Emergence of Public Meaning from a Teleosemantic and Game Theoretical Perspective

KARIM BARAGHITH

Abstract

The generalized theory of evolution suggests that evolutionary algorithms apply to biological and cultural processes like language alike. Variation, selection and reproduction constitute abstract and formal traits of complex, open and often self-regulating systems. Accepting this basic assumption provides us with a powerful background methodology for this investigation: explaining the emergence and proliferation of semantic patterns, that become conventional. A teleosemantic theory of public (conventional) meaning (Millikan 1984; 2005) grounded in a generalized theory of evolution explains the proliferation of public language forms in terms of their adaptive proper function. It has also been suggested, that the emergence of meaning, can be formalized with game-theoretical tools (Skyrms 2010) within signaling systems of coordination. I want to show how closely related these approaches are, both in terms of explanandum and of outcomes. To put it in a nutshell: If the emergence of public meaning can be satisfyingly explained in terms of signaling games, then the cultural evolutionary dynamics will serve as an adequate model to describe their proliferation. Public or conventional meaning (in contrast to personal meaning) can be fully understood in terms of its evolutionary function in a population of communicators. Furthermore, I want to argue how this understanding of conventional meaning could lead us to a strong semantic holism.

Keywords: public meaning, teleosemantics, signaling games, convention, semantic holism
Introduction

“The information-theoretic approach [Skyrms] links content to how a sign manages to inform a receiver; the teleofunctional view [Millikan] links it to what the sign is supposed to do, and how the world must be if the sign is to be able to do its job.”

- P. Godfrey-Smith [13, p.52]
In this paper, I want to discuss and compare two explanatory approaches on the evolution of conventional meaning: Teleosemantics (TS) and Signaling Games (SG). It has been argued, that the solution to the “mystery” of determining meaning will be found in a functional theory like Ruth Millikan’s TS [27], which I will examine in section 1, where the distinction between personal and public meaning (PM) is introduced, as well. An appropriate theory of PM should allow us to understand more about the proliferation of public linguistic forms and show how signals fulfill the evolutionary function that they were selected for. The signaling game (SG), a game theoretical model originally dating back to [23], has been used by recent philosophers [38] [15] [19] for that purpose. I will get to that in section 2. The meaning of public representations becomes conventional, where the behavioral regularities of the involved agents emerge from the process of stabilization of a game-theoretical strategy. There are as many kinds of meaning as there are signaling systems, which emerge and proliferate in different contexts or environmental niches. It will be shown how this statement fits in the framework of the teleosemantic approach. I intend to show that the explanatory systems of Millikan’s teleosemantics with producers and consumers and Skyrms’ SG involving senders and receivers are to a large extend functionally isomorph, because they share a common evolutionary background paradigm. Although both – Skyrms and Millikan – attempt selection analyses of meaning, their starting points differ considerably, as do their methods. Nevertheless, juxtaposing them can provide valuable insights. While Millikan’s approach stands more in the tradition of a very carefully worked out and elaborate conceptual analysis, Skyrms’ explanations of meaning is a formal investigation using game theory, computer simulations and even provides the opportunity of empirical research. But as I want to highlight in this paper, their theories are more than closely related – especially the case of PM. I also want to examine some applications. One possible bridge between the two theories can be found in the implementation of intension and extension to signaling systems. Another one is an evolutionary story of the emergence of the distinction between normative and descriptive statements of language. Both examples will be dealt with in section 3. In section 4, I will briefly discuss the question if adopting both, SG and PM potentially lead us to adopt some kind of semantic holism. The last section concludes.
1 A Teleosemantic Theory of Public Meaning

There are many different approaches to explain content or meaning. Ruth Millikan [27] and [29], but also Fred Dretske [8], David Papineau [32] or Karen Neander [31] developed naturalistic accounts of intentionality that made use of Darwinian evolution.

In the case of language, the latter has been instantiated in the form of cultural evolution [26] [39], understood in this paper as one of several instantiations of generalized Darwinism or a generalized theory of evolution [1], [37]. According to this framework, which serves as theoretical background for this investigation, evolutionary algorithms apply to biological and cultural processes (like language) alike, where variation, selection and reproduction can be seen as abstract and formal features of complex, open and often self-regulating systems. Ruth Millikan’s TS can be seen as the most carefully worked-out approach of semantic phenomena within naturalistic philosophy. According to Peter Godfrey-Smith perhaps even within any kind of philosophy [13, p.41]. Important for our purpose is the conceptual distinction between two kinds of meaning (section 1.1) and the “teleological element”, the use of a biological notion of function (section 1.2), grounded in history.

Millikan’s work is notable for its scope and complexity, which means that only a very small segment can be presented here, namely the one which is important for our investigation. [36] provide a lot more and deeper insights into this body of research. At the very beginning, an important conceptual distinction has to be made between two different kinds of meaning.

1.1 Personal and Public Meaning

According to Ruth Millikan [29, p.54], there are at least two relevant kinds of meaning: the speakers meaning on the one and the conventional, linguistic or public meaning (which again divides in several sub-kinds) on the other hand. It is the distinction between the cooperative function (1) of a language and the function (2) that individual speakers may associate with a sign or meaning-bearing entity. According to Millikan, (1) is fully extensional and seems to be a “distributed structure”, which seemingly cannot be found in the “head” of one person alone. But if we do not want to be semantic Platonists, believing in something akin to a Popperian “world 3” [33, p. 153], this statement has to be explained very carefully (cf. section 4).

Millikan [29, p. 55] and [38, p. 7] explicitly rely on David Lewis
seminal work [23], when talking about linguistic conventions. This supports our claim, that both, Millikan and Skyrms share another peace of theoretical background. Conventions in general do solve coordination problems. What exactly is a convention? According to [29, p. 56]:

“A convention is merely a pattern of behavior that is (1) handed down from one person, pair, or group of persons to others – the pattern is reproduced – and (2) is such that, if the pattern has a function, then it is not the only pattern that might have served that function about as well. Thus if a different precedent had been set instead, a different pattern of behavior would probably have been handed down instead.”

Conventions are historical entities, which involve some amount of arbitrariness. This definition matches almost perfectly into a framework of evolutionary SG (section 2).

For now, it is important to note, that there is a crucial distinction between speakers meaning and PM, or (as you might also put it conceptually) between a mental and a behavioral part of a representational system. The speakers meaning is something internal, PM something external, which – in a more sociological context – has also been called “public representation” [39, p. 61]. For the rest of the paper, when I use the term “meaning”, it will always refer to this public or conventional meaning, unless indicated otherwise. Note that most criticism of TS are related to arguments like those of [6] or [9], who indicate lack of explanatory power of teleofunctional approaches regarding mental representations and content and therefore personal meaning. Content – they claim – cannot be fully understood by only looking at its adaptive function or its history. Without arguing against these objections and “swamp men arguments” of any kind – I shall emphasize that the case may be quite different for public or conventional meaning. The reason for this difference lies in a different explanandum. If and how a TS can solve the problem of intentionality, mental content and personal meaning is a different question, that e.g. [31, p. 381] addresses, but will not be dealt with in this paper. Skyrms argues in a similar direction. As Millikan clearly emphasizes [29, p. 58], the functions of PM are not on the same level as either speaker purposes or hearer purposes. Conventional meaning is selected for being satisfactory at once to both partners in communication, it has to balance speaker with hearer interests. In the next section however, it will be shown that exactly this kind of PM can be seen as implementing a (cultural) evolutionary function within a generalized theory of evolution, cf. e.g. [37].
1.2 PM as Proper Function in Cultural Evolution

The conceptual debate about functions has a long philosophical tradition, cf. [2] and in many of its facets, it is deeply connected with evolutionary explanations and teleology (like in teleosemantics). An evolutionary analysis of the concept of function is able to explain functions as naturalistic processes [27, p. 28]. In an evolutionary context, something has a function if and only if it is an adaptation for a certain environment. As Schurz [37, p. 157] puts it, any causal effect E of a subsystem of an (biological or cultural) evolutionary system S constitutes an evolutionary function F, iff

1. E is a reproduced, hereditary trait that has been positively selected,

2. namely because E contributed to the fitness of S in its phylogenetic history.

A frequently used example is the function F of the heart S, to pump blood E. Our evolutionary ancestors evolved hearts or heart-like organs because that increased their ability to absorb oxygen from the environment (primitive organisms like cnidarian, e.g. jellyfish or sea stars do not have hearts), which helped them to increase their fitness (in terms of producing more offspring). Note that not all features of an organism are necessarily functional. The blood pumping heart also beats, but that sound has hardly a function, it is a mere side effect. This important distinction prevents us from premature adaptionism, according to which all (typical) traits have to be considered as adaptations.

The kind of function that is described here is etiological and has been named “proper function” by Millikan [27, p. 17]. This core concept of Millikan’s philosophy describes what things do when they work “properly”, namely when they do what they were selected for. In the case of malfunctions (if the heart does not pump blood properly anymore), the subsystem still keeps its evolutionary function, but is not able to perform it anymore.

In which way is all of that important for language and meaning? Central to Millikan’s philosophy of mind and language is the idea that semantic representations and the mechanisms that produce them have this etiological function as well. We saw that functions are adaptations. The PM of a representation constitutes a kind of adaptation or co-adaptation between speakers and hearers. As Millikan puts it [29, p. 57]:
“Speakers within a language community are, simply, adapted to an environment in which hearers are responding, sufficiently often, to the forms speakers produce in ways that reinforce these speaker productions. Correlatively, hearers in the community are, simply, adapted to conditions under which speakers, sufficiently often, produce these language forms in circumstances such that making conventional responses to them aids those hearers.”

The evolutionary function of PM lies in the mutual co-adaptation of speakers and hearers of a language community in a specific cultural context. What selection directly acts on is the way that producers create icons and the way consumers interpret them. In the next section, a formal model is presented, which is able to capture this notion of co-adaptation and the emergence of PM in a more formal and empirically testable way.

2 Evolutionary Signaling Games of Coordination

In this section, I shall introduce a model from game theory, which is a commonly used formal method for the investigation of rational agent based behavior in psychology, sociology or economics and was extended to the domain of evolutionary systems, as well.4

2.1 Lewis’ Signaling Games in an Evolutionary Framework

Signaling games (SG) of coordination model semantic interactions between a sender (S) and a receiver (R), (whereas Millikan speaks of producer and consumer) emphasizing the social aspect of language and addressing the question of how meaningful communication and conventional meaning can emerge. This view describes meaning as serving a certain function of language, namely to facilitate coordinated behavior. Meaning can therefore be seen as a consequence of pragmatic factors. The similarity to Millikan’s work is striking. As [13, p. 45] notes about Skyrms’ SG’s:

“In other models, the sender and receiver roles are separate and fixed, […]]. In the models sketched so far, “common interest” between sender and receiver is assumed. Similarly, Millikan’s discussions of “producer–consumer” systems in-
clude the requirement that these two devices “cooperate” with each other, as a consequence of evolutionary design.”

2.1.1 Philosophical Origins and Intentions

David Lewis’ seminal work [23] was the starting point for SG’s in philosophy. It challenges Quine’s counterarguments against the logical positivist’s claim that conventions of meaning form the basis of logical truth and logical inference, cf. [18, p. 2]. Quine doubted that the truth of a sentence is given by a factual component and a pragmatic/linguistic one, because analytic statements seem to lack any factual component. What remains is the linguistic component, simply governed by conventions of meaning (PM) but according to Quine, there can be no non-circular explanation of these conventions. Lewis argued against that by introducing SG, where the emergence of a convention would not require explicit agreement. They might be stable outcomes of repeated nonverbal interactions.

[38] [15] [16] [18] [19] [17] [30] [41] [43] among others analyzed SG in evolutionary terms. Bruner et. al. [4] created an interesting empirical application of the Lewisian framework of SG by testing participants, who faced a semantic coordination problem under specific conditions in the lab. They played iterated SG – and proved several predictions, which were based on computer simulations about the behavior of the trajectories in the population dynamics in biased and unbiased SG to be true. Signaling systems and thereby conventional meaning emerged between the participants of the experiment, at least in the simple Lewis’ standard framework.

Evolutionary explanations avoid difficulties that methodological explanations in terms of rational choice – like Lewis’ account – have to deal with at a very fundamental level. Although agents in evolutionary game theory are not as rational as in standard game theory, they still maximize their payoff, while the environment selects their strategies (i.e. their behavior). Individuals do not need to have control over their strategy (e.g. it can be an automatic or subconscious behavior) and need not necessarily to be aware of the game. It is just a question of how well the strategy proves successful in comparison to others, i.e. how often it gets reproduced.
2.1.2 Ingredients and Features of the Model

SG is based on a coordination problem between world states and acts and allows for a probabilistic association between world states, signals and acts. Although communication can be very complex, we will mostly concentrate on the most simple case here, which consists of a set of two world states $W = \{\sigma_1, \sigma_2\}$, two messages $M = \{m_1, m_2\}$ that S can communicate to R, and two acts $A = \{\alpha_1, \alpha_2\}$ by which R can respond to the received message. A pure S strategy is a function $s : W \rightarrow M$ from states to messages; a pure R strategy is a function $r : M \rightarrow A$ from messages to acts. There are 4 pure S and 4 pure R strategies and consequently 16 pure strategy profiles. Next to pure strategies, there are also mixed strategies.

Figure 1: A simple two state – message – act SG of coordination. Arrows represent the strict Nash equilibria, leading to the best payoff for both players.

Figure 1 [4] depicts the game in extensive form, which shows all possible strategies. We assume a prior probability of $P(\sigma_1) = 0.5$ which shows that both states occur with the same probability. Nature (or culture) flips a fair coin. It is important to note, that only S can observe $\sigma_i$. R has no information about it but only if R chooses the “right” corresponding $a_i$ communication is successful and both sustain a payoff of 1, denoted as $U = \binom{1}{1}$ in figure 1. They have to cooperate and coordinate their behavior, which means we have to look at the combination of their strategies.

[22, p. 321] gives a nice example of such a situation. Imagine two
persons (S and R) living on an island. To get food, they have to go fishing, but they only succeed, if they do this together (maybe the fishing net is too heavy for one person to carry). Only S can detect if and where fish can be found, and where they should go fishing at the sea because he has some kind of experience for that. R has an interest in knowing if \( \sigma_1 \) (fish) or \( \sigma_2 \) (no fish) is the case at the particular place in the particular time. So S has to communicate it to him by sending either \( m_1 \) (“There is fish.”) or \( m_2 \) (“There is no fish.”). R in turn has to react to this message in a compliant way, by choosing between \( a_1 \) (go fishing with S) or \( a_2 \) (do nothing). Only if this matching is correct, both will eat fish most certainly at the end of the day. Otherwise, it is just a matter of coincidence.

If, as we assume, each \( \sigma_i \) occurs with equal probability \( p = 0.5 \), then 16 combined and equally probable strategies are possible:

- \( s_1 : m_1 \) if \( \sigma_1, m_2 \) if \( \sigma_2 \) \( \rightarrow \) \( r_1 : a_1 \) if \( m_1, a_2 \) if \( m_2 \)
- \( s_2 : m_2 \) if \( \sigma_1, m_1 \) if \( \sigma_2 \) \( \rightarrow \) \( r_2 : a_2 \) if \( m_1, a_1 \) if \( m_2 \)
- \( s_3 : m_1 \) if \( \sigma_1, m_1 \) if \( \sigma_2 \) \( r_3 : a_1 \) if \( m_1, a_1 \) if \( m_2 \)
- \( s_4 : m_2 \) if \( \sigma_1, m_2 \) if \( \sigma_2 \) \( r_4 : a_2 \) if \( m_1, a_2 \) if \( m_2 \)

The four S strategies (which specify for each state what signal is to send) are listed on the left side, the four R strategies (which specify how act depending on the received message) on the right. The black arrows indicate the two possible combinations \( (s_1, r_1) \) of the strategies, which are beneficial for both players (especially under frequent iteration of the game) and therefore constitute the two *strict Nash equilibria*, because each player would do worse by deviating from his strategy, given that the other player does not. According to [23], strict Nash equilibria in SG are called *signaling systems*. The *meaning* of the messages that constitute signaling systems have a very high probability to become conventional in a given population of senders and receivers.

<table>
<thead>
<tr>
<th></th>
<th>( r_1 )</th>
<th>( r_2 )</th>
<th>( r_3 )</th>
<th>( r_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( s_1 )</td>
<td>1/1 (I)</td>
<td>0/0</td>
<td>1/2/1/2</td>
<td>1/2/1/2</td>
</tr>
<tr>
<td>( s_2 )</td>
<td>0/0</td>
<td>1/1 (II)</td>
<td>1/2/1/2</td>
<td>1/2/1/2</td>
</tr>
<tr>
<td>( s_3 )</td>
<td>1/2/1/2</td>
<td>1/2/1/2</td>
<td>1/2/1/2</td>
<td>1/2/1/2</td>
</tr>
<tr>
<td>( s_4 )</td>
<td>1/2/1/2</td>
<td>1/2/1/2</td>
<td>1/2/1/2</td>
<td>1/2/1/2</td>
</tr>
</tbody>
</table>

Table 1: The payoff matrix of the average payoff in the standard SG.
Table 1 shows the averaged payoffs for the combined strategies. $s_1, r_1$ (I) and $s_2, r_2$ (II) are the only two signaling systems (strict Nash equilibria); $(s_3, r_3), (s_4, r_3), (s_3, r_4)$ and $(s_4, r_4)$ form non-strict Nash equilibria, also referred to as pooling equilibria in the literature. They are states of partial communication, where S sends always $m_1$, regardless what $\sigma_i$ he observes or R undertakes $a_1$ regardless the given $m_i$. It is clear that under such circumstances, both will receive a payoff from time to time (on average: in half of the cases) but this happens randomly and will not prove stable under iteration. $(s_2, r_1)$ and $(s_1, r_2)$ are states of no communication, where no coordination takes place, at all.

2.1.3 Connection to Information Theory

If a signal $m_j$ is said to carry information, [38, p. 34], then the amount, i.e. the quantity of information is measured by the amount of change to the probabilities ($P$) of any element of $W = \{\sigma_1, \ldots, \sigma_i\}$. Let $P(\sigma_i)$ and $P(m_j|\sigma_1) > 0$ be the prior probabilities that $\sigma_i$ occurs and if so, $m_j$ is sent.

$$P(\sigma_i|m_j) = \frac{P(\sigma_i) \times P(m_j|\sigma_i)}{\sum_{\tau} P(\sigma) \times P(m_j|\sigma)}$$

shows the Bayesian likelihood of $\sigma_i$ given $m_j$ [11, p. 361]. It is very important to note, that only the amount, not the content of information is addressed here, since two different messages can move the initial probabilities by the same amount, but in different directions (e.g. increase or decrease it), as [38, p. 34] shows. He calculates the average information of states in a signal. Furthermore, he intends to show that this approach is quite sufficient for determining the content of a specific message (ibid.: p. 40), given some small mathematical extensions.

However, it has heavily been doubted that this measurement of information alone is sufficient for a real theory of meaning. [29, p. 13] herself argues in the opposite direction, when talking about the content of a representation, which – for her – cannot be measured by this classical information concept. This is a crucial difference between TS and SG.

I do not which to take a side here, but in section 4 of this paper, the notion of information by moving probabilities will appear again, when introducing the idea of semantic holism.
2.2 Iterated Signaling Systems as Evolutionary Stable Strategies

The dynamics that constitute language conventions are good candidates to be maintained in a cultural population because they are part of a local evolutionary-stable-strategy (ESS), which is formally defined as follows [25]:

Let $x$ and $y$ be mixed strategies and let $U$ be the payoff, respectively the reproductive success of a strategy in comparison to another strategy. $U(x, y)$ represents the payoff for strategy $x$ playing against strategy $y$. Then $(x)$ is an ESS if and only if two conditions are fulfilled:

1. $U(x, x) \geq U(x, y)$ holds for all $x \neq y$

This means that if a population of agents plays strategy $x$ and if $x$ is an ESS, no "invader", playing $y$, can enter the population and do better with that.

2. If $(x, x) = U(x, y)$ then $U(x, y) > U(y, y)$

That means: should the invader $y$ play equally successful against $x$ like $x$ would play against itself, then $y$ has to be less successful against itself then against $x$. Both conditions make the infiltration of a sufficiently small group of competing strategies impossible. A locally stable ESS for social institutions (like a public language) is hard to change because movements away from current practice will prove disadvantageous for the agents that choose to play other strategies unless change of conditions.

In the case of language, this means that if certain communicative interactions (words, sentences, speech acts and their meanings) have been selected in a certain population in the past and have proven stable, it will not be rational to change this behavior. For example: If I travel to China and start using Chinese words and sentences in a completely different manner as Speakers of the Chinese language do, this will not prove advantageous for me and it is very hard to imagine that my behavior will proliferate successfully, given no further conditions. The selective conditions of the environment of the population will prevent that.

Adopting an evolutionary perspective in the case of SG, we have to emphasize the historical process leading to a stable convention, which is usually done by a dynamical analysis [19, p. 9]. It is assumed, that each participant of the game has a repertoire of $s$ and $r$ strategies. Transferred into replicator dynamics, iterated signaling systems become ESS. Evolution leads the population to converge to one of the two signaling systems given this set of initial conditions. There can be states of no
communication, partial communication and perfect communication between S and R. Perfect communication is stable, partial communication is vulnerable to neutral drift and no communication is always unstable.

If an adaptive process governs individuals, which are repeatedly confronted with a coordination problem as in SG, one or another signaling system (ESS) will establish, reproduce and become conventional. Small and random fluctuations at the beginning of the process may decide, which one of the two [37, p. 334]. Some forms of the communication (words or sentences respectively their meaning) will proliferate over time, some forms will die out. Players could also learn from each other and blend their information together before it gets reproduced. This would be a small extension of the model [38, p. 92]. Although mi is initially meaningless, it becomes meaningful if agents use it in ways described by a signaling system. No assumptions about an already existing common understanding of the coordination problem and the other players are required. The dynamics are agnostic about the cognitive capacities of the agents. It can be concluded, that the sender and the receiver coevolve and coadapt and the signal’s PM derives from the conventions that facilitate the coordination of both. However, my point in this section was mainly to communicate the general idea behind SG and compare it with a TS understanding of PM, not to describe it in all necessary mathematical detail, which has been done elsewhere. For an educational overview, c.f. [38, p. 72].

It is taken for granted, that the structure of human languages is far more complex than it is captured in these simple signaling systems described so far. But the model is simple and allows for many possible extensions, such as conflict of interest, costly signals [43] or a different probability distribution over states, signals and acts as in [22], and thereby the possibility to depict internal dispositions of senders or receivers. This brings us closer to the debate about speakers meaning and how mental content could be included in this framework. [41] even suggests non-convergent adaptive dynamics, since his investigation seems to show that information transmission can be possible even in the case of meaningless messages in equilibria.

As [18, p. 413], emphasizes: “There is at least one functional aspect of human languages that can fundamentally be expressed in terms of signaling systems: communication facilitates social coordination.”
2.3 Structural Aspects of the Evolution of Signaling Systems in Nature and Culture

Before we will proceed to the next section and focus on some synthetic applications of TS and SG, which can help us to understand them in terms of one another, let me briefly mention some structural aspects of SG in nature. Up to this point, only the functional aspects of PM have been depicted, but since evolution is an interplay between structure and function, it is worth taking a look on the mechanisms that helped signaling evolve in nature.

Skyrms [38, p. 29] echoes the following example. A predatory social bacterium (Myxococcus xanthus) has evolved some interesting microbiological features of both signaling and the emergence of multicellularity [21]. The latter is thought to arise through a gradual process of increasing cooperation between cells. M. xanthus is capable of forming multicellular structures when invading a colony of prey bacteria, such as Escherichia coli. A particular signaling mechanism thereby induces “chemotaxis”, i.e. the movement of an organism in response to a chemical signal. The bacteria colony then forms an aggregation in form of a fruiting body. This development is induced by starvation of some bacteria, which produce chemical molecules, “A-Signals”. A-signal-production mutants have a defect in any one of five asg genes. asgA encodes a protein with a two-component receiver domain followed by a histidine protein kinase domain (ibid.: p. 77). This finding can shed light on the genetic evolution of this particular signaling system. Receivers react to that specific signal if a certain threshold of the chemicals is reached and start their aggregation by operating in one of two possible ways, either forming a fruiting body with differentiations of spores or growth, slowed to a rate compatible with the level of nutrient remaining (ibid.: p.78). In further stages of the process, which is lethal for many bacteria even if it appears to be an ESS, different signals (e.g. “C-signals”) are introduced and serve different functions. But the evolutionary sender-receiver framework of a Lewisian SG can clearly be found in this M. xanthus population structure, even if we will hardly face that much conventionality in such a microbiological setting. To the contrary, it appears a lot like natural salience [38, p. 31]. But when taking a closer look, we can still encounter a sense of the plasticity of signaling. Another case, where this may appear in its most vivid sense when nature (or culture) invents new signals for new environmental situations, where selection operates on a variety of possible signals.

In cultural evolution, things get more messy, since at least some re-
flections on “cultural fitness” (fitness of the agent or fitness of a trait?) cf. [34], cultural reproduction (imitation, emulation or more sophisticated forms of social learning?) and other issues have to be taken into account. Modeling cultural evolution is much more complex and difficult than its biological counterpart and this makes it hard to talk about the structural side of SG and TS. For example: how are payoffs measured in cultural evolution? In the end, this is an empirical question, depending on what you are interested in as a researcher. As [38, p. 55] notes about cultural evolution: “Care in interpretation is required”. A recent empirical study, that proved quite interesting is given by [24] from the MIT. Using a multi-generational signaling game as a laboratory model, they investigated how perceptual and cognitive constraints influenced cultural transmission of melodic structures. Principles like “symmetry, proximity and good continuation” resulted in the accumulation of structural regularities across their cultural “generations” of test subjects, who optimized learnability.

3 Synthetic Applications

In Millikan’s terms, a representation (signal) mediates between producers (senders) and consumers (receivers). Although the terminology differs in detail, the systems can be seen as more or less isomorphic.\(^{10}\) Millikan’s is more fine-grained, she focusses on the selected function (cf. above, section 1.2) of a representational system, and “despite the fact that Skyrms conceives of sender and receiver as part of a whole and Millikan emphasizes the separateness of producer and consumer” [15, p. 198]. In both frameworks, the meaning of signals is determined by the matching of signal production and consumption under selection and iteration. A “synthetic project” (similar to Harms’) of a similar kind can be found in [13, p. 42]. While he focuses on internal or personal meaning, Harms emphasizes the conventional parts of it. When adopting a sender-receiver model of PM, producer and consumer become different agents (mostly but not necessarily), whereas in Millikan’s account they can also be special parts of one and the same representational system, e.g. the brain or parts of the immune system\(^ {11}\) and the signaling system can be regarded as a “distributed structure”. In this section, the terms “intension” and “extension” will be applied to this evolutionary notion of PM in SG. In the philosophy of language, these closely related terms play a crucial role in explanations of meaning and content.
3.1 Intension and Extension in Signaling Games

Since Gottlob Frege [12], meaning is taken as a two-folded concept, consisting of intensional and extensional elements (called “Sinn” and “Bedeutung”; sense and reference). The emergence of these relations can be seen in clear steps. Therefore, the application of those terms to signaling systems can be seen as one important bridge between Millikan’s (more or less classical) theory of meaning and Skyrms’ SG. According to [15, p. 192], the functional basis of any explanation of PM is about the “rules” governing conventional tracking and motivation. Harms investigation stands in the very direction of this paper, also emphasizing the deep similarities between Millikan’s TS and Skyrms’ SG (cf. ibid.: p. 194).

This leads us to the implementation of “intension” and “extension”. According to [15, p. 202], any semantic convention can be seen as a “function-stabilizing mechanism” in an evolutionary sense. This mechanism – Harms refers to it as a “rule” (ibid.: p. 203) – is a map from conditions to processes. To get a more precise notion of the concept, Harms characterizes it set-theoretically as a set of ordered pairs, restricted in a way that each environmental condition maps to one process alone, which is the general structure of a mathematical function. This stands in deep accordance with [29, p. 54], who describes conventional “semantic-mapping functions” (“functions” here in the mathematical sense) as determining truth- and other kinds of satisfaction-conditions.

The rule (R) for an adaptive trait (AT) is generally given by:

\[ R_{AT} = \{ (c_i, p_i) | AT \text{ sel } p_i \text{ in } c_i \} \]

where \( c \) stands for “condition”, \( p \) for “process” and \( sel \) is a “selection operator”. The formula reads literally: “The rule for an adaptive trait is the set of all ordered pairs which condition in the first place and the process in the second, such that the adaptive trait was selected for performing this process in that condition” [15, p. 203]. The process is something that the mechanism does, the conditions may be proximal causes in which the rule maps to a causal chain or states of the world in which the process were performed in an adaptive way.

When implementing the terms “intension” and “extension”, at least two sub-rules can be derived from (1).
3.1.1 Extension

The intension of a (potentially) meaningful signal points to the consequence within the system (you may also call it implication, entailment or interpretation), whereas the extension points out to the world or world-states (like reference, representation or satisfaction). The intension is the relationship between the internal functions of a representational system, the extension explains what the signal indicates. In the framework of SG, a possible sub-rule of (1) for a signal producing mechanism P (extension) is a correspondence map from the states of the world $W = \{\sigma_1, \ldots, \sigma_i\}$ to the signals $S = \{m_1, \ldots, m_i\}$.

(2) $R(\text{extension})_P = \{\langle \sigma_i, s_i \rangle | P \text{ sel } m_i \text{ in } \sigma_i\}$

The extension of a signal can be seen as the adaptive target, which is determined by behavioral efficiency. The truth conditions for a signal are the normal conditions in which the “proper function” (Millikan) is executed. Note that the same selective formalization as in (1) is used to describe the special case of signal production.

3.1.2 Intension

For the interpretation C of the message by the receiver, a second sub-rule can be derived where signals elicit a set of behavioral responses/acts $A = \{a_1, \ldots, a_i\}$:

(3) $R(\text{intension})_C = \{\langle m_i, a_i \rangle | C \text{ sel } a_i \text{ when } m_i\}$

The sub-rules (2) and (3) are able to implement the concepts of “intension” and “extension” in SG [15, p. 204]. What a signal “means”, extensionally and intensionally, constitutes the content. In the system, there are rules for the issuance of the signals and rules for their response. The two sets of rules must fit together as part of a unified design. PM reaches over both communicative agents in their evolutionary dynamic and cannot be fully understood by only looking at one of them, e.g. S. Only if S’s and R’s behavior reach a certain level of coordination, signals can be said to be “meaningful”, and this PM also involve the receivers corresponding act. Note that the difference of intension and extension tells us nothing about the difference between speakers meaning and PM and differs from this distinction.
3.2 The Distinction between Normative and Descriptive Statements

It is good philosophical consensus that ought-statements are not derivable from is-statements and vice versa. The first would be a naturalistic fallacy, the second a moralistic fallacy. I do not want to argue against both. However, a full and appropriate theory of meaning should help us to understand the relationship between what we call “is” and “ought”, and the evolutionary emergence of these particular signals. In this section, I am going to say a little about this relationship.\textsuperscript{13}

3.2.1 Primitive Content or “Pushmi-Pullyu” Representations

A basic assumption is that difference between indicatives and imperatives is regarded as follows: imperatives motivate behavior while indicatives indicate world states. Let us start with a well-known example from the literature on SG: the alarm calls of vervet monkeys. [5] showed that vervet monkeys use to play SG to warn group members of local predators that they became adapted to. In the standard case\textsuperscript{14}, they have three alarm call signals:

- “cough” ($m_1$) for an eagle ($\sigma_1$)
- “bark” ($m_2$) for a leopard ($\sigma_2$)
- “chutter” ($m_3$) for a snake ($\sigma_3$)

Three different kinds of behavior are executed by the receiver monkeys:

- “hide under bush” ($a_1$) if ($m_1$)
- “run up a tree” ($a_2$) if ($m_2$)
- “stand upright and check ground” ($a_3$) if ($m_3$)

That means the game is entirely symmetric. It is easy to imagine that once the behavior became conventional, a confusion would prove fatal.

From a human (philosopher’s) perspective, we could ask ourselves: What exactly do the three signals \textit{mean}? Do they mean the word state that they refer to (extension) or the behavior that ought to be the correct response (intension)? Does $m_1$ mean “Eagle” or “Everyone, hide under a bush”? To put the question in another way: are the signals indicatives or imperatives? The answer is that they are both. The vervet signals have descriptive and directive functions at the same time. Simple SG
are not structured in a way that would allow us to categorize them in terms of is and ought.

Millikan has a special name for such a kind of representation: “pushmi-pullyu” representation [29, p. 166]. She emphasizes: “purely descriptive and purely directive representations are forms requiring a more sophisticated cognitive apparatus to employ them than is necessary for these primitives.” (ibid). The more primitive a language system is, the less can is and ought clearly be distinguished.\(^{15}\) Harms calls the content of such animal signals “primitive content” (PC), featuring indicative and imperative functions at the same time [15, p. 189]. In simple systems, the extension is not specified through subject-predicate structure but via condition of efficacy. The intension is just the behavior, following from the signal. According to him, PC is the key to understand the distinction between is and ought, since it enables us to characterize animal signals like warning cries semantically although they are not translatable into human languages. [18] takes up this point and provides a formal analysis of the difference between indicatives and imperatives via “deliberation”, which shows again the deep similarities between TS and SG. But how and when do indicatives and imperatives part, e.g. when does a signal become “truly” normative?

Harms [15, p. 206] applies the same selective formalization as in the case of “intension” and “extension” – cf. equation (2) & (3). The key is to introduce another layer of semantic adaptation, a regulatory mechanism. Imperative statements regulate social behavior via the enforcements of “rules” and conventions.\(^{16}\)

Suppose, a vervet monkey tries to “deceive” the others by sending a false signal – e.g. \(m_1\) – even though no eagle is near. He may do that because he wants a tempting food resource all for himself and expects the rest of the group to hide under bushes when receiving his signal.\(^{17}\) In this context, we can think of a kind of stabilizing mechanism SM, e.g. given by a corrective signal cs, which forms a higher semantic level in the regulatory adaptive hierarchy. It could be a kind of behavior of cheater detectors, maybe something like aggressive behavior as “punishment” for the cheater. An evolutionary “arms race” between cheaters and detectors will probably start and the signaling system will increase in complexity. But the general “rules” of adaptive implementation stays the same.

Formally speaking [15, p. 208]: The rule for any adaptive trait [cf. equation (1)] fails, when a process (message) is not accompanied by the condition (world state), which it was selected for. The “failure” of the rule can be depicted as:

(4) Failure of \( R_{AT} = (\sigma_i \land \neg m_i) \) where \( \langle \sigma_i, m_i \rangle \in R_{AT} \)

For a kind of SM, the correspondence rule for its corrective signal \( CS = \{cs_1, \ldots, cs_i\} \) is given by:

(5) \( R_{SM} = \{\langle \sigma_i \land \neg m_i \text{ where } \langle \sigma_i, m_i \rangle \in R_{AT, cs}\} \mid \text{SM sel cs when } (\sigma_i \land \neg m_i)\} \)

That reads: “The rule for a stabilizing mechanism is the set of all ordered pairs with the failure of the rule for an adaptive trait in the first place and a corresponding corrective signal in the second place, such that the stabilizing mechanism was selected for sending the corrective signal if a message is not sent when the relevant world state occurs.”

If the adaptive trait (i.e. the signal-act pair) fails to execute its proper function (Millikan), SM itself is selected for detecting this failure of an adapted behavior and motivating a corrective signal/action to restore it. Already quite early in her work, Millikan [27, p. 23] speaks of “first-order and higher order reproductive established families”. Conventions are are first-order families, whereas corrective signals could be seen as higher-order families. Often, they are not direct copies of one another but still serve the same proper function, namely to restore a first-order proper function. Cs would be “true” (where “true” means “adaptively successful”), if the convention it enforced was violated, so that it would be about the historically established convention and an individual’s relationship to it [17, p. 449].

4 Semantic Holism in Signaling Games?

“For a large class of cases of the employment of the word ‘meaning’—though not for all—this word can be explained in this way: the meaning of a word is its use in the language.”

– L. Wittgenstein [42, 43]

In this last chapter, I want to discuss the following question: Does a synthetic theory of PM, which incorporates TS as well as SG, lead us to some kind of semantic holism? I think it does, but in a slightly different sense then it is regularly seen.

Semantic holism (related but not isomorph to “semantic externalism”) argues that many parts of languages can only be said to carry meaning in terms of other larger segments of this very language, they cannot be understood in isolation. Usage of words and sentences are highly interconnected, and draw their meaning from this connectedness,
which does not even have to involve another word, but could also be any kind of (repeatable) behavior. This idea can be traced back to [42]. For a more detailed discussion on semantic and mental holism, cf. [3]. For a defense of meaning holism, cf. [14]. Here, I want to emphasize that in the case of SG, this behavior can be found in the receivers' acts. Surprisingly, we find hardly any evidence for semantic holism in SG literature at all,\(^\text{18}\) which is why I want to address the issue here by asking if the framework of SG leads us in exactly this direction.

Just to make myself more clear and avoid misunderstanding: by semantic holism, I do not mean, that a sign (signal) carries information about external events (world states), i.e. reference about the world. This is quite uncontroversial and can be captured within the notion of “extension” (section 3.1.1). I mean something quite stronger, namely that the signal is not the only entity in a SG, that carries information and meaning, i.e. content.

![Figure 2](image.png)

Figure 2: The two signaling systems of the standard Lewisian SG, depicted as dashed ellipses. These two possible meaning conventions are state-message-act triples, involving S’s message reference as well as R’s behavioral response as crucial adaptive elements. Therefore and in this sense, conventional meaning is a “distributed structure”.

As figure 2 depicts, the two meaningful conventions – the two signaling systems – are constituted by the sender’s messages and the corresponding receiver’s acts. The reason why signaling system one or two evolved, its’ conventional meaning, lies is the mutual co-adaptation of the behavior of both communication partners. If one wanted to explain why and how this convention has been established in a community of (proto-) speakers, one had to tell a story about the successful iteration –
successful because of the payoff – of the sender sending a message and the receiver responding to it in a correct way. It is crucial to understand, that the receivers act does not have to be a behavior (like running up a tree), it can also be an utterance, like a verbal answer or gesture.

This is why one could argue that PM is a property of the signaling system in total, not only the sender’s message. In the particular evolutionary history, not only the signals got iterated and implemented a proper function (i.e. TS-meaning), the acts were iterated and adapted as well. They too, carry information, if information flow is understood in moving probabilities of word-states and signals, cf. section 2.1.3. This is so obvious that it almost seems trivial, yet it is not really mentioned in the literature. If “anything that changes the probability of another event carries information about it” [13, p. 43] then the receivers acts (a1/a2) also carry information, namely about the previous signals and/or the world states. In an iterated game where senders and receivers learn to play properly (i.e. moving towards a signaling system), it is the act of the receiver leading to the mutual benefit (payoff) of both.

As an exemplar case, consider the vervet monkeys, again. For a receiver monkey to become a sender monkey in a later stage of the game, the acts of receivers can also function as a kind of signal (it is not really a signal) for others. Imagine a group member D observing only the receiver’s respond to a signal, not the signal itself:

\[
S \Rightarrow R \Rightarrow D
\]

R’s behavior – e.g. if a leopard is near (σ2) – will be climbing a tree as fast as possible (a2). This behavior will be D’s primary source of imitation, his emulative stimulus. If she also climbs a tree, she reacted properly given σ2, although information was not directly transmitted between her and S. Certainly, information about the world (σ2) was transmitted, namely between R and D (and indirectly between S and D, as well). In the future, D will run up a tree if σ2 occurs. She took part in the convention, even if she overheard m2. But if that is true, then a2 carries as much information as m2, at least in iterated games or in scenarios, where roles are interchangeable, or in a little more complex networks of information transmission. And if that is true, then a very strong meaning holism is required in SG (as well as in TS). In this sense, PM is a physically distributed structure, ranging from S to R (and in our case: D). The convention as a whole carries meaning and information, not only the message.

Critiques of semantic holism resp. defenders of semantic atomism [7] [10] often emphasize that semantic holism is unable to explain how two
speakers can mean the same thing when using the same linguistic expression. Therefore it is unclear, how language learning and communication is even possible, given the limits of our cognitive capacities and that a sentence can only be understood in terms of the totality (the “whole”) of the given language. This may be true, but a theory of SG can probably help to clarify the picture. As we saw in this paper, speakers do not really have to mean the same thing (personal meaning) in order to participate in a public and conventional language, only a certain amount of coordination is required. Furthermore, understanding a signal’s meaning does not need to involve the totality of a given language on the one hand (understanding to what specific behavior/act is was connected in its evolutionary history is enough), but can also not be restricted to what the signal indicates (world-state) on the other hand. If one has to give a name to this special kind of semantic holism, one could label it “evolutionary holism“ or just “conventional holism”.

**Conclusion**

The synthetic conclusion of this paper can be summarized as follows: A set of statistical processes, e.g. the combined strategy \( s_1 \) (\( m_1 \) if \( \sigma_2, m_2 \) if \( \sigma_2 \)) / \( r_1 \) (\( a_1 \) if \( m_1 \); \( a_2 \) if \( m_2 \)) is a semantic convention and therefore carries PM (given as a holistic distributed structure), iff the advantageous trait (payoff) came from coordinated and iterated reproduction in a signaling game.

Comparing two prominent functional approaches on the evolution of conventional meaning – TS and SG – shows the deep similarities between them, but also some minor differences, esp. in methodology and in their understanding of “information”. Clearly, SG has a lot more to offer in terms of empirical applicability. It follows that PM (in contrast to personal meaning) can be fully understood in terms of its cultural evolutionary function in populations of communicators, though it remains an empirical question on what structural features the functional and formal dynamics of the particular signaling game are built upon.

Furthermore, the classical concepts of *intension* and *extension* can be implemented in this synthesis of Millikan’s and Skyrms’s framework, via the two-folded understanding of an evolutionary semantic convention, which indicates information about the world and motivates behavior at the same time. Since this semantic links were forged by adaptation, one could also think about a second level of regulative mechanisms, that were selected to enforce already existing conventions failing to execute their
proper functions and these higher corrective levels can be seen as one possible origin of the semantic is-ought distinction.

Last but not least, the synthetic framework can be read in favor of a specific semantic holism, which seems to indicate that in the case of PM, meaning is not only a property of signals in equilibria, but a property of the whole convention. This “conventional holism” as I would call it, could probably solve some issues of classical semantic holism, but that would overextend the scope of this particular investigation.

**Acknowledgements**

I would also like to thank two anonymous referees for helpful comments on an earlier version of this paper.

**Notes**

1 Millikan distinguishes three sub-kinds of PM [29, p. 54]: “(1) conventional linguistic cooperative functions, to be called ‘stabilizing functions’. (2) conventional semantic-mapping functions (‘functions’ in the mathematical sense) which determine truth- and other kinds of satisfaction-conditions. (3) methods of identifying, to be called ‘conceptions’ and ‘conceptual components’, that govern individual speakers’ grasps of referents and of truth- or satisfaction-conditions, hence help to determine their dispositions to use and understand various conventional language forms.”

2 For an objection, c.f. e.g. [35].

3 “[. . . ] we do not have to endow the sender and receiver with a pre-existing mental language in order to define a signaling system.” [38, p. 7]

4 A basic game theoretical introduction for biological systems can be found in Maynard Smith, J. [25]. For a proper introduction into (evolutionary) game theory and several applications for a generalized evolutionary framework, cf. [37, p. 323].

5 Two questions become pressing for Lewis in this context: 1. How is a conventional language, reps. a convention of any kind, established in the first place? 2. How is it maintained in a population? If the two candidates of the convention are entirely symmetric (concerning their costs and payoffs) this is far from clear. If there is no reason to prefer one of the two strategies, why will both players choose the same one and why will they not switch without “communicating their communication” before. Any form of “preplay communication” [37, p. 353] will not solve this problem because it is easily shown that it leads to an infinite regress if you want to explain the very emergence of communication. It is also not clear, how much knowledge the players need to have about each other to recognize an equilibrium as salient. The answer to these problems lies in the switch from standard game theory (which is based on the assumptions of rational choice) to evolutionary game theory.
6 A Nash equilibrium is a combination of strategies in which no player gains something by changing its strategy if the other player does not change it as well. A strict Nash equilibrium is a Nash equilibrium where the player would even do worse by changing.

7 To be precise, Skyrms [38, p. 36] uses the Kullback-Leibler distance to estimate an overall measure of information of signals and states, given by a weighted average. It is given by: $I_{states}(signal) = \sum_i p_{sig}(state_i) \log[p_{sig}(state_i)/p(state_i)]$

   How much a signal moves the probability of the second state (in a two state-act SG) can be determined relatively to how much it moves the first and vice versa. One can plot the average signal’s information as a function of the probability of the first state given the signal.

8 Replicator Dynamics or population dynamics are the most commonly used method for the formal study of evolutionary development in biological as in cultural evolution. Based on differential equations, which model the change in frequency of competing variants (e.g. genes or cultural traits) over time, the status ($s$) change of a system is given by a difference equation like $\Delta s/\Delta t$ at some point in time ($t$), which is a function $\varphi$ of the system state and a set of parameters $\alpha$. For more details, cf. [37, p. 275].

9 How can iterated reproduction look like in cultural evolution? First of all, evolutionary SGs suggest that not the agents but their behavior will be reproduced. But in the simple Lewisian Game described so far, things are subject to a high amount of idealization. Transmission mechanisms themselves can introduce many biases in the process, such as content bias (based on intrinsic attractiveness), model bias (prestige, age, similarity) or frequency dependent bias (conformity or anti-conformity). All these “forces” strongly influence change in cultural reproduction of the population dynamics over time. For an overview, cf. [26, p. 57]. However, these biases could in principle always be implemented into the framework of SG. The key is to introduce probabilities for the activities of sender and receiver. If, for example, within a network of senders and receivers, a very prestigious sender is imitated by another one of lower status, the probability for him to reproduce the signal will be higher than average (model bias).

10 “Neither [Millikan nor Skyrms] assumes words, combinatorial syntax […]. Rather, both assume that the most basic vehicle of meaning is a sort of monolithic signal […].” [15, p. 198], brackets added.

11 S/R Configurations are found within and across organisms, but as [13, p. 46] notes correctly: “In general, as we move further from the within-organism case, breakdowns of common interest become more likely.”

12 “So generally, the intentional properties of any conventionally meaningful sign include both (1) relations regarding externally mapped object(s) or what a representation “corresponds” to and (2) relations regarding internally mapped representations or events, logical consequences, definitional implications, and the like. We have substantial vocabularies for both sorts of relationships.” [16, p. 33].

13 This section can be regarded as a sub topic of the evolution of moral norms and moral behavior. A game theoretical analysis of the evolution of morality in the broad sense and the evolution of moral signals in the narrow sense can also be found in [17, p. 434].

14 There are plenty other examples of signaling systems special kinds of systems for vervet monkeys, and also for other animal species. For an overview cf. [18, p. 413].
That does not mean, that they disappeared in complex modern human languages, where many utterances still are pushmi-pullyu representations, e.g. the sentence: “Johnny, we do not eat peas with our hands!” [29, p. 167]

An interesting analogy from biological evolution are structural and regulatory genes, i.e. regulatory mechanisms for a preexisting system.

This is not a thought experiment. Monkeys doing exactly this kind of cheating have been empirically observed in the wild [38, p. 74].

For example, [38, p. 40] discusses a variety of cases, e.g. how SG can help us determine the quantity of information as well as the informational content of signals about the word as well as “about the act” (ibid.: p. 35). But he hardly discusses the fact, that also acts can carry information.

Simple learning dynamics can easily be implemented in SG, e.g. reinforcement learning [38, p. 12].

Karim Baraghith
DCLPS (Duesseldorf Center for Logic and Philosophy of Science)
University of Duesseldorf
Universitaetsstr. 1
40225 Duesseldorf, Germany
<kbaraghith@phil.uni-duesseldorf.de>

References


[35] A. Rosenberg, „How Jerry Fodor slid down the slippery slope to Anti-Darwinism, and how we can avoid the same fate,“ *European Journal for Philosophy of Science* 3/1, pp. 1-17, 2013.


