J Log Lang Inf
DOI 10.1007/s10849-016-9236-9

# Compositional Signaling in a Complex World

**Shane Steinert-Threlkeld**[1]



**Abstract** Natural languages are compositional in that the meaning of complex expressions depends on those of the parts and how they are put together. Here, I ask the following question: *why* are languages compositional? I answer this question by extending Lewis–Skyrms signaling games with a rudimentary form of compositional signaling and exploring simple reinforcement learning therein. As it turns out: in complex worlds, having compositional signaling helps simple agents learn to communicate. I am also able to show that learning the meaning of a function word, once meanings of atomic words are known, presents no difficulty.

**Keywords** Signaling games · Compositionality · Reinforcement learning · Evolution · Negation

*…a singing creature, only associating thoughts with tones.*

von Humboldt (1836)

## 1 Introduction

In the epigraph, Humboldt points to an awe-inspiring feature of human language: that it allows us to use sensible devices to transmit thoughts across the gap between human minds. Perhaps even more remarkable than this feat is *how* we accomplish it. As Humboldt later puts it, through language we "make infinite use of finite means" (pp. 98–99). This has been taken to be an early statement of the productivity of our

✉ Shane Steinert-Threlkeld
shanest@stanford.edu

1 Department of Philosophy, Stanford University, 450 Serra Mall, Stanford, CA 94305, USA

Published online: 18 February 2016

linguistic competence, which has in turn been used to motivate the famous principle of *compositionality*. Roughly:[1]

(C) The meaning of a complex expression is determined by the meaning of its parts and how the parts are put together.

A sizable body of literature in philosophy and linguistics has debated the truth of (C).[2] Most arguments for (C) take the form of inferences to the best explanation: the truth of (C) gives us the best explanation of the learnability, systematicity, and productivity of our understanding of natural languages.[3] These last two properties refer, respectively, to the fact that anyone who understands a sentence like 'dogs like cats' automatically understands 'cats like dogs' and the fact that competent speakers can understand novel expressions that are built out of known simpler expressions and syntactic rules. Here, my aim is not to evaluate any of these arguments in particular or to answer the question of whether (C) is true. Rather, I will simply suppose that (C) is true. Given that, I ask a further question: *why* is (C) true? In other words, what explains the fact that natural languages are compositional?

At first glance, it is hard to know what form an answer to this *why*-question would even take. The kind of answer I am interested in is broadly *evolutionary*. A crude first pass would transform one of the above arguments for the truth of (C) into an answer to the present question. Those arguments all take the form:

1. Our linguistic abilities have property P.
2. The best explanation for (1) is that (C).
3. Therefore, (C).

By starting at the same point but taking a different fork in the argumentative road, one can turn this into an evolutionary argument along the following lines:

1. Our linguistic abilities have property P.

2'. That (1) would have been good for our ancestors.

3'. Therefore, linguistic abilities with property P are likely to have evolved.

Unfortunately, this general form of argument is too unconstrained. Plenty of properties—say, the ability to run 50 miles per hour—would be good to have but have not evolved. In addition to constraints coming from the laws of physics and ecological pressures, the mechanisms of evolution—natural selection, in particular—operate locally in the sense that they can only select for genetic variations already present in a population of organisms.[4] Moreover, the mechanisms for introducing variation

[1] The earliest statement in the Western tradition appears to be in Frege (1923). Pagin and Westerståhl (2010a) find a similar statement in an Indian text dating from the fourth or fifth century.

[2] A related literature also focuses on making the statement of (C) more precise so as to avoid worries that it is trivial and/or vacuous. See, for an overview, §2.1 of Pagin and Westerståhl (2010b). In what follows, I will assume that (C) has been formulated in an appropriately more precise way.

[3] For learnability, see Davidson (1964). For systematicity and productivity, see Fodor and Pylyshyn (1988), Fodor (1987), and Fodor (1998). Some also take the principle to be a methodological one, guiding inquiry in semantics. See Szabó (2013), Pagin and Westerståhl (2010a, b), and Janssen (1997) for overviews of all of these proposals.

[4] See chapter 3, section 5 of Bergstrom and Dugatkin (2012) for a discussion of constraints on natural selection. Futuyma (2010) considers more constraints than just a bottleneck in variation.

(drift, mutation, etc.) are unlikely to make every possible beneficial feature available. Haldane (1932) puts the point quite nicely:

> A selector of sufficient knowledge and power might perhaps obtain from the genes at present available in the human species a race combining an average intellect equal to that of a Shakespeare with the stature of Carnera. But he could not produce a race of angels. For the moral character or for the wings he would have to await or produce suitable mutations. (p. 110)

More generally, we can say that evolutionary answers to the *why*-question are constrained by a *how*-question: how, actually, did natural languages come to be compositional? Because the kind of evidence needed to answer this question—e.g. detailed archaeological records—is hard to come by and incomplete, the strongest constraint will be a different *how*-question: how, possibly, could natural languages come to be compositional? In this paper, I attempt to answer the *why*-question using models that are simple and widely-used enough to plausibly satisfy this how-possibly constraint. In particular, I will add a very rudimentary form of compositionality to Lewis–Skyrms signaling games and show that compositionality can help simple agents learn to signal in increasingly complex environments. This will exhibit a sense in which it would be good to have a compositional language. But the model of signaling and of learning used are simple enough that this learnability advantage could have conferred a fitness advantage in the prehistory of our species. Thus, while the results to be presented do not present a story about the gradual emergence of compositionality, they are likely to be compatible with any such story.

The paper is structured as follows. In the next section, I introduce signaling games and the simple form of reinforcement learning that I will use later. Then, in Sect. 3, I introduce the *negation game*, which enriches basic signaling games with non-atomic signals that are interpreted compositionally. Following that, Sect. 4 presents simulation results of reinforcement learning in the negation game in increasingly large state spaces. We will see that compositional interpretation only makes learning easier in large state spaces and that how much compositionality helps learning strongly correlates with the number of states. From there, I drop the assumption that negation is built-in and ask whether the meaning of a function word can be learned. In a suitable set-up, I am able to prove that negation is learnable and offer simulation results that show that such learning happens very quickly. After that, I compare my work with other proposals in the literature and conclude with some future directions.

## 2 Signaling Games and Reinforcement Learning

Lewis (1969) introduced signaling games to understand how stable linguistic meaning could be conventional. In the simplest case, there are two agents—the Sender and the Receiver—who have common interests. The Sender can see which of two states of the world (call them $s_1$ and $s_2$) obtains and has two signals (call them $m_1$ and $m_2$) to send to the Receiver. The Receiver cannot see which state of the world obtains, but only which signal was sent by the Sender. Then, the Receiver can choose to perform one of two acts (call them $a_1$ and $a_2$). It is assumed that $a_1$ is appropriate in $s_1$ and $a_2$ in

$s_2$ in that both the Sender and Receiver receive a benefit if $a_i$ is performed in $s_i$, but not otherwise. Ideally, then, the Sender will send $m_1$ only in one state and $m_2$ only in the other and the Receiver will perform the appropriate act upon receiving $m_i$. There are two ways of this happening, corresponding to the two so-called *signaling systems*. Schematically:

$$\begin{tikzcd} s_1 \arrow[r] & m_1 \arrow[r] & a_1 \\ s_2 \arrow[r] & m_2 \arrow[r] & a_2 \end{tikzcd}$$

$$\begin{tikzcd} s_1 \arrow[dr] & m_1 \arrow[dr] & a_1 \\ s_2 \arrow[ur] & m_2 \arrow[ur] & a_2 \end{tikzcd}$$

If we view the meaning of $m_i$ as the state(s) in which it gets sent, this shows that the meaning depends on which signaling system the Sender and Receiver choose to adopt.[5] In this sense, it is conventional.

These ideas can be developed slightly more formally, in a way that will help for later. Writing $\Delta(X)$ for the set of probability distributions over a finite set $X$, we introduce a string of definitions.

**Definition 1** (*Signaling Game*) A signaling game is a tuple $\langle S, M, A, \sigma, \rho, u, P\rangle$ of states, messages, acts, a sender $\sigma : S \to \Delta(M)$, a receiver $\rho : M \to \Delta(A)$, a utility function $u : S \times A \to \mathbb{R}$, [6] and a distribution over states $P \in \Delta(S)$. $\sigma$ and $\rho$ have a common payoff, given by

$$\pi(\sigma, \rho) = \sum_{s \in S} P(s) \sum_{a \in A} u(s, a) \cdot \left( \sum_{m \in M} \sigma(s)(m) \cdot \rho(m)(a) \right) \tag{1}$$

We will also refer to $\pi(\sigma, \rho)$ as the *communicative success rate* of $\sigma$ and $\rho$. If $\sigma$ is not probabilistic but rather a deterministic function, we call it a *pure sender*. *Mutatis mutandis* for $\rho$ and receiver.

The most well-studied class of signaling games are generalizations of the two-state game described above, where there is an equal number of states, signals, and acts, with exactly one act appropriate for each state. I will call these Atomic $n$-games.

**Definition 2** (*Atomic n-game*) The *Atomic n-game* is the signaling game where $|S| = |M| = |A| = n$, $u(s_i, a_j) = 1$ iff $i = j$, 0 otherwise, and $P(s) = 1/n$ for every $s \in S$.

[5] This corresponds to one of two types of informational content for a signal identified in chapter 3 of Skyrms (2010) and reflects the idea that a proposition is a set of possible worlds.

[6] In not having the utility function depend on the signal sent, we are assuming that no signal is more costly to send than any other.

**Definition 3** The *Lewis game* is the Atomic 2-game.

**Definition 4** A *signaling system* in a signaling game is a pair $(\sigma, \rho)$ of a sender and receiver that maximizes $\pi(\rho, \sigma)$.

The reader can verify that the two signaling systems depicted at the beginning of this section are indeed the only ones in the Lewis game according to this definition. In fact, they are the only two strict Nash equilibria of the Lewis game. More generally, Trapa and Nowak (2000) show that the Atomic $n$-game always has strict Nash equilibria, all of which have the following form. By currying, we can view $\sigma$ as an $n \times n$ stochastic matrix $S$, with $S_{ij} = \sigma(s_i)\left(m_j\right)$ and $\rho$ as an $n \times n$ stochastic matrix $R$, with $R_{kl} = \rho(m_k)(a_l)$. Then: $(\sigma, \rho)$ is a stict Nash equilibrium iff $S$ is a permutation matrix[7] and $R = S^T$.[8] These strict equilibria have $\pi(\sigma, \rho) = 1$, which is the maximum value that $\pi(\sigma, \rho)$ can obtain. There are, however, other Nash equilibria in the Atomic $n$-game which have lower communicative success rates.

### 2.1 Reinforcement Learning in Signaling Games

Lewis' full analysis of conventional meaning in terms of signaling systems makes strong assumptions about common knowledge and rationality of the Sender and Receiver. Thus, while it provides a nice conceptual analysis, it cannot offer an explanation of how agents might *arrive at* or come to use a signaling system. In a long series of work, Skyrms (1996, 2004, 2010) and colleagues have weakened these assumptions and explored various dynamic processes of learning and evolution in the context of signaling games. Here, I outline one exceedingly simple form of learning called Roth–Erev reinforcement learning.[9]

Now, instead of the signaling game being played once, we imagine a Sender and Receiver repeatedly playing the game. We also imagine that each is equipped with some urns: the Sender has an urn for each state with balls in it labeled by signals, while the Receiver has an urn for each signal with balls in it labeled by acts. When the Sender finds out what state obtains, it draws a ball from the corresponding urn and sends the signal written on the ball. When the Receiver hears the signal, it draws a ball from the corresponding urn and performs the act written on the ball. We assume initially that each urn has one ball of every appropriate kind. If the act was appropriate for the state, the Receiver adds a ball of that act to the signal's urn and the Sender adds a ball of that signal to the state's urn. Otherwise, nothing happens. The addition of balls of the same type after successful play of the game makes it more likely that the same choices will be made given the same state of affairs in the future, thus reinforcing the choices that led to successful signaling. In the context of the how-possibly constraint mentioned in the introduction, it must be noted that Schultz et al. (1997) show that

[7] A matrix with only 1s and 0s, where each row has a single 1 and distinct rows have 1s in distinct columns.

[8] Here, $M^T$ denotes the transpose of $M$, i.e. $\left(M^T\right)_{ij} = M_{ji}$.

[9] See Roth and Erev (1995). Sutton and Barto (1998) provides a detailed introduction to reinforcement learning.

dopamine neurons in certain areas of primate brains appear to implement a similar reinforcement learning procedure.[10]

Slightly more formally and generally,[11] we keep track of the *accumulated rewards* of the players' choices. That is, we have functions $ar_{\sigma,t} : S \times M \to \mathbb{R}$ and $ar_{\rho,t} : M \times A \to \mathbb{R}$. At $t = 0$, these are set to some initial values, usually 1. They are then incremented by

$$ar_{\sigma,t+1}\left(s_i, m_j\right) = ar_{\sigma,t}\left(s_i, m_j\right) + u\left(s_i, a_k\right)$$
$$ar_{\rho,t+1}\left(m_j, a_k\right) = ar_{\rho,t}\left(m_j, a_k\right) + u\left(s_i, a_k\right)$$

where $s_i$, $m_j$, and $a_k$ were the state, message, and act played at time $t$. How, though, do the Sender and Receiver choose their signals and acts? The simple idea captured by the urn metaphor is that they do so in accord with their accumulated rewards; that is, to the extent that those choices have been successful in the past. In other words:

$$\sigma_{t+1}\left(s_i\right)\left(m_j\right) \propto ar_{\sigma,t}\left(s_i, m_j\right)$$
$$\rho_{t+1}\left(m_j\right)\left(a_k\right) \propto ar_{\rho,t}\left(m_j, a_k\right)$$

The simplicity of this learning method does not prevent it from being effective. Consider, first, the Lewis game. In simulations with every urn containing one ball of each type and payoffs equaling one, after 300 iterations, the Sender and Receiver have $\pi\left(\sigma, \rho\right) \approx 0.9$ on average.[12] In fact, for this simplest case, one can prove that Roth–Erev learning converges to a signaling system.

**Theorem 1** (Argiento et al. 2009) *In the Lewis game, with probability 1,*

$$\lim_{t\to\infty} \pi\left(\sigma_t, \rho_t\right) = 1$$

*Moreover, the two signaling systems are equally likely to occur: with probability 1/2,*

$$0 = \lim_{t\to\infty} \frac{\sigma_t\left(s_1, m_1\right)}{\sigma_t\left(s_1, m_2\right)} = \lim_{t\to\infty} \frac{\sigma_t\left(s_2, m_2\right)}{\sigma_t\left(s_2, m_1\right)} = \lim_{t\to\infty} \frac{\rho_t\left(m_1, a_1\right)}{\rho_t\left(m_1, a_2\right)} = \lim_{t\to\infty} \frac{\rho_t\left(m_2, a_2\right)}{\rho_t\left(m_2, a_1\right)}$$

*while with probability 1/2, the limits of the reciprocals are 0.*

When considering the Atomic $n$-game for larger $n$, simulations have somewhat mixed results. Barrett (2006) finds that for $n = 3, 4$, after $10^6$ iterations, agents have $\pi\left(\sigma, \rho\right) > 0.8$ at rates of .904 and .721, respectively. For $n = 8$, that number drops to .406. Nevertheless, the agents always perform significantly better than chance. I will present my own simulation results for this situation later and so do not dwell on them

[10] That seminal paper has spanwed a large body of research. See Schultz (2004) and Glimcher (2011) for overviews.

[11] For instance, the ball-in-urn metaphor essentially assumes that the utility function only has integer values.

[12] See Skyrms (2010), p. 94.

here. Analytically, Hu et al. (2011) show that there is a large class of non-strict Nash equilibria of the Atomic $n$-game (the so-called partial pooling equilibria) to which Roth–Erev learning converges with positive probability.[13]

## 3 The Negation Game

Having seen the basics of signaling games and reinforcement learning, we are in a position to use these tools to answer our original question: *why* would compositional languages arise? To address this question, we first need to extend the Lewis–Skyrms signaling games to handle rudimentary forms of compositional signaling. In particular, we will focus on having a signal corresponding to *negation*.

To get an intuition for how the model will work, consider vervet monkeys.[14] These monkeys have three predators: leopards, eagles, and snakes. It turns out that they have three acoustically distinct signals—a bark, a cough, and a chutter—that are typically sent only when a predator of a particular type is present. Moreoever, when a vervet receives one of the signals, it appears to respond in an adaptive way: if it hears a bark, it runs up a tree; if a cough, it looks up; if a chutter, it looks down. Schematically, their behavior looks like:

$$\text{leopard} \longrightarrow \text{bark} \longrightarrow \text{climb tree}$$

$$\text{eagle} \longrightarrow \text{cough} \longrightarrow \text{look up}$$

$$\text{snake} \longrightarrow \text{chutter} \longrightarrow \text{look down}$$

As the diagram above suggests, a signaling system in the Atomic 3-game provides a natural model for this signal/response behavior.

More speculatively, suppose that a vervet on the ground saw that a group-mate had run up a tree as if a leopard were present. The vervet on the ground, however, knows that no leopard is present. It would be useful if it could signal such a fact to the monkey in the tree so that it would come down and the group could get on with its business. To signal that no leopard was present, the vervet on the ground could attempt to use a completely new signal. Alternatively, it could send a signal along the lines of “no bark”, relying on the fact that the other vervets already know how to respond to barks. This latter solution seems *prima facie* superior since it leverages the existing signaling behavior. Moreover, Zuberbühler (2002) finds a syntax of exactly this simple kind in Campbell monkeys: they have alarm calls for leopards and eagles, which can be prefixed with a low ‘boom boom’ sound to indicate that the threat is not immediate.[15]

[13] Similar results hold for the closely related replicator dynamics. See Huttegger (2007), Pawlowitsch (2008). Hofbauer and Huttegger (2008, 2015) study the replicator-mutator dynamics for these games.

[14] See Seyfarth et al. (1980).

[15] See Schlenker et al. (2014) for a detailed semantic analysis of this form of signaling. They in fact find that the basic alarm calls have roots with a morphological suffix.

That some monkeys have such syntax makes it plausible that the following model will fit well with theories of the gradual emergence of compositionality.

To model signaling behavior of this sort, let us introduce *the negation game*. For a given $n$, there are $2n$ states and acts with the same utility function as in the Atomic $2n$-game. Now, however, instead of all messages being simple/atomic, the Sender can also send a *sequence* of two signals. The Sender has $n$ "basic" signals $m_1, \ldots, m_n$ and can also send signals of the form $\boxminus m_i$ for some new signal $\boxminus$ and $1 \leq i \leq n$. Now, in the extended vervet case, and with negation in general, there are certain logical relations among the states. The model here will capture a few basic intuitions about the relations imposed by negation:

- Every state has a negation.
- The negation of a state is distinct from the state.
- Distinct states have distinct negations.

The relevant mathematical notion capturing these intuitions is that of a *derangement*: this is a bijection (a one-to-one and onto function) with no fixed points, i.e., no $x$ such that $f(x) = x$.[16] So, writing $[n] = \{1, \ldots, n\}$, we also assume that we are equipped with a derangement $f : [2n] \to [2n]$.[17] We will consider $f$ as a function on both the states and acts by writing $f(s_i) := s_{f(i)}$ and $f(a_i) := a_{f(i)}$. Two examples of state spaces with derangements may be illuminating.

*Example 1* There are six states: for each of a leopard, snake, and eagle, one state indicates the presence and another the absence thereof. The function sending presence states to the corresponding absence state and *vice versa* is a derangement, depicted here:

$$\begin{tikzcd} \text{leopard} \arrow[r, bend left] & \overline{\text{leopard}} \arrow[l, bend left] \end{tikzcd}$$

$$\begin{tikzcd} \text{snake} \arrow[r, bend left] & \overline{\text{snake}} \arrow[l, bend left] \end{tikzcd}$$

$$\begin{tikzcd} \text{eagle} \arrow[r, bend left] & \overline{\text{eagle}} \arrow[l, bend left] \end{tikzcd}$$

*Example 2* Let $W$ be a set of *worlds*. Then $\mathcal{P}(W)$, the set of subsets of $W$, is the set of propositions based on those worlds. *Set complement* is a derangement on $\mathcal{P}(W)$, often used to specify the meaning of negation: $[\![\neg\varphi]\!] = W \setminus [\![\varphi]\!]$.

These examples show how the notion of a derangement captures the minimal core of negation in terms of the three intuitions above. Because of this, I will use the term

[16] See Hassani (2003) for the definition and applications.

[17] To fully model classical logical negation, we would also need to require that $f$ is an *involution*, i.e. that $f(f(i)) = i$. But since our syntax is so impoverished that sending a "double negation" signal is not possible, we can omit this requirement. It is also doubtful that natural language negation satisfies double negation elimination. This remark about involutions does, however, explain why we have $2n$ states and acts: a permutation that is an involution will only have cycles of length $\leq 2$. And being a derangement requires that there are no cycles of length 1. Together, this means that $[n]$ only has derangements which are involutions if $n$ is even.

'minimal negation' in this context. The question now becomes: how can the Sender and Receiver exploit this structure to communicate effectively?

First, consider the Sender. For simplicity, let us suppose that this is a pure sender, so that it always sends one signal in each state. Without loss of generality, we can even suppose that it sends $m_i$ in $s_i$. Nature might inform the Sender, however, that state $s_k$ obtains for some $k > n$. What signal, then, should it send? We can use the derangement $f$ which captures the idea of minimal negation. For some $i$, $s_k = f(s_i)$. The Sender will therefore send $\boxminus m_i$. To put it more formally in order to foreshadow the generalization to a Sender using a mixed strategy, $\sigma(f(s_i)) = \boxminus\sigma(s_i)$.

What about the Receiver? It will "interpret" $\boxminus$ as a minimal negation in the following way. When it receives a signal $\boxminus m_i$, it looks at the act it would take in response to $m_i$. Suppose that's $a_i$. Instead, however, of taking that act, it performs $f(a_i)$. Again, to put it formally: $\rho(\boxminus m_i) = f(\rho(m_i))$. Here already we see the rudiments of compositionality: a complex signal is 'interpreted' as a function of the interpretation of its part. Now, consider a Sender and Receiver playing in this way. If $s_k$ obtains, the Sender sends $\boxminus m_i$. The Receiver then will play $f(\rho(m_i)) = f(a_i) = a_k$, which is the appropriate act.

This simple model shows how sending signals corresponding to minimal negations can enable the Sender and Receiver to achieve communicative success in a large and logically structured state space. As presented, however, we required that both agents played pure strategies and that they had an agreement on the meanings of the basic signals. We can generalize the above definitions to allow for mixed Sender and Receiver strategies. Once we have done that, we can use Roth–Erev reinforcement to try and *learn* those strategies.

**Definition 5** (*Negation n-game*) A *Negation n-game* is a tuple $\langle S, M, A, \sigma, \rho, u, P, f\rangle$ where $|S| = |A| = 2n$, $f : [2n] \to [2n]$ is a derangement, and $u\left(s_i, a_j\right) = 1$ iff $i = j$, 0 otherwise. $M = \{m_1, \ldots, m_n\} \cup \{\boxminus m_i : 1 \le i \le n\}$. Moreover, we require that:[18]

$$\sigma\left(f\left(s_j\right)\right)(\boxminus m_i) \propto \sigma\left(s_j\right)(m_i) \tag{2}$$

$$\rho(\boxminus m_i)\left(f\left(a_j\right)\right) = \rho(m_i)\left(a_j\right) \tag{3}$$

Payoffs are given by (1) as before.

We can understand these constraints in terms of the urn model described above. Consider, first, the Sender. Now, the Sender's urns also contain balls labelled $\boxminus$. If one of these balls is drawn from an urn (say, $s_j$), then the Sender looks at the urn for $f^{-1}\left(s_j\right)$, draws a non-$\boxminus$ ball (say, $m_i$) and then sends $\boxminus m_i$. This Sender behavior exactly implements the constraint (2). The Receiver's behavior is even simpler: for signals of the form $m_i$, it simply draws a ball from the appropriate urn as before.

[18] Or equivalently:

$$\sigma\left(s_j\right)(\boxminus m_i) \propto \sigma\left(f^{-1}\left(s_j\right)\right)(m_i)$$

$$\rho(\boxminus m_i)\left(a_j\right) = \rho(m_i)\left(f^{-1}\left(a_j\right)\right)$$

When, however, encountering a signal $\boxminus m_i$, the Receiver draws a ball from the $m_i$ urn (labelled, say, $a_k$) and then performs $f(a_k)$. This Receiver behavior exactly implements the constraint (3). In this way, the Sender uses and the Receiver interprets $\boxminus$ as minimal negation. Given this model of behavior, implementing Roth–Erev reinforcement learning for the Negation $n$-game is as simple as reinforcing all the choices made in a given iteration. That is, we perform the following reinforcements:

$$\begin{aligned} ar_{\sigma,t+1}\left(s_j, \boxminus\right) &= ar_{\sigma,t}\left(s_j, \boxminus\right) + u\left(s_j, f(a_k)\right) \\ ar_{\sigma,t+1}\left(f^{-1}\left(s_j\right), m_i\right) &= ar_{\sigma,t}\left(f^{-1}\left(s_j\right), m_i\right) + u\left(s_j, f(a_k)\right) \\ ar_{\rho,t+1}\left(m_i, a_k\right) &= ar_{\rho,t}\left(m_i, a_k\right) + u\left(s_j, f(a_k)\right) \end{aligned}$$

Learning in the Negation $n$-game provides a model to answer the following question: given the ability to use a signal to mean minimal negation, can the Sender and Receiver learn how to use the atomic words to communicate effectively?

## 4 Experiment

We are now in position to test whether compositional languages are in some sense better. In particular, we will ask whether such languages are easier to learn than languages with only atomic signals. Consider again Example 1, where the vervets want to communicate about both the presence and absence of leopards, snakes, and eagles. *Prima facie*, it seems like having minimal negation around would make learning easier: once signals for the three predators are known, the signals for their absence are also automatically known by prefixing with the negation signal. By contrast, with only atomic signals, three new unrelated signals would need to be introduced to capture the states corresponding to the lack of each predator. Of course, similar thoughts apply for more or fewer than three predators, so I will compare atomic versus compositional languages when there are different numbers of states.

### 4.1 Methods

To test the hypothesis that compositional languages will be easier to learn, I ran 100 trials of Roth–Erev learning for each of the Atomic $2n$-game and the Negation $n$-game for each $n \in \{2, 3, 4, 5, 6, 7, 8\}$. The initial value of accumulated rewards was 1 for every argument. Similarly, $u\left(s_i, a_j\right) = 1$ when $i = j$. Each trial consisted of 10,000 iterations. This relatively low value was chosen so that differences in speed of learning may also be apparent. I measured $\pi(\sigma, \rho)$ for each trial at the end of the iterations of learning. The code for running these simulations and performing the data analysis, in addition to the actual data, may be found at http://github.com/shanest/learning-evolution-compositionality. Let $\textsc{Atom}_{2n}$ (respectively, $\textsc{Neg}_n$) refer to the set of 100 such payoffs of the Atomic $2n$-game (respectively, Negation $n$-game). I will use $\overline{(\cdot)}$ to refer to the mean of a set of values.

**Table 1** Means and $t$ tests for Roth–Erev reinforcement learning

| $n$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|
| $\overline{\text{ATOM}_{2n}}$ | 0.914 | 0.863 | 0.819 | 0.758 | 0.699 | 0.617 | 0.506 |
| $\overline{\text{NEG}_n}$ | 0.851 | 0.816 | 0.786 | 0.746 | 0.706 | 0.661 | 0.605 |
| $\text{DIFF}_n$ | −0.064 | −0.046 | −0.033 | −0.013 | 0.007 | 0.044 | 0.099 |
| $t$ | 3.666 | 3.461 | 2.897 | 1.117 | −0.724 | −4.825 | −10.88 |
| $p$ | 0.0003 | 0.0007 | 0.0042 | 0.2654 | 0.4698 | 0.000003 | 0.6e−22 |

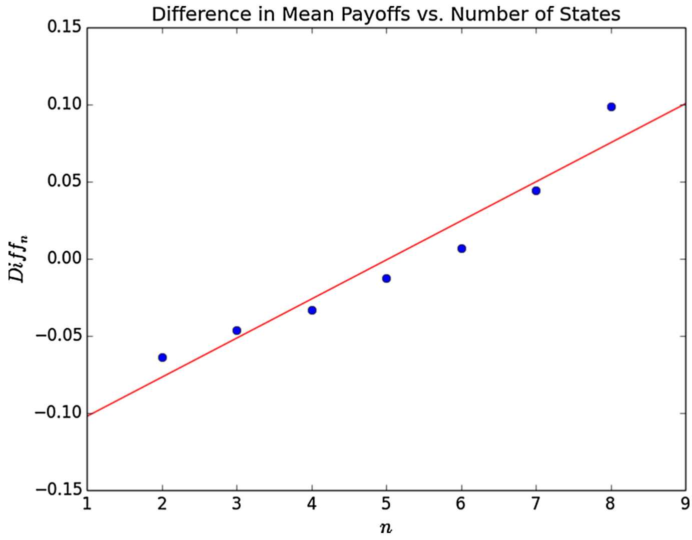


**Fig. 1** Fitting a line to $\text{DIFF}_n$

## 4.2 Results

To measure the effectiveness of Roth–Erev learning, the mean payoffs of the trials of each game for each $n$ were computed. Welch's $t$ test was used to test the hypothesis that those means are different. These results are summarized in Table 1, where $\text{DIFF}_n = \overline{\text{NEG}_n} - \overline{\text{ATOM}_{2n}}$.

To measure the effect, if any, of $n$ on the effectiveness of Roth–Erev learning in the two types of game, a simple linear regression of $\text{DIFF}_n$ on $n$ was run. This yielded $r = 0.9658$ and $p = 0.00041$, indicating a very strong positive linear correlation. Figure 1 shows the line of best fit overlayed on the data.

## 4.3 Discussion

There are two main take-aways from these results. First, having a compositional language does not always make learning easier. To see this, note that a negative value of $\text{DIFF}_n$ means that agents learn to communicate more successfully with the atomic language than with the minimal negation language for that value of $n$. This occurs for $n = 2, 3, 4, 5$. Moreover, the $p$ values from the Welch's $t$ test show that for $n = 2, 3, 4$, this difference in mean success is in fact statistically significant. Thus, when there are not many states (4, 6, and 8 respectively) for the agents to distinguish, an atomic language is easier to learn. By the time we reach $n = 7$ (14 states), however, the agents using a compositional language do perform statistically significantly better, as evidenced by the positive value for $\textsc{Diff}_7$ and the very low $p$ value there. This trend also continues for the $n = 8$ (16 state) case.

This leads to the second observation: the very strong positive correlation between $n$ and $\textsc{Diff}_n$ ($r = 0.9658$) shows that it becomes increasingly more advantageous from a learning perspective to have a compositional language as the number of states increases. Thinking of the number of states as a proxy for the complexity of the world about which the agents must communicate, we can summarize these two results in a slogan: compositional signaling can help simple agents learn to communicate in a complex world.

Taken together, these results present a precise model identifying one ecological pressure that could explain the evolution of compositionality: the need to communicate about a large number of situations. Moreover, the surprising fact that compositional languages only confer a learning benefit in large state spaces shows another sense in which the crude evolutionary arguments from the introduction are too unconstrained. Recall that premise (2') claimed that, for example, productive understanding would have been 'good' for our ancestors. Perhaps one could even have identified the sense of goodness with learnability. These simulations, however, show that the situation is not so simple: the sense in which a compositional language is good depends very strongly on the size of the state space. One could not have figured that out without analysis of the kind presented here.

Note that this experiment examined only the most rudimentary form of compositional semantics. Nevertheless, the fact that this form only exhibited advantages in a suitably complex world suggests that even more sophisticated languages (with, for instance, binary operators and a partially recursive syntax) will only be advantageous when the world in which the agents are embedded also is substantially more complex. Moreover, that compositionality only confers an advantage in large state spaces may explain why it is rare in the animal kingdom: many species may not have pressure to talk about many different situations.

## 5 Learning Negation

While the above experiment shows how having a rudimentary form of compositional signaling can be beneficial, it leaves open the question: where did the compositional signaling come from? That is, can the agents *learn* to use a signal as negation?

To model the learning of negation itself, we need to relax the definition of the Negation $n$-game so that the Sender must learn how to use $\boxminus$ and the Receiver must choose how to interpret the signal $\boxminus$. Instead of putting one derangement in the model, we will put in a whole set of functions $\mathcal{F}$. The Receiver will have distributions over the acts for each basic signal and a distribution over $\mathcal{F}$ for $\boxminus$. Its choice on how to interpret $\boxminus m_i$ will directly refer to this latter distribution which reflects how it will interpret $\boxminus$ as a function.

**Definition 6** (*Functional n-game*) A *Functional n-game* is a tuple $\langle S, M, A, \sigma, \rho, u, \mathcal{F}\rangle$ where $S, M, A$, and $u$ are as in the definition of a Negation $n$-game (see Definition 5). $\mathcal{F}$ is a set of functions $[2n] \to [2n]$ (not necessarily bijections). We require that exactly one $g \in \mathcal{F}$ is a derangement. For the Receiver, we require that

$$\rho\left(\boxminus m_i\right)\left(a_j\right) = \sum_{g\in\mathcal{F}} \rho\left(\boxminus\right)(g) \cdot \sum_{k:g(k)=j} \rho\left(m_i\right)\left(a_k\right) \tag{4}$$

Intuitively, this constraint captures the following behavior. The Sender no longer uses $\boxminus$ explicitly as minimal negation; its behavior is unspecified. Now, instead of the Receiver automatically interpreting $\boxminus$ as minimal negation, it first *chooses* a function by which to interpret it. In terms of the urn models, the Receiver now has another urn labeled $\boxminus$. Balls in this urn, however, are labeled with functions $g \in \mathcal{F}$. When the Receiver receives a signal $\boxminus m_i$, it draws a function $g$ from the $\boxminus$ urn, an act $a_i$ from the $m_i$ urn, and performs $g(a_i)$. Constraint (4) exactly captures that behavior. Note that when $\mathcal{F}$ is a singleton containing a lone derangement $f$, Constraint (4) reduces to Constraint (3) because $\rho\left(\boxminus\right)(f) = 1$ and there will be only one $k$ such that $f(k) = j$ since $f$ is a bijection.

To look at the emergence of minimal negation, we want to capture a natural idea about how agents would learn to use a function word: they are already capable of communicating with atomic signals and then try to *introduce* the functional element. To model this kind of situation, we need a few more definitions. As before, let $S$ be a set and $f$ a derangement on $S$. Call a set $X \subset S$ *complement-free* iff $X \cap f[X] = \emptyset$. If $X$ is a maximal complement-free subset (in the sense of not being contained in another complement-free set), we will call it a *complementizer* of $S$. As an example of a complementizer, note that both the sets {leopard, eagle, snake} and its complement are complementizers with respect to the derangment mentioned in Example 1 above. Note that if $|S| = 2n$, then all complementizers of $S$ have cardinality $n$. For a set $S$ of size $n$ and a subset $I \subseteq [n]$, write $S \upharpoonright I = \{s_i : i \in I\}$.

The situation we are interested in is the following: the agents are playing a Functional $n$-game. For some complementizer $X$ of $[2n]$ with respect to the derangement $f \in \mathcal{F}$, $\sigma$ and $\rho$ in addition constitute a signaling system on the subgame generated by $S \upharpoonright X$ and $A \upharpoonright X$ in which $\sigma$ only sends basic signals from $\{m_1, \ldots, m_n\}$. In this setting, the Sender needs to choose when to send complex signals $\boxminus m_i$ and the Receiver needs to choose how to interpret $\boxminus$. We will suppose that the Receiver chooses between a few natural responses to hearing new signals with $\boxminus$. The Receiver might (a) ignore $\boxminus$, (b) treat $\boxminus$ as a new atomic word, or (c) treat $\boxminus$ as minimal negation. These options can be modeled by different functions: (a) corresponds to the identity function $id(j) = j$, (b) to a constant function $c_i(j) = i$ and (c) to a derangement $f$. Therefore, to model the Receiver choosing among these natural options for interpreting $\boxminus$, we assume that $\mathcal{F}$ contains exactly those three functions.

Call a Functional $n$-game with the above restrictions a *basic negation learning setup of size n*. Our question now is: does simple Roth–Erev reinforcement learning allow the Sender to start using $\boxminus$ like minimal negation and the Receiver to learn to interpret $\boxminus$ as minimal negation? We answer this question using both analytic and simulation results.

### 5.1 Analytic Result

It turns out that in this setup, we can actually *prove* that Roth–Erev learning works.

**Theorem 2** *In a basic negation learning setup of size n, if $i \in X$ for the $c_i \in \mathcal{F}$, then:*

1. *With probability 1, for the derangement $f$, $\lim_{t\to\infty} \rho_t (\boxminus) (f) = 1$. In other words, the probability that $\boxminus$ is interpreted as minimal negation by the Receiver converges to 1.*
2. *With probability 1, $\sigma$ converges to a strategy where Constraint (2) holds. In other words, with probability 1, the Sender learns to use $\boxminus$ as minimal negation.*

*Proof* Theorem 4 of Beggs (2005) states that probabilities and empirical frequencies converge to 0 for strategies which do not survive iterated removal of strictly dominant strategies (when there are a finite number of players and actions) under Roth–Erev learning.

For the first part, this means that it suffices to show that choosing the derangement $f$ is strictly dominant for the Receiver. To see this, note that $\sigma$ only sends signals of the form $\boxminus m_i$ for states in $f[X]$. If the Receiver chooses $c_i$, then, the payoff will be 0 since $i \in X$. Similarly, since $m_i$ is only sent in $s_i \in X$, if the Receiver chooses $id$, it will perform an act in $A \upharpoonright X$ and so receive a payoff of 0. Choosing the derangement is thus strictly dominant.

For the second part, recall that in a basic negation learning setup, $\sigma$ and $\rho$ are a signaling system on the subgame restricted to the complementizer $X$. This means that for $s_i \in X$, w.l.o.g., $\sigma(s_i)(m_i) = 1$. Now, consider $f(s_i)$. We must show that $\boxminus m_i$ is the only signal that survives iterated removal of dominant strategies. Suppose that $\sigma$ sends a basic signal $m_j$ in $f(s_i)$. By assumption, for some $k \in X$, $\rho\left(m_j\right) = a_k \neq f(a_i)$ since $i \in X$ and $X \cap f[X] = \emptyset$, so the payoff is 0. So no basic signal survives. Now, since we know that the Receiver will interpret $\boxminus$ as $f$, only $\boxminus m_i$ will lead to the Receiver performing $f(a_i)$ and to the Sender receiving a positive payoff. We thus have that $\sigma(f(s_i))(\boxminus m_i)$ converges to 1, which yields (2) above because $\sigma(s_i)(m_i) = 1$. $\square$

This proof depends on the assumption that $i \in X$. If $i \in f[X]$, then choosing $f$ is only *weakly* dominant over choosing $c_i$ and so the result of Beggs (2005) no longer applies. For this reason, and because the above convergence result does not tell us about *rate* of convergence, I also ran simulations of the basic negation learning setup.

### 5.2 Simulation Results

I ran 100 trials of only 1000 iterations of the basic negation learning setup for each $n \in \{2, 3, 4, 5, 6, 7, 8\}$ both with $i \in X$ and $i \in f[X]$. Table 2 shows the average payoffs and probability of interpreting $\boxminus$ as negation after the 100 trials.

**Table 2** Statistics for Roth–Erev reinforcement in the basic negation learning setup

| | $n$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| $i \in X$ | $\pi(\sigma, \rho)$ | 0.995 | 0.989 | 0.978 | 0.962 | 0.937 | 0.899 | 0.855 |
| | $\rho(\boxminus)(f)$ | 0.995 | 0.995 | 0.995 | 0.994 | 0.993 | 0.993 | 0.991 |
| $i \in f[X]$ | $\pi(\sigma, \rho)$ | 0.904 | 0.875 | 0.870 | 0.818 | 0.821 | 0.756 | 0.701 |
| | $\rho(\boxminus)(f)$ | 0.816 | 0.844 | 0.807 | 0.666 | 0.922 | 0.760 | 0.830 |

These results show two things. First, learning happens *very fast* in this setup. This holds especially true for the Receiver, who has uniformly high values for $\rho(\boxminus)(f)$ in the $i \in X$ case. The Sender appears to need more time to fully learn its strategy in the large state spaces. Second, having $i \in f[X]$ does result in lower values for both the payoff and for $\rho(\boxminus)(f)$. Moreover, $t$ tests show that all of these differences are in fact statistically significant. This provides some initial reason to doubt that Theorem 2 generalizes to the case when $i \in f[X]$. Nevertheless, the Theorem and these simulations show that the task of learning the meaning of a function word when the meanings of atoms are known is rather easy even for very simple learners.

## 6 Comparison to Other Work

While the present paper does present a precise answer to the question of why natural languages are compositional, it is not the first to attempt to do so. Therefore, I will now discuss a few earlier proposals from the literature. In each case, I find that something distinctive about the nature of function words has been left out.

Nowak and Krakauer (1999) study the emergence of compositionality via a kind of natural selection of signal-object pairings. For our purposes, the most interesting case is their last, where they consider a state space consisting of pairs of two objects and two properties, for four total combinations. These combinations can be specified by four atomic words $w_1, w_2, w_3, w_4$ or with pairs $p_i o_j$. Nowak and Krakauer consider a strategy space where players use the atomic words with probability $p$ and the 'grammatical' constructs with probability $1 - p$. They are able to show that the only two evolutionary stable strategies are when $p = 0$ and $p = 1$ and that their evolutionary dynamics evolves to use the grammatical rule with probability 1. In trying to use syntactic structure to mirror structure in the state space, the present approach does have something in common with Nowak and Krakauer's. There are, however, two reasons that their proposal does not go far enough. First, they only analyze mixtures of a single atomic and a single compositional signaling system but do not analyze whether one or the other form of signaling makes it easier to arrive at a signaling system. One would like a dynamic that considers more than two possible signal-state mappings. Secondly, they are not explicitly interested in function words but only in subject-predicate structure. Both models are needed for a full story of the evolution of complex signaling: their emphasis is on the structure of minimally meaningful signals, while mine is on how function words can be used to modify the meanings of already meaningful signals.

In a series of papers, Barrett (2006, 2007, 2009) considers 'syntactic' signaling games with reinforcement learning where the number of states and acts exceeds the number of signals but where there is more than one sender. These are equivalent to games in which there is one sender who sends a sequence of signals for each state (the length of the sequence is fixed and the strategies are independent for each position in the sequence). Perfect signaling is achieved when sender-receiver strategies settle on a coding system with a sequence that elicits the ideal act in each state. For example, consider the 4-state case with 2 senders, each of which has 2 signals. An example of optimal signaling behavior has Sender 1 use $m_1$ in $\{s_1, s_2\}$ and $m_2$ in the other two states, while Sender 2 uses $m_1$ in $\{s_1, s_3\}$ and $m_2$ in the other two states. If the Receiver can take the intersection of the sets of states indicated by the two messages received, it can learn the exact state and so perform the correct act. Barrett found that in simulations involving 4-state, 4-act, 2-signal, 2-sender signaling games, successful signaling[19] is achieved by reinforcement learning approximately 3/4 of the time and that in 8-state, 8-act, 2-signal, 3-sender games, successful signaling is achieved 1/3 of the time.

The difference, then, between this approach and our own is that there is no explicit signal that serves as a function word in Barrett's set-up; rather, signal concatenation is implicitly treated like conjunction. Of course, natural language contains many such function words, meaning that his model cannot provide a full account of the emergence of natural language compositionality. Moreover, as Franke (2014) observes, the Receiver in a syntactic game appears to not actually interpret length-2 signals compositionally: "Although we can describe the situation as one where the meaning of a complex signal is a function of its parts, there is no justification for doing so. A simpler description is that the receiver has simply learned to respond to four signals in the right way" (p. 84). In other words, the Receiver's dispositions treat the complex signals as if they were atomic. Nevertheless, there is some similarity between the present paper and Barrett's work. The task in Experiment 1 can now be recast as the Sender having to learn a complementizer $X$ on the state space. The simulation results show that this is actually a somewhat harder task than learning two complementary partitions.

A final proposal comes from Franke (2014) where the explicit concern is with using reinforcement learning in signaling games to explain compositional meanings. Franke adds 'complex' signals of the form $m_{AB}$ (as well as 'complex' states of the form $s_{AB}$). Formally, these are just new atomic signals. The sense in which they bear a relation to the basic signals is captured in a distance $s = d(m_{AB}, m_A) = d(m_{AB}, m_B)$ which intuitively represents a kind of similarity (and similarly for the set of complex states of the form $s_{AB}$). Franke modifies Roth–Erev learning by incorporating (i) spill-over and (ii) lateral inhibition. Spill-over refers to the fact that non-actualized message/state pairs (for the sender) are reinforced in proportion to their similarity to the played one (and similarly for the receiver). Lateral inhibition involves lowering the accumulated rewards for non-actualized pairs when the actualized pair was successful. Franke shows that with these two modifications, creative compositionality can arise in the following sense: if Sender/Receiver play only using 'basic signals', they can still learn to have

[19] For Barrett, this means $\pi(\sigma, \rho) > 0.8$.

$\sigma(s_{AB})(m_{AB}) > \sigma(s_{AB})(m_j)$ for all $m_j \neq m_{AB}$ and so on for the other complex states and signals. Since the 'complex' signal is the one most likely to be sent when the 'complex' state is seen for the first time, this purportedly captures creativity: using a new signal for a new state.[20]

There are two main differences between Franke's work and the present approach. First, in his model, signals do not have genuine syntactic structure which then gets compositionally interpreted. One does not find complex signals with words treated like function words as one does in the Negation $n$-game. Secondly, his model claims to provide an explanation for the *emergence* of compositional signaling, whereas the present paper has simply identified an evolutionary advantage that compositional signaling would confer. One would like a model that incorporates both of these aspects: a story about the emergence of genuinely syntactically structured signals with compositional interpretation.

## 7 Conclusion and Future Directions

In this paper, I have offered a potential story that answers the question of *why* natural languages are compositional by enriching signaling games with the most rudimentary form of compositionality and exploring basic learning of such languages. The current answer is that compositional languages help simple agents learn to communicate more effectively in a complex world. I also demonstrated that learning the meaning of a function word after knowing the meanings of atomic words is not difficult.

While these two strands present a promising start to answering the *why*-question, much more work remains to be done in the future. Firstly, there are areas of the parameter space of the current games that remain to be explored. For instance, in what ways are the simulation results robust against the form of the utility function? Is there an independently motivated but wide class of such functions for which the results still hold? Secondly, one might wonder about other and more complex forms of compositionality. Can different languages be compared in order to find out whether different kinds of compositional language are better in different circumstances? Similarly, it is natural to continue extending the approach here to languages with multiple logical operators of varying arities (conjunction, for instance) and to start adding a more recursive syntax. Finally, considering such richer forms of compositional signaling might also require exploring more sophisticated learning algorithms than the very simple reinforcement learning considered here.

**Acknowledgments** I would like to thank the organizers and participants at IDAS'14 and especially Stefano Demichelis and Roland Muehlenbernd for stimulating discussion. An earlier version was presented at a workshop on Knowledge, Argumentation, and Games in Amsterdam. For discussion there, I thank especially Michael Franke and Alexandru Baltag. Jeffrey Barrett, Johan van Benthem, Thomas Icard, Chris Potts, Carlos Santana, Brian Skyrms, and Michael Weisberg have provided helpful comments on earlier versions of this paper as did three anonymous referees for this journal.

[20] Though it's worth noting that $\sigma(s_{AB})(m_{AB}) \leq 1/2$ so that $m_{AB}$ never becomes strongly associated with $s_{AB}$.

## References


Argiento, R., Pemantle, R., Skyrms, B., & Volkov, S. (2009). Learning to signal: Analysis of a micro-level reinforcement model. *Stochastic Processes and their Applications*, *119*(2), 373–390. doi:10.1016/j.spa.2008.02.014.

Barrett, J. A. (2006). Numerical simulations of the Lewis signaling game: Learning strategies, pooling equilibria, and the evolution of grammar. Technical report, Institute for Mathematical Behavioral Sciences, 2006. http://escholarship.org/uc/item/5xr0b0vp

Barrett, J. A. (2007). Dynamic partitioning and the conventionality of kinds. *Philosophy of Science*, *74*(4), 527–546.

Barrett, J. A. (2009). The evolution of coding in signaling games. *Theory and Decision*, *67*(2), 223–237. doi:10.1007/s11238-007-9064-0.

Beggs, A. W. (2005). On the convergence of reinforcement learning. *Journal of Economic Theory*, *122*(1), 1–36. doi:10.1016/j.jet.2004.03.008.

Bergstrom, C. T., & Dugatkin, L. A. (2012). *Evolution*. New York: W. W. Norton.

Davidson, D. (1964). Theories of meaning and learnable languages. In *Inquiries into truth and interpretation* (pp. 3–16). Oxford: Oxford University Press. doi:10.1093/0199246297.001.0001

Fodor, J. (1987). *Psychosemantics*. Cambridge, MA: The MIT Press.

Fodor, J. (1998). *In critical condition: Polemical essays on cognitive science and the philosophy of mind*. Cambridge, MA: The MIT Press.

Fodor, J., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture. *Cognition*, *28*(1–2), 3–71. doi:10.1016/0010-0277(88)90031-5.

Franke, M. (2014). Creative compositionality from reinforcement. In E. A. Cartmill, S. Roberts, H. Lyn, & H. Cornish (Eds.), *The evolution of langauge (Proceedings ofEvoLang 10)* (pp. 82–89). Singapore: World Scientific.

Frege, G. (1923). Logische Untersuchungen. Dritter Teil: Gedankengefüge ("Compound Thoughts"). *Beiträge zur Philosophie des deutschen Idealismus*, *III*, 36–51. doi:10.1093/mind/LI.202.200.

Futuyma, D. J. (2010). Evolutionary constraint and ecological consequences. *Evolution*, *64*(7), 1865–1884. doi:10.1111/j.1558-5646.2010.00960.x.

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences*, *108*(42), 15647–15654. doi:10.1073/pnas.1014269108.

Haldane, J. B. S. (1932). *The causes of evolution*. London: Longmans, Green, and co.

Hassani, M. (2003). Derangements and applications. *Journal of Integer Sequences, 6*. http://www.emis.ams.org/journals/JIS/vol6.html.

Hofbauer, J., & Huttegger, S. M. (2008). Feasibility of communication in binary signaling games. *Journal of Theoretical Biology*, *254*(4), 843–849. doi:10.1016/j.jtbi.2008.07.010.

Hofbauer, J., & Huttegger, S. M. (2015). Selection-mutation dynamics of signaling games. *Games*, *6*(1), 2–31. doi:10.3390/g6010002.

Hu, Y., Skyrms, B., & Tarrès, P. (2011). Reinforcement learning in signaling game. http://arxiv.org/abs/1103.5818

Huttegger, S. (2007). Evolution and the explanation of meaning. *Philosophy of Science*, *74*(1), 1–27. doi:10.1086/519477.

Janssen, T. M. V. (1997). Compositionality. In: J. van Benthem, A. ter Meulen (eds.), *Handbook of logic and language*, chapter 7, pp. 417–473. Elsevier Science. doi:10.1016/B978-044481714-3/50011-4.

Lewis, D. (1969). *Convention*. Oxford: Blackwell.

Nowak, M. A., & Krakauer, D. C. (1999). The evolution of language. *Proceedings of the National Academy of Sciences of the United States of America*, *96*, 8028–8033.

Pagin, P., & Westerståhl, D. (2010a). Compositionality I: Definitions and variants. *Philosophy Compass*, *5*(3), 250–264. doi:10.1111/j.1747-9991.2009.00228.x.

Pagin, P., & Westerståhl, D. (2010b). Compositionality II: Arguments and problems. *Philosophy Compass*, *5*(3), 265–282. doi:10.1111/j.1747-9991.2009.00229.x.

Pawlowitsch, C. (2008). Why evolution does not always lead to an optimal signaling system. *Games and Economic Behavior*, *63*(1), 203–226. doi:10.1016/j.geb.2007.08.009.

Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, *8*, 164–212.

Schlenker, P., Chemla, E., Arnold, K., Lemasson, A., Ouattara, K., Keenan, S., et al. (2014). Monkey semantics: two 'dialects' of Campbell's monkey alarm calls. *Linguistics and Philosophy*, *37*(6), 439–501. doi:10.1007/s10988-014-9155-7.

Schultz, W. (2004). Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology. *Current Opinion in Neurobiology*, *14*(2), 139–47. doi:10.1016/j.conb.2004.03.017.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593–1599.

Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behavior*, *28*, 1070–1094.

Skyrms, B. (1996). *Evolution of the social contract*. Cambridge, MA: Cambridge University Press.

Skyrms, B. (2004). *The stag hunt and the evolution of social structure*. Cambridge, MA: Cambridge University Press.

Skyrms, B. (2010). *Signals: Evolution, learning, and information*. Oxford: Oxford University Press.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Bradford: The MIT Press.

Szabó, Z. G. (2013). Compositionality. The Stanford Encyclopedia of Philosophy, Fall. http://plato.stanford.edu/archives/fall2013/entries/compositionality/

Trapa, P. E., & Nowak, M. A. (2000). Nash equilibria for an evolutionary language game. *Journal of Mathematical Biology*, *41*(2), 172–188. doi:10.1007/s002850070004.

von Humboldt, W. (1836). *On language: On the diversityof human language construction and its influence on the mentaldevelopment of the human species*. Cambridge Texts in the Historyof Philosophy. Cambridge University Press.

Zuberbühler, K. (2002). A syntactic rule in forest monkey communication. *Animal Behaviour*, *63*(2), 293–299. doi:10.1006/anbe.2001.1914.