change. But this, too, is beyond the scope of this article.

[5] Biologists have been using the static approach for identifying plausible evolutionary outcomes since at least Enquist, (1985), but only very recently have explicitly dynamic models been applied to signaling games.

Works Cited

Alexander, Jason McKenzie, Brian Skyrms, and Sandy L. Zabell. 2012. Inventing new signals. *Dynamic Games and Applications* 2. 129–145.

Arnold, Kate, and Klaus Zuberbühler. 2006. Language evolution: semantic combinations in primate calls. *Nature* 441. 303.

Austin, John L. 1962. *How to do things with words: the William James lectures delivered at Harvard University in 1955*. Oxford: Clarendon.

Barrett, Jeffrey A. 2007. Dynamic partitioning and the conventionality of kinds. *Philosophy of Science* 74. 527–546.

——. 2009. The evolution of coding in signaling games. Theory and Decision 67. 223–237.

Barrett, Jeffrey A., and Kevin J. S. Zollman. 2009. The role of forgetting in the evolution and learning of language. *Journal of Experimental and Theoretical Artificial Intelligence* 21(4). 293–309.

Benz, Anton. 2006. Utility and relevance of answers. *Game theory and pragmatics*, ed. by Anton Benz, Gerhard Jäger and Robert van Rooij, 195–219. Hampshire: Palgrave.

Benz, Anton, and Robert van Rooij. 2007. Optimal assertions and what they implicate. *Topoi* 26. 63–78.

Bhattacharya, Sudipto. 1979. Imperfect information, dividend policy, and "the bird in the hand" fallacy. *The Bell Journal of Economics* 10. 259–270.

Cressman, Ross. 2003. *Evolutionary dynamics and extensive form games*. Cambridge, Massachusetts: MIT Press.

Enquist, Magnus. 1985. Communication during aggressive interactions with particular reference to variation in choice of behaviour. *Animal Behaviour* 33(4). 1152–1161.

Fox, Michael J, and Jeff S. Shamma. Language evolution in finite populations. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pages 4473–4478. IEEE, 2011.

Franke, Michael. 2011. Quantity implicatures, exhaustive interpretation, and rational conversation. *Semantics & Pragmatics* 4(1). 1–82.

——. 2012. On assertoric and directive signals and the evolution of dynamic meaning. *International Review of Pragmatics* 4(2). 232–260.

——. 2013a. An adaptationist criterion for signal meaning. In Marial Aloni, Michael Franke, and Floris Roelofsen, editors, *The dynamic, inquisitive, and visionary life of* ϕ, $?\phi$, *and* $\Diamond\phi$: *A festschrift for Jeroen Groenendijk, Martin Stokhof, and Frank Veltman*, pages 96–104.

——. 2013b. Game theoretic pragmatics. *Philosophy Compass* 8(3). 269–284.

——. 2014. Creative compositionality from reinforcement learning in signaling games. In Erica A. Cartmill, Sean Roberts, Heidi Lyn and Hannah Cornish /The Evolution of Language: Proceedings of EvoLang 10. 82–89.

Franke, Michael, Gerhard Jäger, and Robert van Rooij. 2011. Vagueness, signaling and bounded rationality. *New frontiers in artificial intelligence*, 45–59. Berlin: Springer.

Godfrey-Smith, Peter. 2011. Signals: evolution, learning, and information, by Brian Skyrms. *Mind* 120(480). 1288–1297.

Grafen, Alan. 1990. Biological signals as handicaps. *Journal of Theoretical Biology* 144. 517–546.

Grice, Paul Herbert. 1975. Logic and conversation. *Syntax and semantics, vol. 3, speech acts*, ed. by Peter Cole and Jerry L. Morgan. 41–58. New York: Academic Press.

Harms, William F. 2004. *Information and meaning in evolutionary processes*. Cambridge: Cambridge University Press.

Hofbauer, Josef, and Karl Sigmund. 1998. *Evolutionary games and population dynamics*. Cambridge: Cambridge University Press.

Hofbauer, Josef and Karl Sigmund 2003. "Evolutionary Game Dynamics". In: *Bulletin of the American Mathematical Society* 40.4, pp. 479–519.

Huttegger, Simon M. 2007a. Evolution and the explanation of meaning. *Philosophy of Science* 74(1). 1–27.

——. 2007b. Evolutionary explanations of indicatives and imperatives. *Erkenntnis* 66. 409–436.

Huttegger, Simon M., and Kevin J. S. Zollman. 2010. Dynamic stability and basins of attraction in the Sir Philip Sidney game. *Proceedings of the Royal Society B: Biological Sciences* 277(1689). 1915–1922.

——. 2013. Methodology in biological game theory. *The British Journal for the Philosophy of Science* 64(3). 637–658.

Huttegger, Simon M., Brian Skyrms, Rory Smead, and Kevin J.S. Zollman. 2010. Evolutionary dynamics of Lewis signaling games: signaling systems vs. partial pooling. *Synthese* 172(1). 177–191.

de Jaegher, Kris and van Rooij, Robert. 2014. Game-Theoretic Pragmatics Under Conflicting and Common Interests. *Erkenntnis* 79 769–820.

Jäger, Gerhard. 2007. The evolution of convex categories. *Linguistics and Philosophy* 30(5). 551–564.

——. 2008. Applications of game theory in linguistics. Language and Linguistics Compass 2(3). 406–421.

Jäger Gerhard. 2014. Rationalizable signaling. *Erkenntnis* 79. 673–706.

Jäger, Gerhard, and Robert van Rooij. 2007. Language structure: psychological and social constraints. *Synthese* 159(1). 99–130.

Jäger, Gerhard, Lars P. Metzger, and Frank Riedel. 2011. Voronoi languages: equilibria in cheap-talk games with high-dimensional types and few signals. *Games and Economic Behavior* 73(2). 517–537.

Kullback, S., and R. A. Leibler. 1951. On information and sufficiency. *The Annals of Mathematical Statistics* 22(1). 79–86.

Lewis, David. 1969. *Convention: a philosophical study*. Cambridge: Harvard University Press.

Millikan, Ruth Garrett. 1984. *Language thought and other biological categories*. Cambridge, Massachusetts: MIT Press.

——. 1995. Pushmi-pullyu representations. Philosophical Perspectives 9. 185–200.

Moran, P. A. P. 1962. *The statistical processes of evolutionary theory*. Oxford: Clarendon Press.

Mühlenbernd, Roland. 2011. Learning with neighbours. *Synthese* 183(1). 87–109.
Nelson, Phillip. 1974. Advertising as information. *The Journal of Political Economy* 82(4). 729–754.
Nowak, Martin A., David C. Krakauer. 1999. The evolution of language. *PNAS* 96. 8028–8033.
O'Connor, Cailin. 2014. The evolution of vagueness. *Erkenntnis* 79 707–727.
Ouattara, Karim, Alban Lemasson, and Klaus Zuberbühler. 2009a. Campbell's monkeys concatenate vocalizations into context-specific call sequences. *PNAS* 106(51). 22026–22031.
——. 2009b. Campbell's monkeys use affixation to alter call meaning. *PLoS ONE* 4(11). e7808, 11.
Page, Karen M. and Martin A. Nowak 2002. "Unifying Evolutionary Dynamics". In: *Journal of Theoretical Biology* 219, pp. 93–98.
Parikh, Prashant. A game-theoretic account of implicature. In Yoram Moses, editor, *TARK '92: Proceedings of the 4th conference on Theoretical aspects of reasoning about knowledge*, pages 85–94, San Francisco, 1992. Morgan Kaufmann Publishers Inc.
——. 2000. Communication, meaning, and interpretation. *Linguistics and Philosophy* 23(2). 185–212.
——. 2001. *The use of language*. Stanford University: CSLI Publications.
Pawlowitsch, Christina. 2007. Finite populations choose an optimal language. *Journal of Theoretical Biology* 249. 606–616.
——. 2008. Why evolution does not always lead to an optimal signaling system. *Games and Economic Behavior* 63(1). 203–226.
Pawlowitsch, Christina, Panayotis Mertikopoulos, and Nikolaus Ritt. 2011. Neutral stability, drift, and the diversification of languages. *Journal of Theoretical Biology* 287. 1–12.
van Rooij, Robert. 2004. Signalling games select Horn-strategies. *Linguistics and Philosophy* 27. 493–527.
Sandholm, William H. 2010. *Population games and evolutionary dynamics*. Cambridge, Massachusetts: MIT Press.
Searcy, William A., and Stephen Nowicki. 2005. *The evolution of animal communication*. Princeton, New Jersey: Princeton University Press.
Searle, John R. 1969. *Speech acts: an essay in the philosophy of language*. Cambridge, UK: Cambridge University Press.
Searle, John R., and Daniel Vanderveken. 1985. *Foundations of illocutionary logic*. Cambridge, UK: Cambridge University Press.
Skyrms, Brian. 1996. *Evolution and the social contract*. Cambridge, UK: Cambridge University Press.
——. 2004. *The stag hunt and the evolution of social structure*. Cambridge: Cambridge University Press.
——. 2010a. The flow of information in signaling games. *Philosophical Studies* 147. 155–165.
——. 2010b. *Signals: evolution, learning, and information*. Oxford: Oxford University Press.
Smith, John Maynard. 1974. The theory of games and the evolution of animal conflicts. *Journal of Theoretical Biology* 47(1). 209–221.
——. 1982. *Evolution and the theory of games*. Cambridge: Cambridge University Press.
——. 1991. Honest signalling: the Philip Sidney game. *Animal Behavior* 42. 1034–1035.
Spence, Michael. 1973. Job market signaling. *The Quarterly Journal of Economics* 87. 355–374.
Taylor, P. D., and L. Jonker. 1978. Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences* 40. 145–156.
Trapa, Peter E., and Martin A Nowak. 2000. Nash equilibria for an evolutionary language game. *Journal of Mathematical Biology* 41(2). 172–188.
Wagner, Elliott O. 2009. Communication and structured correlation. *Erkenntnis* 71(3). 377–393.
—— 2012. Deterministic chaos and the evolution of meaning. *The British Journal for the Philosophy of Science* 63(3). 547–575.
—— 2013. The dynamics of costly signaling. *Games* 4(2). 163–181.
Wärneryd, Karl. 1993. Cheap talk, coordination, and evolutionary stability. *Games and Economic Behavior* 5(4). 532–546.
Weibull, Jörgen W. 1995. *Evolutionary game theory*. Cambridge, Massachusetts: The MIT press.
Zollman, Kevin J.S. 2005. Talking to neighbors: the evolution of regional meaning. *Philosophy of Science* 72(1). 69–85.
Zollman, Kevin J.S. 2011. Separating directives and assertions using simple signaling games. *The Journal of Philosophy* CVIII(3). 158–169.