Journal of Logic, Language and Information
https://doi.org/10.1007/s10849-025-09433-z

Information and Meaning in the Evolution of Compositional Signals

Travis LaCroix[1]

Accepted: 18 April 2025

Abstract

This paper provides a formal treatment of the argument that syntax alone cannot give rise to compositionality in a signalling game context. This conclusion follows from the standard information-theoretic machinery used in the signalling game literature to describe the informational *content* of signals.

Keywords Signalling games · Information transfer · Communication systems · Semantic meaning · Compositionality · Reflexivity

1 Introduction

The signalling game (Lewis, 1969; Skyrms, 2010) is appealed to as a useful model for explaining the evolution of conventional meanings for arbitrary signals. However, when considering the emergence of *language*, there is a gap between the simple communication systems for which the signalling game model accounts and the linguistic communication systems of *Homo sapiens* (LaCroix, 2020b, c). Attempts to bridge this gap have focused on the evolution of *compositional* signals on the assumption that, because compositionality is an apparently unique feature of human-level linguistic communication systems, explaining of the evolution of compositional signalling would constitute significant progress toward explaining the evolution of language.

Several models have been proposed to explain the emergence of compositional signals using the signalling-game framework (Barrett, 2006, 2007, 2009; Franke, 2016; Scott-Phillips & Blythe, 2013; Steinert-Threlkeld, 2016, 2020; Barrett et al., 2020). However, these models often focus on the *syntactic* composition of individual signals. Some researchers have suggested that syntax alone cannot give rise to compositionality (Franke, 2016; Steinert-Threlkeld, 2016; LaCroix, 2020b). One under-appreciated feature of the signalling-game framework is that this model allows us to *decouple* the production of language (the sender strategy) from its interpretation. In effect, the

✉ Travis LaCroix
travis.lacroix@durham.ac.uk

[1] Department of Philosophy, Durham University, Durham, UK

syntactic structure of a complex signal is represented by the sender strategy, and the decomposition of meanings is represented by the receiver strategy.

Using an information-theoretic approach to understand the *meanings* of syntactic signals, this paper provides a formal treatment of the argument that syntactic signalling cannot be genuinely compositional—i.e., in the robust sense required for an adequate explanation of the evolution of compositional language. Section 2 offers some theoretical motivation for the importance of the problem by describing the project of language origins research and the role that compositionality and evolutionary modelling play in this research area. Section 3 provides some formal background for the main claim of this paper, introducing some concepts from information theory (3.1), and how it has been used to elucidate a notion of *semantic information*—or *informational content*—in the context of a signalling game (3.2). In Section 4, I provide a model that explains why syntactic signalling is not compositional. This model shows that the receiver still interprets syntactically complex signals atomically. Therefore, the evolution of syntactic structure does not give rise to systematicity, which is a requirement for linguistic compositionality. Section 5 concludes by considering some implications for modelling compositionality in an evolutionary context and how these implications come to bear on wider questions surrounding the evolutionary origin of language.

2 Signalling and Compositionality

Communication is ubiquitous in nature, whereas language is often taken to be unique to humans. Hence, it is an evolutionary puzzle to explain how language might have evolved in our species. Namely, language origins research seeks to understand how rich linguistic communication systems, like the ones we see in humans, could have evolved out of simpler (non-linguistic) communication systems. One of the main difficulties arising in the study of language origins is a lack of direct evidence: language does not fossilise, and we cannot go back in time to observe the actual precursors of human-level linguistic capacities. Moreover, without a concrete conception of what language is, the question of how it evolved is hopelessly ambiguous.[1]

One approach to language origins is to simplify the question of how language evolved by exploring the characteristics of linguistic communication that differentiate natural language from simpler communication systems. One of the crucial differences between communication and language that researchers often point to is the *productive capacity* or *openness* of natural languages: with a limited vocabulary and a finite set of grammatical rules, natural language allows for the production of an unlimited number of unique expressions. A principle of such productive features of natural language captures how arbitrary sounds can be combined in endless variations to form semantically meaningful and syntactically permissible units—e.g., phonemes form morphemes and words, and words form phrasal expressions and sentences. This is often referred to as the *Principle of Compositionality*, which is typically formulated as follows (Partee 1984; Kamp & Partee 1995; Szabó 2012):

[1] As Jackendoff (2010) suggests, one's theory of language origins depends upon one's theory of language.

Definition 1 *Principle of (Linguistic) Compositionality*
The meaning of a compound [complex] expression is a function of the meaning of its parts [constituents] and the ways in which they are combined [composed].[2]

Simple communication systems that arise in nature lack this property.

Explaining how compositionality could evolve out of a non-compositional communication system is taken as a proxy for explaining how natural language might have evolved from simpler precursors. An adequate explanation requires the satisfaction of (at least) two key desiderata. The first of these is lexical composition—i.e., syntax. As we will see, this is the notion that is usually targeted in evolutionary accounts of compositionality. There is an apparent adaptive advantage for combinatorial capacities in a communication system: namely, fewer elements need to be stored in memory to produce the same possible number of messages, thus allowing for more efficient communication (Nowak & Krakauer, 1999; Nowak & Plotkin, and Jansen,, 2000). In addition to lexical combination, compositionality appears to require some degree of generalisation or systematicity. The idea of systematicity, introduced by Fodor and Pylyshyn (1988), is that "the ability to entertain a given thought implies the ability to entertain thoughts with semantically related contents". Hence, systematicity captures the "function of the meaning of its parts" component of the principle of compositionality, whereas lexical combination captures the "ways in which they are combined" component.

Building on the success of game-theoretic analyses of conventional meaning (Lewis, 1969), researchers have extended the signalling game framework to an evolutionary context, describing how meaningful signals can evolve, even while positing relatively few assumptions regarding the cognitive capacities of the players.

2.1 Signalling Games

The simplest signalling game is one in which there are two players (called the Sender and Receiver), two states of the world (s_0 and s_1), two possible signals or messages (m_0 and m_1), and two possible actions (a_0 and a_1). This is referred to as a 2×2 signalling game (Skyrms, 2010). The sender observes the state and sends a signal to the receiver. The receiver observes the signal and chooses an action. Both players receive some payoff if they coordinate on states and actions. A formal definition is given in Definition 2.[3]

Definition 2 *Signalling Game*
Let $\Delta(X)$ be a set of probability distributions over a finite set X. A *Signalling Game* is a tuple,

$$\Sigma = \langle S, M, A, \sigma, \rho, u, P \rangle,$$

[2] Note that this formulation is taken to be "theory-neutral" in the sense that it requires and entails no specific commitments about, e.g., what "meanings" or "ways of combining" might actually be. This principle arises in virtually any field of study concerned with language and meaning—notably, philosophy, logic, computer science, psychology, and semantics of natural language (Janssen, 2012). For an historical overview of the principle of compositionality in the context of natural languages, see Janssen (2012); Hodges (2012).

[3] For further details, see discussion in Huttegger (2007); Steinert-Threlkeld (2016); LaCroix (2020b).

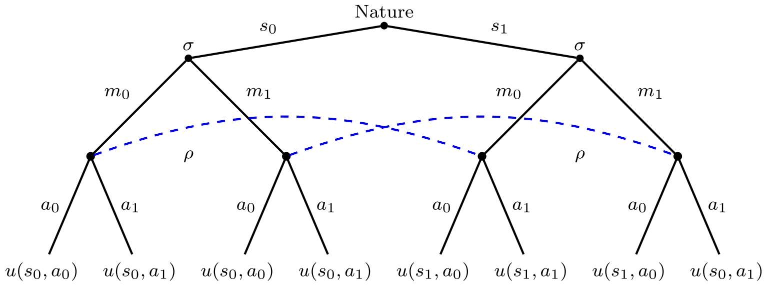


Fig. 1 The extensive form of the simple 2×2 signalling game. Each node denotes a choice point for a given player, and each branch denotes the possibilities available to her at that point. The dotted lines indicate the receiver's information set

where $S = \{s_0, \ldots, s_k\}$ is a set of *states*, $M = \{m_0, \ldots, m_l\}$ is a set of *messages*, $A = \{a_0, \ldots, a_n\}$ is a set of *acts*, with S, M, and A nonempty. $\sigma : S \to \Delta(M)$, is a function from states to a probability distribution over the set of messages that defines a *sender*, $\rho : M \to \Delta(A)$ is a function from messages to a probability distribution over actions that defines a *receiver*, $u : S \times A \to \mathbb{R}$ defines a *utility function*, and $P \in \Delta(S)$ gives a probability distribution over states in S. Finally, σ and ρ have a common *payoff*, given by

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

□

The payoff, $\pi(\sigma, \rho)$, for a particular combination of sender and receiver strategies gives an expectation of the utilities of state-act pairs (given by $u(s, a)$) weighted by the relative probability of a particular state, provided by $P(S)$. This is referred to as the *communicative success rate* of the strategies σ and ρ. The extensive form of the 2×2 signalling game is given in Fig. 1.

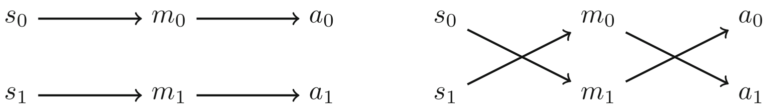
Following the notation of Steinert-Threlkeld (2016), we can introduce the further definition of an *atomic* signalling game—where states, messages, and actions are equinumerous, the utility function is 1 when the act matches the state and 0 otherwise, and nature is unbiased. See Definition 3.

Definition 3 *Atomic n-Game*
The *Atomic n-Game* is a signalling game, Σ, with the following restrictions:

(1) $|S| = |M| = |A| = n$,
(2) $u(s_i, a_j) = \delta_{ij}$, where δ_{ij} is the Kronecker delta,

$$\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ else} \end{cases},$$

and
(3) $P(s) = \frac{1}{n}$ for all $s \in S$.



(A) Signalling system 1 (B) Signalling system 2

Fig. 2 The two signalling systems of the 2×2 signalling game

□

A *signalling system* describes a situation in which the sender and receiver strategies lead to perfect coordination and maximal payoff. The atomic 2-game has exactly two signalling systems, shown in Fig. 2. Following the formal specification in Definition 2, the signalling systems of a signalling game can be defined formally as in Definition 4.

Definition 4 *Signalling Systems*
A signalling system in a signalling game is a pair (σ, ρ) of a sender and receiver that maximises $\pi(\sigma, \rho)$. □

This signalling game model is extended to an evolutionary context by specifying an evolutionary dynamic, which allows us to answer the question whether a sender and receiver (or a population of senders and receivers) can *learn* (or *evolve toward*) a signalling system.

2.2 Evolutionary Dynamics

In an evolutionary model, a dynamic explains how sender-receiver strategies (or populations) change over time. One common dynamic is simple reinforcement learning, described by the following urn-learning metaphor.[4]

We assume the sender has urns labelled s_0 and s_1. Similarly, the receiver has urns labelled m_0 and m_1. At the outset, each sender urn is equipped with one ball for each message—labelled m_0 and m_1. Similarly, each receiver urn contains a ball for each action—labelled a_0 and a_1. In each play, the state is chosen at random. The sender selects a ball at random from the urn corresponding to the state of the world and sends that message to the receiver. The receiver then chooses a ball at random from the urn corresponding to the message received. If the action matches the state of the world, then the sender and the receiver both reinforce their behaviour by returning the ball to the urn from which it was chosen and adding another ball of the same type to the urn from which the original ball was chosen. If the action does not match the state, each player returns the drawn ball to the urn from which it was drawn. The game is then repeated for a newly chosen state. See Fig. 3.

The dynamic shifts strategies to the extent that adding balls to an urn for a successful action shifts the relative probability of picking a ball of that type on a future play of the game. Adding balls to a particular urn changes the conditional probabilities of the

[4] The mean-field dynamics of simple reinforcement learning is mathematically equivalent to the replicator dynamics for population-based evolutionary models (Beggs, 2005; Hopkins & Posch, 2005).

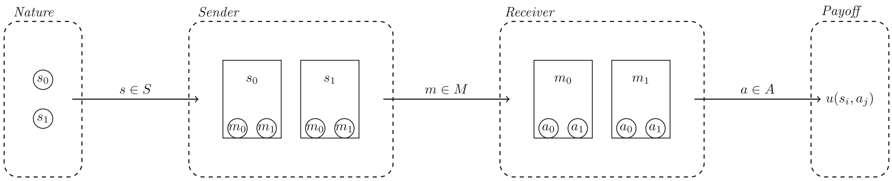


Fig. 3 Simple urn learning model for an atomic 2-game

sender's signals (conditional on the state) and the receiver's acts (conditional on the signal). Thus, the conditional probabilities of the sender's signals and the receiver's actions change over time, and the players become more likely to perform previously successful actions.

2.3 Models of Compositional Signals

The signalling-game framework gives a robust set of models for examining conditions under which we should expect simple communication to appear in nature under several different dynamics, including the reinforcement learning dynamic just described. However, simple signalling of this sort is a far cry from the complex structures present in human language. Several models seek to extend the signalling game to show how compositional signals could evolve under simple dynamics.

Signal-Object Associations Nowak and Krakauer (1999) examine how compositionality might emerge by way of natural selection on signal-object associations. Signals are interpreted as unique sounds. Each individual in the population communicates with every other individual, and rewards are summed. The rewards are interpreted as the fitness of strategy so that a higher payoff implies higher fitness.

They consider a state space consisting of pairs of objects, each with two properties, giving rise to four possible combinations. Their strategy space is constituted by the probability p that players use atomic words and the probability $1 - p$ that players use "grammatical constructions" (i.e., combinations of words). They show that $p = 0$ and $p = 1$ are the only evolutionarily stable strategies. Further, their evolutionary dynamic evolves to use the grammatical rule with probability 1. However, it is essential to note that in their discussion of the emergence of compositional language, Nowak and Krakauer (1999) only analyse whether atomic versus compositional signalling makes it easier to arrive at a signalling system; not *how* such a system might evolve.

Syntactic Signalling Barrett (2006, 2007, 2009) considers a signalling game where there are two senders, each of which can send one of two possible messages, and there are four state-act pairs. In this case, there is an informational bottleneck in the sense that no signal alone can adequately partition nature; however, the two senders together can completely partition nature.[5]

Skyrms (2010) reinterprets this situation as a signalling game in which one sender sends two signals in a particular order, giving rise to a *syntactic* signalling game. (Math-

[5] It is worth noting that the sender and receiver will learn to coordinate in a way that maximises information transfer when there are bottlenecks of this sort, although this is not implied by the dynamics alone; see discussion in LaCroix (2020a).

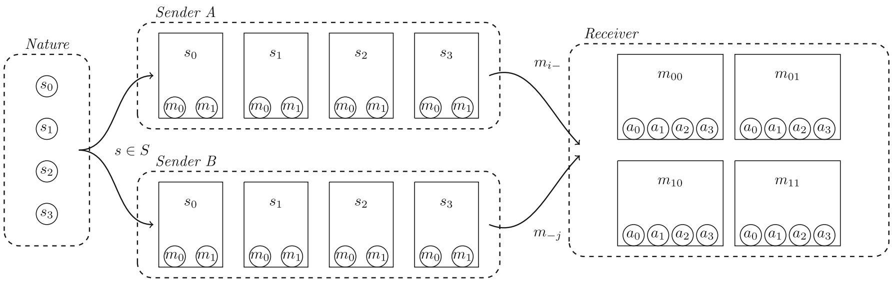


Fig. 4 Simple urn learning model for a syntactic signalling game

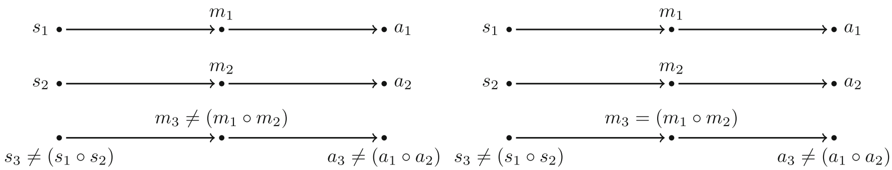
ematically, these two models are equivalent.) The sender and receiver communicate perfectly when they learn a bijective mapping between state-act pairs and sequences of signals. The receiver then needs to interpret the correct action as being given by the intersection of the two signals. See Fig. 4.

Spill-Over Reinforcement and Lateral Inhibition Franke (2014, 2016) uses *spill-over reinforcement* (similar to a mechanism in O'Connor (2014)) and *lateral inhibition* (similar to a mechanism in Steels (1995)) in a model of simple reinforcement learning to try to give an account of what he calls *creative* compositionality. Here, we find complex signals of the form m_{AB}. Formally, these are new atomic signals. However, they bear a similarity to basic atomic signals by the distance, $s = d(s_{AB}, s_A) = d(s_{AB}, s_B)$. On Franke's account, "spill-over" affects the reinforcement of non-actualised state/message pairs proportional to their similarity to the actualised state/message pair (similarly for the receiver and message/act pairs). At the same time, "lateral inhibition" lowers the accumulated rewards for non-actualised pairs when the actualised pair was successful. The "creative" portion of *creative compositionality* has to do with the fact that a new (complex) signal is more likely to be used when a new (complex) state arises—that is, the sender chooses a compositional signal with some likelihood, though she has never seen the complex state before.

Functional Negation The "negation game", introduced by Steinert-Threlkeld (2016, 2017), models a type of functional compositionality. The game is like an $n \times n$ signalling game, except there are $2n$ possible states and acts; thus, the sender has n atomic signals, $m_1, \ldots, m_n$, but the sender can also send signals of the form $\boxminus m_i$ for $1 \leq i \leq n$, where $\boxminus$ functions as a sort of "minimal negation".[6] The model for minimal negation that Steinert-Threlkeld (2016, 2017) employs has much structure built-in. However, he is less concerned here with the question of *how* compositional signals might arise as he is with the question of *why* compositional signals might arise.

Combinatorial Signals Scott-Phillips and Blythe (2013) try to differentiate "combinatorial" or "composite" communication systems from "non-combinatorial" or "non-composite" (i.e., atomic) communication systems. A signalling system, on their account, is *composite* if it contains at least one pair of composite signals—where the combination (concatenation) of two signals, $m_k = (m_i \circ m_j)$, is produced in at least

[6] The mathematical notion here describes a *derangement* $f : [2n] \to [2n]$—namely, a bijective function with no fixed points—for $[n] = \{1, \ldots, n\}$. Further, f is applied to both the states and the acts. So, $f(s_i){:=}s_{f(i)}$ and $f(a_i){:=}a_{f(i)}$.



(A) Non-Composite Communication System (B) Composite Communication System

Fig. 5 Composite versus non-composite communication systems. m_1 and m_2 are atomic signals. In (A), m_3 is atomic, and unique from m_1 or m_2, and so is not a combination of these. In (B), m_3 is composite signal because it is a combination of m_1 and m_2 (e.g., concatenation), and it produces an action unique from either m_1 or m_2

one non-composite state, $s_k \neq (s_i \circ s_j)$; see Fig. 5. A combinatorial communication system, then, is a system that includes at least one pair of fully-composite signals.

The signalling system of putty-nose monkeys is composite in their sense. The presence of eagles elicits a “pyow” signal, which in turn elicits the action *climb down a tree*; the presence of leopards elicits a “hack” signal, which in turn elicits the action *climb up a tree*. However, the absence of food elicits the combinatorial “pyow-hack” signal, which in turn elicits the action *move to a new location* (Arnold & Zuberbühler, 2008).

This model captures a similar notion of *syntactic* combination as in the syntactic signalling game (Barrett, 2007). However, Scott-Phillips and Blythe (2013) stipulate that (atomic) signal-order does not matter in their model, so the meaning of $(m_1 \circ m_2)$ is equivalent to the meaning of $(m_2 \circ m_1)$. Thus, their model fails to capture sensitivity to syntactic structure which is apparent in complex signals in, e.g., bird song and whale song. Barrett (2007) is sensitive to signal order, but complex signals get *interpreted* atomically. Thus, the meaning of a fully composite signal pair need not have anything to do with the meaning of its parts when we consider lexical combination in isolation. In order to account for meanings, we require a separate notion of systematicity which is not accounted for by evolutionary models of “compositional” signalling.

This criticism is treated formally in Section 4. In the next section, I provide some formal background from information theory before describing how a mathematical notion of information has been used to elucidate the content of signals in the signalling game.

3 Information and Meaning

This section provides some formal machinery that will be useful for the main argument in Section 4. I begin by introducing and discussing some formal concepts from information theory (3.1). I then highlight how the latter formalism has been used to provide a concept of semantic meaning in the signalling game framework (3.2).

3.1 Shannon Entropy and Relative Entropy

Shannon entropy measures the degree of randomness in some data set. Higher entropy means a higher degree of randomness, and less entropy means *higher predictability*.

Suppose X is a discrete random variable (RV) with alphabet $\mathcal{X}$ and probability mass function $p(x) = p_X(x) = \Pr\{X = x\}, x \in \mathcal{X}$.[7] The definition for Shannon entropy is given in 5.

Definition 5 *Shannon Entropy*:
The *entropy*, $H(X)$, of a discrete random variable, X, is defined by

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log_b p(x). \tag{1}$$

$\square$

The base of the logarithm, b, determines the unit of measure. For $b = 2, e, 10$, the unit of information is given by *Bit*, *Nat*, or *Hart*, respectively. We assume that $0 \log 0 = 0$. The entropy of a discrete RV does not depend on the alphabet since it is a function of the *distribution* of X; therefore, it depends solely upon the probabilities underlying this distribution. The entropy of an RV, in general, is characterised as a measure of how much information is required, on average, to describe the RV fully. For example, if we consider the set of states in the atomic 2-game as a discrete RV, $S = \{s_0, s_1\}$ with $p(s_0) = p(s_1) = 1/2$, $H(S)$ tells us that we need, on average, 1 bit of information to describe S.

Relative entropy—also known as Kullback-Leibler (KL) Divergence—is understood as a measure of the similarity of two probability distributions, p and q.

Definition 6 *Relative Entropy (Kullback-Leibler Divergence)*:
The *relative entropy*, or the *Kullback-Leibler distance*, from the probability mass function $q(x)$ to $p(x)$ is defined as

$$\begin{aligned} D(p \parallel q) &= \sum_{x \in \mathcal{X}} p(x) \cdot (\log_b p(x) - \log_b q(x)) \\ &= \sum_{x \in \mathcal{X}} p(x) \log_b \frac{p(x)}{q(x)} \end{aligned} \tag{2}$$

$\square$

With these definitions in place, the next section describes how KL-divergence has been used to describe the *semantic information* of a signal in a signalling game.

3.2 Semantic Information and Signalling

Entropy (Definition 5) is not equivalent to, or a measure of, information in the colloquial sense—e.g., the *content* of a signal or message. Since entropy (H) is an average, every message in a repertoire 'has' the same entropy value. However, each message in the repertoire may be *about* different things—i.e., messages may have different

[7] In this case, $p(x)$ and $p(y)$ refer to two different RVs—indeed, two different probability mass functions, $p_X(x)$ and $p_Y(y)$. See discussion in Cover and Thomas (2006).

meanings or *contents*. Thus, the entropy of distinct signals may be identical though the 'information' those signals carry, in the colloquial sense, is different.

Entropy depends upon discrete RVs. However, we note that the elements of the signalling game, described in Definition 2, can be understood as a set of discrete RVs, $\{S, M, A\}$. S is a static RV with some probability distribution—uniform, in the atomic case. At a signalling system, the signals are entirely informative, and the receiver has complete information about the state. Therefore, she can act as though she had observed the state directly. The 'key quantity', described by Skyrms (2010), depends upon a comparison between the (conditional) probability that a particular state obtains *given that* a signal was sent, and the likelihood that we are in that state simpliciter:

$$\frac{p(s_i|m_j)}{p(s_i)}$$

We can define the *quantity* of information a signal, m_j, carries (i.e., about a particular state, s_i) as

$$Q(m_j, s_i) = \log_2 \frac{p(s_i|m_j)}{p(s_i)}.$$

When signals are random, they carry no information. At a signalling system in the atomic 2-game, each signal carries exactly 1 bit of information, corresponding to a reduction of uncertainty from two possible states to one, conditional on the signal.

Skyrms (2010) highlights that signals may carry information about different states. Taking a weighted sum of the probabilities of being in any particular state conditional upon the specific signal, we obtain the following measure of the quantity of information carried by a particular signal, m_j, about the states:

$$\underset{states}{I(m_j)} = \sum_{i=1}^{|S|} p(s_i|m_j) \cdot \log_2 \left(\frac{p(s_i|m_j)}{p(s_i)} \right) \tag{3}$$

This is just the KL-Divergence (Definition 6) of the two probability distributions $P = p(s|m)$, $Q = p(s)$. Signals can also carry information about the acts:

$$\underset{acts}{I(m_j)} = \sum_{i=1}^{|A|} p(a_i|m_j) \cdot \log_2 \left(\frac{p(a_i|m_j)}{p(a_i)} \right) \tag{4}$$

In this context, the relative entropy of a particular signal can be understood as a measure of *additional bits gained* by moving from a prior distribution, $p(s)$, to a posterior distribution, $p(s \mid m)$, in a Bayesian sense. Equations 3 and 4 give the *quantity* of information in a signal. On Skyrms' (2010) account, the quantity of information is a summary number—i.e., the bits carried by a signal in a state; in contrast, the signal's informational *content* is a vector that specifies the information that the signal gives about each state.[8]

This vector is given by

$$\begin{aligned} \underset{states}{I}(m_j) &= \left\langle \log_2\left(\frac{p(s_0|m_j)}{p(s_0)}\right), \log_2\left(\frac{p(s_1|m_j)}{p(s_1)}\right), \cdots, \log_2\left(\frac{p(s_n|m_j)}{p(s_n)}\right)\right\rangle \\ &= \left\langle Q(m_j, s_0),\ Q(m_j, s_1),\ \cdots,\ Q(m_j, s_n)\right\rangle \end{aligned} \tag{5}$$

for the content about the states of a particular signal, m_j.

Suppose there are four initially equiprobable states, and σ is a constant function. In this case, the informational content about the states of each signal is given by the following vectors.[9]

$$\begin{aligned} I(m_1) &= \langle 0, 0, 0, 0\rangle \\ I(m_2) &= \langle 0, 0, 0, 0\rangle \\ I(m_3) &= \langle 0, 0, 0, 0\rangle \\ I(m_4) &= \langle 0, 0, 0, 0\rangle \end{aligned} \tag{6}$$

None of the signals carries any information about the states, so their content is empty everywhere. If we further suppose that the sender and receiver evolve to a signalling system where signal i is sent only in state i (*mutatis mutandis* for when the receiver performs act i only when she receives signal i), then the informational content of each signal at that signalling system is given by the following vectors.

$$\begin{aligned} I(m_1) &= \langle 2, -\infty, -\infty, -\infty\rangle \\ I(m_2) &= \langle -\infty, 2, -\infty, -\infty\rangle \\ I(m_3) &= \langle -\infty, -\infty, 2, -\infty\rangle \\ I(m_4) &= \langle -\infty, -\infty, -\infty, 2\rangle \end{aligned} \tag{7}$$

Now, each signal carries precisely 2 bits of information about the state of nature. The $-\infty$ components tell us which signals end up with probability 0, conditional on the states. Skyrms (2010) suggests that the traditional account in the philosophy of language—where the (declarative) content of a signal is a proposition, and a proposition is a set of possible worlds—is *contained* in this richer information-theoretic account of the content of a signal.[10]

[8] A more robust account of the informational content of a signal, explicitly couched in the language of semantic theory (there called "s-vector semantics"), is given by Isaac (2019).

[9] The same considerations apply when we consider the informational content about the acts for each signal.

[10] Note that some authors have criticised and extended this account. For example, the above characterisation of informational content depends upon how probabilities are *moved* (Skyrms, 2010). Godfrey-Smith (2011) suggests that the content of the signal should say something about *the world* rather than how much the probability of a particular state was moved by the signal's being sent. Birch (2014) highlights that Skyrms' account of informational content falls prey to the *problem of error* (in the same way as the information-

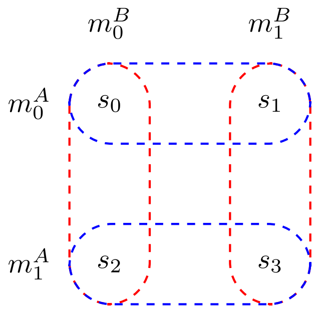


Fig. 6 Fully partitioning states via set intersection

With the formal machinery of Sections 2.1, 3.1, and 3.2 in place, we are now able to understand why syntactic signalling cannot be compositional.

4 Measuring Compositionality

Given the formal definition of semantic information discussed in Section 3.2, we can make exact the argument that syntax alone does not give rise to compositionality. This captures the complaints of Franke (2016); Steinert-Threlkeld (2016), that composite signals are interpreted atomically and so cannot be compositional in the sense that they do not capture intuitions about generalisability conditions for compositional signalling—i.e., systematicity.

Suppose we have a 4×4 syntactic signalling game, with two senders, σ_A and σ_B, and one receiver, ρ. Each sender can send one of two messages, and the receiver is sensitive to which sender sent which message. Suppose further that the senders and receiver have evolved a signalling system so that each sender's signal partitions nature into two sets—$\{s_0, s_1\}$ and $\{s_2, s_3\}$ for σ_A, and $\{s_0, s_2\}$ and $\{s_1, s_3\}$ for σ_B. The signal combinations determine the state via the intersection of these sets. See Fig. 6.

We know the maximal entropy of the system from Definition 5, given by

$$\begin{aligned} H(S) &= -\sum_{s \in \mathcal{S}} p(s) \log_2 p(s) \\ &= -\log_2 \left(\frac{1}{4}\right) \\ &= 2 \text{ bits.} \end{aligned}$$

Thus, an entirely informative length-two signal carries 2 bits of information because it reduces the possible states from 4 to 1.

We defined the *informational content* of a particular signal with respect to the states as a vector. Therefore, we can give the entire informational content of all of the

theoretic approach to content in Dretske (1981)), when we consider what it means for a signal to have *false* propositional content; see also Fodor (1984); Godfrey-Smith (1989); Crane (2003). Although these insights are theoretically valuable, we will ignore them for now. However, see further discussion in Skyrms and Barrett (2019); Shea et al. (2018); LaCroix (2020b).

Table 1 Complete informational content about the states at a signalling system in a 4×4 syntactic signalling game

		States			
		s_0	s_1	s_2	s_3
	m_0^A	1	1	$-\infty$	$-\infty$
Informational	m_1^A	$-\infty$	$-\infty$	1	1
Content	m_0^B	1	$-\infty$	1	$-\infty$
	m_1^B	$-\infty$	1	$-\infty$	1

Table 2 Complete informational content in simple signals about the acts at a signalling system in a 4×4 syntactic signalling game

		States			
		s_0	s_1	s_2	s_3
	$m_0^A \frown m_0^B$	2	$-\infty$	$-\infty$	$-\infty$
Informational	$m_0^A \frown m_1^B$	$-\infty$	2	$-\infty$	$-\infty$
Content	$m_1^A \frown m_0^B$	$-\infty$	$-\infty$	2	$-\infty$
	$m_1^A \frown m_1^B$	$-\infty$	$-\infty$	$-\infty$	2

signals explicitly as a matrix. Each row is the informational *content*, as described in Section 3.2, of a particular message; see Table 1.

Further, we can see that a particular state is *wholly determined* by all and only the messages that carry information about that state. Therefore, s_0 is entirely determined by the combination of m_0^A and m_0^B, rather than, e.g., the combination of m_0^A and m_1^B, because the latter carries no information about state 0 when in combination with m_0^A. The syntactic combination of a syntactic length-two signal carries complete information about a particular state; see Table 2.

Now, consider the same 4×4 syntactic signalling game at a signalling system, as described above, but suppose that σ_B spontaneously changes her signal m_0^B to a new signal, $m_?^B$. We might imagine two distinct explanations for this change:

(1) σ_B simply uses a novel signal in lieu of m_0^B.
(2) B forgets the meaning of signal m_0^B.

These two situations might be modelled in various ways. For example, 'forgetting' the meaning of signal m_0^B, as in case (2), can be modelled by 'emptying' all of the balls from the s_0 and s_2 urns for σ_B and adding a novel ball labelled $m_?^B$ to those urns. But, the descriptions of case (1) and (2) will be functionally equivalent under the assumption that the meaning of the *other* signal does not change for σ_B, as we shall see.

In case (1), the arbitrary signal, $m_?^B$, has simply replaced m_0^B; they mean the same thing. Under the urn-learning metaphor described in 2.1, this can be modelled by taking every ball labelled m_0^B in each of the state urns for σ_B and re-labelling them $m_?^B$. This re-labelling does not change that the senders already convened upon a signalling system that perfectly partitions the states of nature; however, the receiver must now learn the meaning of $m_?^B$. This situation is a cue-reading game (Barrett & Skyrms, 2017).

Table 3 Complete informational content about the states at a signalling system in a 4×4 syntactic signalling game with a novel signal identical to the old signal

		States				
		s_0	s_1	s_2	s_3	
	m_0^A	1	1	$-\infty$	$-\infty$	
Informational	m_1^A	$-\infty$	$-\infty$	1	1	
Content	$m_?^B$	1	$-\infty$	1	$-\infty$	$\leftarrow$ Novel Signal
	m_1^B	$-\infty$	1	$-\infty$	1	

In case (2), σ_B needs to *re*-coordinate so the new signal successfully partitions nature when combined with σ_A's signal. This is similar to a normal signalling context since the σ_B must re-learn when to send this novel signal (given the meanings of all the other signals are fixed, the correct strategy is to send the new signal in the same context as that in which the prior signal was used), and ρ must additionally learn the meaning of the novel signal. However, since the meaning of m_1^B is fixed, it follows that even if we 'reset' the urns for s_0 and s_2 with one each of $m_?^B$ and m_1^B, the conditional probability that s_0 obtains given that m_1^B is sent is effectively 0. Therefore, the informational content vectors about the states remain unchanged under either interpretation. This is shown in Table 3.

However, the signals also carry information about the acts. Assuming that m_0^B is replaced with $m_?^B$, this can be modelled by effectively throwing out the receiver urns that have a token of m_0^B and *creating* new urns that are labelled identically to the old urns, except with each token of m_0^B replaced with $m_?^B$. That is, the receiver urns labelled $m_i^A \frown m_0^B$ are replaced with new urns, labelled $m_i^A \frown m_?^B$. Each new urn contains one ball for each possible act—i.e., the receiver needs to re-coordinate with the senders on which action she should take when receiving a syntactic signal containing a token of the $m_?^B$ signal. Again, we can re-calculate the information that each of the concatenated signals contains about the acts, as in Table 4.[11] That is to say, any composite signal containing a token of the novel signal *carries no information* about the acts.

If the concatenated signals were compositional, this should not happen. Consider that, regardless of the new signal's meaning, m_0^A is only sent for a_0 or a_1. Therefore, the conditional probability that a_2 or a_3 should obtain, given that the receiver has received a length-two string starting with m_0^A, is 0. The probability of a particular act being appropriate simpliciter is still the chance probability, 0.25. What does this mean

[11] Recall from Section 3.2 that signals carry information about both the states (calculated in Table 3) and acts 4. Each entry in these tables is just an application of Equations 3 and 4, respectively.

$$\begin{aligned} \underset{acts}{I}(m_0^A \frown m_?^B) &= \left\langle \log_2\left(\frac{p(a_0|m_0^A \frown m_?^B)}{p(a_0)}\right), \ldots, \log_2\left(\frac{p(a_3|m_0^A \frown m_?^B)}{p(a_3)}\right)\right\rangle \\ &= \left\langle 0,\ 0,\ 0,\ 0\right\rangle \end{aligned} \tag{8}$$

Table 4 Complete informational content in simple signals about the acts at a signalling system in a 4×4 syntactic signalling game

		Acts a_0	a_1	a_2	a_3
	$m_0^A \frown m_?^B$	0	0	0	0
Informational	$m_0^A \frown m_1^B$	$-\infty$	2	$-\infty$	$-\infty$
Content	$m_1^A \frown m_?^B$	0	0	0	0
	$m_1^A \frown m_1^B$	$-\infty$	$-\infty$	$-\infty$	2

for the informational content of the concatenated signal? It is given by

$$\left\langle \log_2 \left(\frac{p(a_0 | m_0^A \frown m_?^B)}{p(a_i)} \right) \right\rangle, \qquad i \in \{1, 2, 3, 4\}.$$

Substituting the values for the conditional and unconditional probabilities, we have

$$\left\langle \log_2 \left(\frac{1/2}{1/4} \right), \log_2 \left(\frac{1/2}{1/4} \right), \log_2 \left(\frac{0}{1/4} \right), \log_2 \left(\frac{0}{1/4} \right) \right\rangle,$$

which resolves to the informational content vector

$$\langle 1, 1, -\infty, -\infty \rangle .$$

However, this makes no sense: m_0^A alone gives us 1 bit of disjunctive information—namely, about $a_0 \vee a_1$. If the ρ interprets the concatenation of m_0^A and $m_?^B$ compositionally, indeed, the second part of the length-two signal would not give her any *new* information regarding the disjunction $a_0 \vee a_1$—namely, unlike before, where the novel signal provides disjunctive information so the union of the two signals uniquely determines a single state. There is no reason why changing the second signal should take information away from the entire composite signal. The receiver interprets the signal as an atomic whole, which provides no information about the act.

This shows that the signals are not *interpreted* compositionally. However, it also highlights that they are compositional for the senders (or, for the states, if you prefer). This is because there is a notion of *independence*—concerning the information that the signal carries about the states—that does not hold for the information that the signal carries about the acts.

We assumed that the states were fixed in the previous example, and only one of the signals changed its meaning. We saw that this has no effect on the informational content of the signal concerning the states, but the receiver counter-intuitively loses the information that should have been contained in the unchanged signal. The same argument holds if, instead of supposing the lexicon is altered, it is merely extended—i.e., if a novel state, a novel signal to represent that state, and a novel action to perform in that state are introduced into the signalling game.

To tell an intuitive story, we might suppose that σ_A sends a verb, and sender B sends a *noun*. Suppose there are two distinct action contexts and two distinct object contexts.

Thus, we have the 4×4 syntactic signalling game, as before. Suppose now that a *novel object* context is added to the game. The noun-sender accommodates this by adding a new signal to her lexicon and sending that in the novel context. The receiver must learn what is appropriate given this new signal; however, given that the verb context has not changed, she should gain *some* information. This argument captures precisely the *systematicity* feature of compositional communication: if the receiver knows the meaning of 'pick up x' and the meaning of 'the book', but not the meaning of 'put down x', then she might understand the command 'pick up the book', though she does not understand the meaning of 'put down the book'. Even so, she may still understand that the latter expression has *something* to do with the book.

The preceding argument captures the criticism of Franke (2014) that the syntactic signalling model "misses a key feature of compositionality, namely that it is a flexible and potentially creative ability to associate novel expressions with novel meanings" and that the " meaningfulness that arises in Lewis-style signaling models is holophrastic. By highlighting that we can decouple the informational content for the sender and the receiver, this argument clarifies the informal suggestion that even though "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" Franke (2014).[12] The analysis offered demonstrates that, on the syntactic signalling model, the informational content of a signal is different for the sender and the receiver. Hence, Skyrms' (2010) claim that "[t]he information in a complex signal is a function of the information in its parts" is only true for the sender.

However, it is worth noting that the main argument offered in Section 4 is neutral with respect to the dynamic in question; instead, the key insight about the informational content of syntactic signals follows from an analysis of a given signalling behaviour *at a point in time*—e.g., after the sender and receiver have coordinated on a signalling system. As a result, attempts to explain compositionality by altering the *dynamic*, but not the structure of the game, will fail for the same reason.

Hence, although this argument specifically concerned the syntactic signalling model given in Barrett (2006, 2007, 2009), the same considerations apply to the model for combinatorial systems of communication proposed by Scott-Phillips and Blythe (2013). Since they explicitly focus on composition *qua* syntactic structure, this system cannot give rise to a genuine notion of compositional signalling. The same is true for spill-over reinforcement (Franke, 2016). If we add a novel signal to a pre-established signalling system that has evolved via spill-over RL, the receiver loses information in any string containing the novel signal; therefore, Brochhagen (2015) is correct in pointing out that the agents are not sensitive to a generalisation condition for compositionality—namely, the relations between constituent parts are not generalisable.[13]

In particular, for a complex signal to be compositional, there needs to be a systematic association between simplex elements and the complex elements of which they are constituents. To account for productivity, structural properties that are common

[12] See also discussion in Franke (2016).

[13] Recall that the models extensions proposed by Nowak and Krakauer (1999); Steinert-Threlkeld (2016) do not purport to demonstrate *how* compositional signals evolve.

between components of complex signals must be recognisable (and indeed recognised) for it to be possible to learn how to (de)compose two such elements in such a way that this can be generalised over their classes. In particular, as we have seen, if each combination of parts needs to be learned case by case and mentally stored in a lexicon for interpretation, then this will not provide any advantage *to the receiver*.

This is enough for *syntactic* compositionality—composition on the part of the sender. However, this leads to what Steinert-Threlkeld (2020) calls *trivial compositionality*:

> *Trivial Compositionality*
> A communication system is trivially compositional just in case complex expressions are always interpreted by the intersection (generalised conjunction) of the meanings of the parts of the expression. (Steinert-Threlkeld, 2020, 3-4)

He highlights that the models that discussed by Nowak and Krakauer (1999); Barrett (2007, 2009); Mordatch and Abbeel (2018) all share the following underlying assumptions:

(A1) Agents communicate about a fixed set of states.
(A2) Optimal communication consists in correctly identifying the true member of the state space.
(A3) Messages are fixed-length sequences of signals from fixed sets.

He then *proves* that if a model carries all three of these assumptions, the composition that emerges from the model will necessarily be trivial.[14] Namely, these three conditions are jointly sufficient for trivial compositionality.

By considering the fact that informational content about the states and the acts can be decoupled, the preceding argument adds some nuance to this distinctions between trivial and non-trivial compositionality. In particular, any adequate account of compositionality has to explain *both* syntax and semantics. Since prior proposals have focused on the production of syntactically complex signals (i.e., the sender's behaviour), they only account for the former. However, one useful feature of the signalling game's formal structure is that we can decouple the sender and reciever strategies. Hence, genuine compositionality requires that the sender composes signals syntactically, but also that the receiver *decomposes* them on interpretation. Because the argument is based on the sender (reciever) behaviour at a point in time, this implies that changes to the dynamic alone will not resolve the issue. Instead, an explanation of how compositionality evolves requires a different game architecture altogether. Some examples are given in 5.

[14] Note that Steinert-Threlkeld (2016) and Barrett et al. (2020) drop assumption (A3); Steinert-Threlkeld (2020) drops (A1).

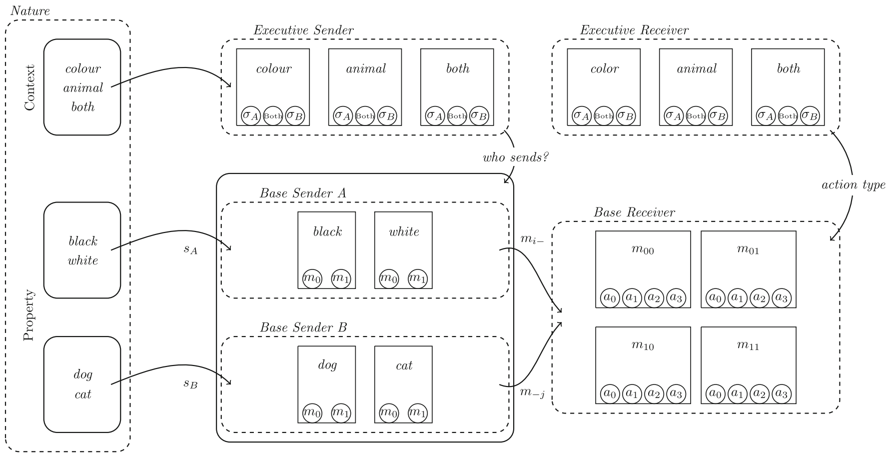


Fig. 7 Simple urn learning model for a hierarchical, co-evolutionary, syntactic signalling game

5 Moving Forward

Explanations of the evolution of compositionality have fixed upon the evolution of compositional *syntactic* structure alone while failing to attend to the importance of semantics and systematicity that is required for an adequate evolutionary explanation of the origins of *language*. The insights of Section 4 suggest that focusing exclusively on syntax in discussing the evolution of compositionality under the signalling-game framework is misguided. Part of the purpose of this paper was to demonstrate that extensions of the simple signalling model that seek to explain compositionality often fail to do so. Implicit in this argument is the suggestion that the emphasis on compositionality, more generally, in the language origins literature is a mistake, even if compositionality is a distinguishing feature of natural language. Hence, an alternative explanatory target is required.

Barrett et al. (2020); LaCroix (2022) show how compositional signalling might evolve in a hierarchical signalling game with two basic senders, one executive sender, one basic receiver, and one executive receiver. See Fig. 7. On this hierarchical model, the basic senders and receiver play a standard syntactic signalling game (the base game), where each sender sends a signal for one of two properties—*colour* and *animal*. Moreover, there is a *context* that determines which signal is relevant for the actions. Hence, the executive sender and receiver—called *hierarchical agents*—can learn to influence the behaviour of the basic senders and receiver—called *basic agents*.

However, it is not the compositionality of the signals that drives compositionality in this signalling system. This can be seen by the fact that the base game (constituted by the base senders and base receiver) is functionally equivalent to the 4×4 syntactic signalling game, which does *not* evolve compositional signalling, as we have seen. Instead, it is the *reflexivity* and *modularity* of the executive sender and receiver that drives compositionality in this context, insofar as the ball that the executive sender chooses *refers* to a component of the base game (LaCroix, 2020b); see Fig. 8. The

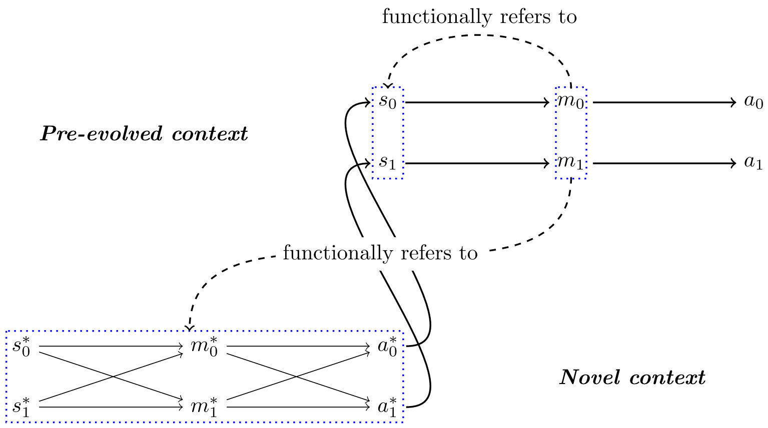


Fig. 8 Simple model of proto-reflexivity in a signalling game context

hierarchical game structure, understood as an evolutionary context in which a complex signalling game arises, is itself composed of simpler "modules"—i.e., the base game and the hierarchical game. Barrett and Skyrms (2017); Barrett (2023) and others have begun to explore how such modules or hierarchical structures can *self-assemble* to create more complex game structures.

Moreover, reflexivity, itself, gives rise to functional composition as a by-product of these processes. On this account, signals "may become functionally referential, referring to concrete objects in the world. Once individuals are able to make use of protoconcepts, they can refer to abstracta. Therefore, they can refer to communicative contexts, giving rise to protoreflexivity" (LaCroix, 2021). In addition, such proto-reflexive abilities mean that the players can influence future communicative behaviour *via* communication. Hence, simple communicative capacities evolve alongside cognitive capacities.

References

Arnold, Kate, & Zuberbühler, Klaus. (2008). Meaningful Call Combinations in a Non-Human Primate. *Current Biology, 18*, R202–R203.

Barrett, Jeffrey. (2006). *Numerical Simulations of the Lewis Signaling Game: Learning Strategies, Pooling Equilibria, and Evolution of Grammar*. Institute for Mathematical Behavioral Science: Technical Report.

Barrett, Jeffrey. (2007). Dynamic Partitioning and the Conventionality of Kinds. *Philosophy of Science, 74*, 527–546.

Barrett, Jeffrey. (2009). The Evolution of Coding in Signaling Games. *Theory and Decision, 67*, 223–237.

Barrett, Jeffrey A. (2023). Self-assembling games and the evolution of salience. *British Journal for the Philosophy of Science*, 74(1).

Barrett, Jeffrey A., Cochran, Calvin, & Skyrms, Brian. (2020). On the Evolution of Compositional Language. *Philosophy of Science, 87*(5), 910–920.

Barrett, Jeffrey A., & Skyrms, Brian. (2017). Self-Assembling Games. *The British Journal for the Philosophy of Science, 68*(2), 329–353.

Beggs, A. W. (2005). On the Convergence of Reinforcement Learning. *Journal of Economic Theory, 122*, 1–36.

Birch, Jonathan. (2014). Propositional Content in Signalling Systems. *Philosophical Studies, 171*(3), 493–512.

Brochhagen, Thomas. (2015). Minimal Requirements for Productive Compositional Signaling. In *CogSci*, pages 285–290.

Cover, Thomas M., & Thomas, Joy A. (2006). *Elements of Information Theory*. John Wiley & Sons, Hoboken, 2 edition.

Crane, Tim. (2003). *The Mechanical Mind: A Philosophical Introduction to Minds, Machines and Mental Representation*. Routledge, London, 2 edition.

Dretske, Fred. (1981). *Knowledge and the Flow of Information*. The MIT Press.

Fodor, Jerry. (1984). Semantics, Wisconsin Style. *Synthese, 59*, 231–250.

Fodor, Jerry A., & Pylyshyn, Zenon W. (1988). Connectionism and Cognitive Architecture: A Critical Analysis. *Cognition, 28*(1), 3–71.

Franke, Michael. (2014). Creative Compositionality from Reinforcement Learning in Signaling Games. In Cartmill, Erica A., Seén Roberts, Heidi Lyn, and Hannah Cornish, editors, *The Evolution of Language: Proceedings of the 10th International Conference*, volume 10 of *Evolang*, pages 82–89. World Scientific, Singapore.

Franke, Michael. (2016). The Evolution of Compositionality in Signaling Games. *Journal of Logic, Language and Information, 25*(3), 355–377.

Godfrey-Smith, Peter. (1989). Misinformation. *Canadian Journal of Philosophy, 19*, 533–550.

Godfrey-Smith, Peter. (2011). Signals: Evolution, Learning, and Information by Brian Skyrms (Review). *Mind, 120*(480), 1288–1297.

Hodges, Wilfrid. (2012). Formalizing the Relationship Between Meaning and Syntax. In Wolfram Hinzen, Edouard Machery, & Markus Werning (Eds.), *The Oxford Handbook of Compositionality* (pp. 245–261). Oxford: Oxford University Press.

Hopkins, Ed. & Posch, Martin. (2005). Attainability of Boundary Points Under Reinforcement Learning. *Games and Economic Behavior*, 53: 110–125.

Huttegger, Simon M. (2007). Evolution and the Explanation of Meaning. *Philosophy of Science, 74*, 1–27.

Isaac, Alistair M. C. (2019). The semantics latent in shannon information. *The British Journal for the Philosophy of Science, 70*(1), 103–125.

Jackendoff, Ray. (2010). Your theory of language evolution depends on your theory of language. In Richard K. Larson, Viviane Déprez, & Hiroko Yamakido (Eds.), *The Evolution of Human Language: Biolinguistic Perspectives* (pp. 63–72). Cambridge University Press.

Janssen, Theo M. V. (2012). Compositionality: Its Historic Context. In Wolfram Hinzen, Edouard Machery, & Markus Werning (Eds.), *The Oxford Handbook of Compositionality* (pp. 19–46). Oxford: Oxford University Press.

Kamp, Hans, & Partee, Barbara. (1995). Prototype Theory and Compositionality. *Cognition, 57*, 129–191.

LaCroix, Travis. (2020a). Communicative bottlenecks lead to maximal information transfer. *Journal of Experimental & Theoretical Artificial Intelligence, 32*(6), 997–1014.

LaCroix, Travis. (2020b). *Complex Signals: Reflexivity, Hierarchical Structure, and Modular Composition*. PhD thesis, University of California, Irvine.

LaCroix, Travis. (2020c). Evolutionary Explanations of Simple Communication: Signalling Games and Their Models. *Journal for General Philosophy of Science, 51*, 19–43.

LaCroix, Travis. (2021). Reflexivity, functional reference, and modularity: Alternative targets for language origins. *Philosophy of Science, 88*(5), 1234–1245.

LaCroix, Travis. (2022). Using logic to evolve more logic: Composing logical operators via self-assembly. *British Journal for the Philosophy of Science, 73*(2), 407–437.

Lewis, David (2002/1969). *Convention: A Philosophical Study*. Blackwell, Oxford.
Mordatch, I., & Abbeel, P. (2018). Emergence of grounded compositional language in multi-agent populations. In *The Thirty-Second AAAI Conference on Artificial Intelligence*. AAAI.
Nowak, Martin A., & Krakauer, David C. (1999). The Evolution of Language. *Proceedings of the National Academy of Sciences, 96*, 8028–8033.
Nowak, Martin A., & Plotkin, Jansen. (2000). The Evolution of Syntactic Communication. *Nature, 404*, 495–498.
O'Connor, Cailin. (2014). The Evolution of Vagueness. *Erkenntnis, 79*(4), 707–727.
Partee, Barbara Hall. (1984). Compositionality. In F. Landman & F. Veltman (Eds.), *Varieties of Formal Semantics, page 281–311*. Dordrecht: Foris.
Scott-Phillips, Thomas C., & Blythe, Richard A. (2013). Why is Combinatorial Communication Rare in the Natural World, and Why is Language an Exception to this Trend? *Journal of the Royal Society Interface, 10*(88), 1–7.
Shea, Nicholas, Godfrey-Smith, Peter, & Cao, Rosa. (2018). Content in Simple Signalling Systems. *The British Journal for the Philosophy of Science, 69*(4), 1009–1035.
Skyrms, Brian. (2010). *Signals: Evolution, Learning, & Information*. Oxford: Oxford University Press.
Skyrms, Brian, & Barrett, Jeffrey A. (2019). Propositional Content in Signals. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 74*, 34–39.
Steels, Luc. (1995). *A Self-Organizing Spatial Vocabulary. Artificial Life, 2*(3), 319–332.
Steinert-Threlkeld, Shane. (2016). Compositional Signaling in a Complex World. *Journal of Logic, Language, and Information, 25*(3–4), 379–397.
Steinert-Threlkeld, Shane. (2017). *Communication and Computation: New Questions About Compositionality*. PhD thesis, Stanford University.
Steinert-Threlkeld, Shane. (2020). Towards the Emergence of Non-trivial Compositionality. *Philosophy of Science, 87*(5), 897–909.
Szabó, Zoltán Gendler. (2012). The Case for Compositionality. In Wolfram Hinzen, Edouard Machery, & Markus Werning (Eds.), *The Oxford Handbook of Compositionality* (pp. 64–80). Oxford: Oxford University Press.