The evolution of compositionality and proto-syntax in signaling games

Michael Franke

Abstract. Compositionality is a key design feature of human language: the meaning of complex expressions is, for the most part, systematically constructed from the meanings of its parts and their manner of composition. This paper demonstrates that rudimentary forms of compositional, even proto-syntactic communicative behavior can emerge, without stipulating sophisticated or purposeful agency, from a variant of reinforcement learning applied to signaling games. This helps explain how compositionality could have emerged *gradually*: if unsophisticated agents can evolve prevalent dispositions to communicate compositional-like, there is a direct evolutionary benefit for adaptations that exploit the systematicity in form-meaning mappings more rigorously.

> It is astonishing what language can do. With a few syllables it can express an incalculable number of thoughts, so that even a thought grasped by a terrestrial being for the very first time can be put into a form of words which will be understood by someone to whom the thought is entirely new. This would be impossible, were we not able to distinguish parts in the thoughts corresponding to parts of a sentence, so that the structure of the sentence serves as the image of the structure of the thought.
>
> (Frege, 1923)

Language is undoubtedly a key element in the success of mankind to transmit complex knowledge from one generation to the other. Language enables such gradual extension of culturally accumulated knowledge in part because of its flexibility to express a virtually limitless space of ideas. What gives a language this expressive power is a semantic property called *compositionality*. A language is compositional (in the relevant sense considered here) if the meanings of (most of) its composite expressions (e.g., phrases made up of words) are systematically derived from the meanings of their parts and the way in which these parts are combined.

This paper is concerned with the question how compositionality could have evolved, and it offers a mathematical model, couched in game theory, that shows how the very beginnings of rudimentary compositionality in signal use can arise by a variety of reinforcement learning. Obviously, since compositionality is such a fundamental design feature of language, many other contributions exist in the literature that also try to show how compositional language use could have evolved. The contribution of this paper is to show how early compositional-like signal use can arise without a presupposition of questionable cognitive capacities or behavioral learning dispositions of language using agents; capacities and dispositions that are themselves not independently justified by evolutionary arguments. Only once compositional-like signal use is present (even if infrequent), will there be grip for biological selection of traits that exploit and refine compositional form-meaning mappings. This argument will be spelled out in Section 1, where I defend a general *methodological minimalism* in evolutionary modeling, according to which the presupposed cognitive abilities of agents should be kept minimal in order to explain how a particular to-be-explained feature can evolve gradually.

Section 2 then introduces *signaling games* and recapitulates relevant results from learning and evolutionary dynamics in these games. Section 3 discusses the Barrett-Skyrms model of compositional signaling and argues that it fails to capture an element of creativity, spontaneity or stimulus generalization. Against this background, Section 4 introduces *spill-over reinforcement learning* and shows that spontaneous compositional-like signal use can arise under favorable parameter settings, but is not a guaranteed outcome. Section 5 further explores the potential of spill-over reinforcement learning to model order-sensitive signaling, which may be conceived as a form of proto-syntax.

1 Compositionality & methodological minimalism

It is important to start with a general remark about what compositionality is. The statement that human language *is* compositional is possibly misleading, and perhaps even false. Even if human language was indeed compositional, this could never be directly observed. Rather, compositionality is an assumption —and a good one as such, in terms of its explanatory power— that serves to explain a number of observables about human language, namely its productive and interpretative flexibility, its systematicity and its learnability, among others (Pagin and Westerståhl, 2010a,b). This is also reflected in the opening quote from Frege's *Gedankengefüge*: it is the assumption that language is compositional that explains productive expression of new ideas and the ability to understand novel sentences. (Frege's suggestion that such productivity and systematicity *require* compositionality can be called into question; after all, even if we cannot presently conceive of a better explanation, that does not mean that our best explanation is necessarily true.)

This remark about the nature of compositionality is important, because it tells us what we need to be looking for in a model that aspires to account for the evolution of compositionality. Since compositionality as such cannot be directly observed in nature, we cannot observe it directly in a model either. But we can observe in the behavior of modeled language users those observable features of human language that make us postulate compositionality as an explanatory principle. In particular, we should look for conditions under which agents evolve behavioral dispositions of using complex signals spontaneously and creatively in a systematic way to express complex meanings that they have never encountered before, and similarly evolve behavioral dispositions to react to complex signals in a way that is systematically related to their reactions to the parts of the complex signals. It is this *creative compositionality*, as a behavioral disposition of signaling agents, that I would like to account for, using minimal assumptions about the agents' cognitive abilities.

Why the fuss about minimal cognitive abilities? Because otherwise there is likely a gap in our understanding of the gradual emergence of compositionality. To see this, it helps to consider alternatives.

The standing assumption in evolutionary modeling of language origins is that behavioral patterns of language use that give higher chances of communicative success lead to higher expected fitness and will therefore increase in proportion over time. Chance of communicative success is itself a function that depends on the behavior of the other members in the population. This is why language use is a *game problem*: the success of the speaker's behavior depends on the listener's behavior and vice versa. As a consequence, this means that adaptations that change the way language is used and interpreted face a *threshold problem*: the number of mutations required before a novel trait, like compositional language use, can be fitness increasing is rather high, especially in realistic scenarios where each agent only interacts with a limited number of other agents and where spurious cognitive abilities may incur a maintenance cost. To wit, a single agent who speaks a private language that is bad for communicating with everyone else would likely die a lonesome death, even if this particular communicative behavior would be highly efficient if used by more than one individual. This highlights that an explanation of the *gradual* emergence of compositionality is important.

Previous accounts have tried to account for the evolution of compositionality, but have not sufficiently explained its gradual emergence. Often, a distinction is made between vertical and horizontal evolution. *Vertical evolution* happens over many generations where *biological selection* pushes population towards a higher proportion of fitter behavioral types. In contrast, *horizontal evolution* takes place within a few overlapping generations, or usually just one generation, by individual learning and *cultural adaptations*, due to the plasticity of the agents' behavioral dispositions. The evolution of compositionality has been addressed both as an issue in vertical and horizontal evolution.

Adopting a vertical perspective, Martin Nowak and colleagues used evolutionary game theory to show that compositional language use is evolutionarily advantageous in the presence of noise (Nowak and Krakauer, 1999; Nowak, Plotkin, and Jansen, 2000): if signals are transmitted through a noisy channel, then the encoding of meanings in a compositional way makes the reception of signals less error-prone. If we accept this premis, it is clear that compositional language has an evolutionary advantage over non-compositional language. But the *threshold problem* remains unsolved: how could compositional language use arise and spread in the first place, e.g., in a small population?

Accounts that adopt a horizontal perspective on language evolution appear to offer more grip on the threshold problem, because they zoom in, so to speak, on the agent level and the agents' respective behavior. There are consequently many accounts using agent-based simulations that aim to address the evolution of compositionality (e.g.

Oliphant and Batali, 1997; Batali, 1998; Gong, 2007; Tria, Galantucci, and Loreto, 2012). It is impossible to do justice to all contributions that file under this approach. Instead, I would like to pick one example, not because it is particularly worthy of critique —in fact, my target is quite ingenious, informative and has consequently been quite influential—, but because it serves to make a conceptual point about hidden assumptions in agent-based modeling.

Batali (1998) uses recurrent neural networks to model agents' production and interpretation behavior and shows that compositional language use can arise from repeated interaction in his setup. Each agent is represented as a neural network that maps a character string onto a bit vector that represents a point in a semantic space. A network thus represents the agent's interpretation strategy. Following the lead of Hurford (1989), Batali assumes that an agent's production behavior is built on top of his interpretation behavior: when an agent wants to express a meaning given by a bit vector, he constructs a character string by, essentially, performing a heuristic search through the space of all character strings and selecting the first one that the agent himself would interpret as a meaning that is close enough to the one that he wants to communicate. This is sometimes called an *obverter strategy* and is also implemented in many instantiations of iterated learning models that produce compositional form-meaning mappings (e.g. Kirby and Hurford, 2002; Kirby, 2002; Smith, Kirby, and Brighton, 2003).

On the face of it, neural networks may appear to be the most innocuous and natural modeling choice, without danger of endowing agents with a surplus of unwarranted abilities. But the obverter strategy is far from being innocuous. It implements production behavior that is goal-oriented and geared towards efficient communication. It also implements an implicit bias towards a close alignment between interpretation and production (which was the original reason for Hurford (1989) to introduce it). And it even carries an implicit bias towards compositional meaning mappings, because it will preferably choose form F for meaning M if the mapping of F to M is similar to mappings of other form-meaning pairs that is has been trained on.

Agent-based simulations like Batali's (and much subsequent work in the iterated learning tradition) are very insightful, because they give sufficient conditions for the emergence of form-meaning mapping. Knowing about multiple sufficient conditions is important to obtain a complete picture, and each successful computational model will tell us about more sets of sufficient conditions. But as long as the sufficient conditions tested include assumptions like the obverter strategy, for whose evolution there is no independent justification ready at hand, there is also something missing. It is in this sense, that an approach to horizontal evolution of compositionality should make *minimal* assumptions about the agents' sophistication, in order to show how the phenomenon could gradually evolve: once proto-compositionality emerges (at least sporadically) in the behavior of unsophisticated signalers, there is an incentive to evolve learning or usage strategies that exploit these features, riding piggy-back on the clever-seeming behavior of dumb agents.

To conclude, we are left with two desiderata. For one, we would like to see agents that are in some sense capable of reacting to *novel* stimuli in a systematic, compositional-like way. For another, we would like to make minimal assumptions about these agents' abilities to cognize, learn and tune their behavior towards more communicative efficiency. In the following I would like to demonstrate that behavioral dispositions towards seemingly clever, compositional-like creative signal use can emerge in signaling games under a variant of reinforcement learning, in which agents lack the ability to sharply discriminate between similar meanings and similar forms.

2 Signaling games & reinforcement learning

Signaling games were invented by David Lewis to give a naturalistic account of conventional meaning (Lewis, 1969). A signaling game features two players, a sender and a receiver. One round of play proceeds as follows. A chance move determines which state of the world obtains. The sender observes the state, the receiver does not. The sender then selects a signal, or message, to send to the receiver. The receiver observes the message and chooses an act in response. In the simplest case, the play was a success for both players just in case the act matches the state.

If T, M, and A are (finite) sets of states, messages and acts respectively, then behavioral strategies for sender and receiver are functions:

$$\sigma \in (\Delta(M))^T \qquad\qquad \rho \in (\Delta(A))^M .$$

Behavioral strategies map choice points of agents to a probability distribution over the choices they have at that choice point. The sender makes a choice which messages to send for each state; the receiver chooses an act for each message.

Let $\mathrm{U} : T \times A \to \mathbb{R}^{\geq 0}$ be a non-negative utility function that only depends on the state and the receiver's act. The expected utility of a strategy pair is:

$$\mathrm{U}(\sigma, \rho) = \sum_{t \in T} \sum_{m \in M} \sum_{a \in A} \Pr(t) \cdot \sigma(t, m) \cdot \rho(m, a) \cdot \mathrm{U}(t, a) ,$$

where $\Pr(t)$ is the occurrence probability of state t. A strategy pair $\langle \sigma, \rho \rangle$ is an *evolutionary stable state* (in a two-population setting, where sender and receiver behavior can be independently adjusted) iff it is a strict Nash equilibrium, i.e., iff there is no σ' with $\mathrm{U}(\sigma', \rho) \geq \mathrm{U}(\sigma, \rho)$ and there is no ρ' with $\mathrm{U}(\sigma, \rho') \geq \mathrm{U}(\sigma, \rho)$ (Selten, 1980).

Under the simplifying assumption that we have the same number of states, messages and acts, and that $\mathrm{U}(t_i, a_j) = 1$ if $i = j$ and 0 otherwise, signals can be said to acquire a clear-cut meaning in evolutionary stable states, being associated with exactly one corresponding state-act pair (Wärneryd, 1993; Blume, Kim, and Sobel, 1993). Much recent work has been devoted to complementing the picture of meaning evolution in signaling games by supplying accounts of abstract evolutionary dynamics that show under which conditions populations of agents can be expected to converge on a meaningful use of initially meaningless signals (e.g. Huttegger, 2007; Pawlowitsch, 2008; Huttegger et al., 2010; Huttegger and Zollman, 2011; Franke and Wagner, 2014).

Next to studying vertical evolution of signaling strategies, there is also an increasing understanding of horizontal evolution in terms of agent-based learning dynamics (c.f. Skyrms, 2010; Huttegger and Zollman, 2011, for overview). A particularly appealing notion of learning, since austere and well-established, is reinforcement learning (c.f. Barrett, 2007, 2009; Mühlenbernd, 2011, for applications to signaling games). In its simplest form, reinforcement learning in signaling games can be conceived of as a dynamic process using Polya-urns. Players have one urn for each choice point, filled with balls corresponding to each act. For instance, at any point in time, $B(t, m)$ is the number of balls for message m in the sender's urn for state t, and similarly $B(m, a)$ is the number of balls for act a in the receiver's urn for message m. Urns map onto behavioral strategies:

$$\sigma(t, m) = \frac{B(t, m)}{\sum_{m' \in M} B(t, m')} \qquad \rho(m, a) = \frac{B(m, a)}{\sum_{a' \in A} B(m, a')} .$$

Under basic Roth-Erev reinforcement learning (RL), sender and receiver adapt their behavioral dispositions after each play of the game (Roth and Erev, 1995). If the last round of play involved t^*, m^*, and a^* and yielded a payoff of $\mathrm{U}(t^*, a^*) = u^*$, then the new urn content $B'(\cdot)$ is like the old one $B(\cdot)$ except that u^* balls for the chosen type are added to the urns that were visited, so as to increase the chance of repeating more successful behavior in the future:

$$B'(t, m) = \begin{cases} B(t, m) + u^* & \text{if } t = t^* \text{ \& } m = m^* \\ B(t, m) & \text{otherwise} . \end{cases} \tag{1}$$

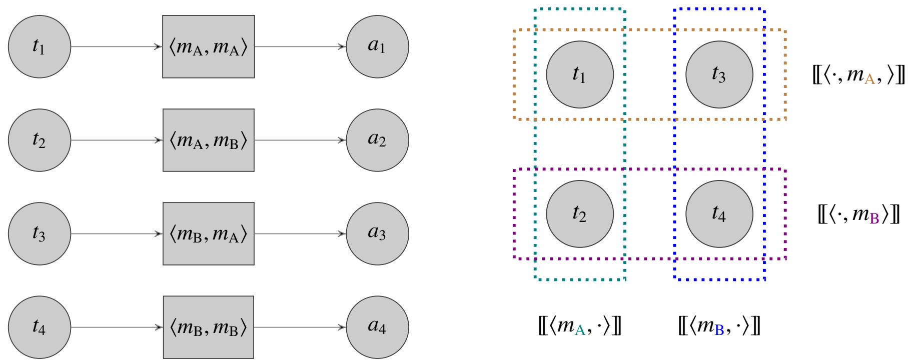
Likewise for the receiver.

Basic reinforcement learning between a single sender and a single receiver leads to meaningful signaling behavior with certainty in the limit if the signaling game has two messages, acts and equiprobable states and Lewis-style utilities (Argiento et al., 2009). For more states, messages and acts, simulations show that convergence to perfectly communicative signaling is also high, but decreasing with the size of the game. For three messages, acts and equiprobable states, Barrett (2007, 2009) reports that ca. 95% of simulation runs, starting with urns that all had one ball for each choice, converged to perfectly communicative signaling behavior. Mühlenbernd (2011) showed that larger populations of reinforcement learners also acquire successful signaling behavior, albeit with the possibility of showing regional variations across their network of interaction.

All of this is conceptually interesting, especially because RL makes very little assumptions about cognitive abilities of agents. Reinforcement learners need not be aware that they are playing a game, or that there is any other player around. Reinforcement learning also doesn't presuppose that agents know what the game is about, so to speak, in that they would know the utilities. Agents are merely trying to gradually optimize their behavior by repeating more frequently what showed good results in the past. It is also a misconception to think that reinforcement learning is a form of supervised learning. The only requirement is that agents receive feedback about their payoff from the last round of play. If this information is not always available, so be it; then our reinforcement learner will update his urns only when it is available. If information about success or failure is never available, then adjustments of behavioral dispositions towards more successful behavior seem hardly possible at all.

3 Compositionality from reinforcement learning

Reinforcement learning can also lead to behavior that looks like a compositional use of signals. Barrett (2007, 2009) introduced an extension of Lewis' signaling games with



(a) Stable strategy profile.

(b) Compositional meaning of signals.

Figure 1: Example of a strategy pair for a syntactic game that achieves perfect information transmission (left), together with a possible decomposition of the signal's meaning (right).

two senders instead of one. Skyrms (2010) saw that we can equally treat this setup as one in which the same sender sends two signals consecutively. The receiver responds to a pair of signals in either set-up. There are then cases in which RL leads agents to use pairs of signals *as if* each of its components has a singular meaning that is combined, by meaning intersection, into the meaning of the pair.

In a *syntactic game*, as Barrett calls it, there are four world states, $t_1 \ldots t_4$, and four receiver acts, $a_1 \ldots a_4$. As before, the utility function for sender and receiver is the same, namely $\mathrm{U}(t_i, a_j) = 1$ if $i = j$ and 0 otherwise. There are also two messages, m_A and m_B. Unlike before, the sender can send a pair of messages in each state, and the receiver conditions his reaction on each pair. There are thus four complex signals, and therefore perfect, non-redundant information transfer is possible. A strategy profile that achieves this is pictured in Figure 1a. Numerical simulations show that basic RL almost always leads to such an evolutionary stable state.

What is interesting about this case is that it allows us to decompose the meanings of complex signals, as manifested by the behavior of agents, into parts associated with each component. Figure 1b shows this for the strategy pair of Figure 1a. In general, in order to achieve perfect information transmission, the pair of signals must communicate two bits of information. Since each component of the pair has two values, m_A and m_B, it is as if each position is independently communicating one bit of information. More concretely, in the example in Figure 1b, it is as if the first component of the signal pair gives an answer to the question, conceived as a partition of states, $\{\{t_1, t_2\}, \{t_3, t_4\}\}$; the second component answers the *orthogonal* question $\{\{t_1, t_3\}, \{t_2, t_4\}\}$ (c.f. Lewis, 1988, for this notion of orthogonality). In this sense, basic RL is able to evolve behavior that looks, from the outside, as if complex signals are meaningful in a systematic way, combining the meaning of their parts by meaning intersection.

Since only basic RL is used, it is clear that the second desideratum with which Section 1 concluded is met. There is no sense in which agents' cognitive abilities have been unduly amplified in order to see the desired behavior emerge. But the first desideratum is not met. The story so far leaves no room for the claim that agents have acquired a disposition to use complex signals compositionally beyond the ones that they have already seen. Similarly, there is no reason to assume that the component parts of complex signals are independently meaningful *to the agents*. In other words, since syntactic games constitute a closed system, there is no role for productivity and generalization. Such a closed system offers little to reduce the threshold problem, because any mutant learner type with an ability to generalize to novel situations would still play against a population of agents who will only be able to use complex signals in a compositional-like way that they have learned as if they were unstructured wholes. The problem, then, lies in the nature of basic RL in that it does not allow for any sort of stimulus generalization.

4 Spill-over reinforcement learning

There are many variants of RL beyond the basic variety. There is *negative reinforcement* that penalizes unsuccessful behavior by subtracting balls from urns. (We did not allow for this in our formulation of basic RL in Equation (1), because we required utilities to be non-negative.) There is forgetting of past reinforcements (c.f. Barrett and Zollman, 2009; Mühlenbernd, 2011, for applications to signaling games). There is *lateral inhibition* in which balls are removed from an urn if they are not of the type of the successful act that was played during the last round, and in which balls of the successful type are removed from other urns (c.f. Franke and Jäger, 2012, for an application to signaling games). But there is also a natural way of endowing reinforcement learners with the potential to acquire behavioral dispositions for choice points that have never been encountered before. I call this *spill-over reinforcement learning* (c.f. O'Connor, 2014, for another application to account for vagueness in signaling games).

Spill-over RL is a conservative generalization of basic RL. It does not presuppose that agents are any smarter or more sophisticated than under basic RL. If anything, it presupposes that they might be less sophisticated. The basic idea of spill-over RL, expressed in terms of urns and balls, is that whenever balls of type b are added to some urn u, balls of type b' are also added to urns u' in proportion to how similar b is to b' and u is to u'. Such spill-over to similar choice points and acts is what can be expected of agents who are not able to properly distinguish the relevant contingencies.

To implement spill-over RL for signaling games, let's assume that there is a given similarity metric on the state space, and one on the message space. Let $0 \leq \mathrm{Sim}(t, t') \leq 1$ be the similarity of state t and t'. For our purposes, it suffices to assume that similarity is symmetric and maximal for identical states. The same applies to similarity between messages and between acts.

If t^*, m^* and a^* have been used in the last round of play for payoff $\mathrm{U}(t^*, a^*) = u^*$, then spill-over RL has the sender update his behavioral dispositions like so:

$$B'(t,m) = B(t,m) + \mathrm{Sim}(t,t^*) \cdot \mathrm{Sim}(m,m^*) \cdot u^* \,. \qquad (2)$$

Likewise for the receiver. It is obvious that spill-over RL reduces to basic RL in case all non-identical states and messages are maximally dissimilar. If not, reinforcements spill-over, so to speak, to similar choice points and actions, as if the agent is not able to distinguish properly between the effects of some act and related acts, or to distinguish the effect of some act at similar choice points. It is also clear that spill-over RL allows behavioral dispositions to change at choice points that have never been encountered before, and also for choices that have never been utilized before. In this sense, spill-over RL does allow for a very low-level form of stimulus generalization, without this necessarily being any higher cognitive attempt of learning a rule or a pattern. A similar stimulus generalization happens in neural networks, as in Batali's (1998) approach, where similar inputs tend to trigger similar outputs simply due to the nature of representation and information flow in a simple neural network (c.f. O'Connor, 2014).

To get an impression of how spill-over RL works in signaling games, let's look at the simplest but non-trivial extension of the syntactic games of Barrett and Skyrms in which something like creative compositionality could evolve. Let there be three states, labeled A, B and C, that we consider simple. Let there be three states that we consider complex, labeled AB, AC, and BC. For the time being, let's interpret complex states as some composition of the simple states. For instance, if t_{A} is a state that involves birds, and if t_{B} is a state that involves water, then state t_{AB} is a state that involves both birds and water. Notice that under this interpretation of states, only the presence of A and B matter, but not their order. We will look at an extension of this basic model in which order is relevant for states and messages in Section 5. Other interpretations than the one suggested here are possible, and we will come back to this issue in Section 6.

As for messages, let's assume that there are similarly three simple and three complex messages, labeled just like the states. For the time being, interpret simple messages as plain emission of a single sound or the production of a single ostensive sign; complex messages are pairs of signals where order of production of signals does not matter. In other words, the message m_{AB} should be conceived of as a production of signals m_{A} and m_{B} in close temporal succession in arbitrary order. As stated before, there is much room for the interpretation of states and messages, and we will explore this issue critically later on.

In all that follows, I will simply assume that the receiver's acts *are* the states, and that payoffs are as in Lewis' games: $\mathrm{U}(t, t') = 1$ if $t = t'$ and 0 otherwise.

In order to apply spill-over RL to a signaling game with

$$T = \{t_{\mathrm{A}}, t_{\mathrm{B}}, t_{\mathrm{C}}, t_{\mathrm{AB}}, t_{\mathrm{AC}}, t_{\mathrm{BC}}\} \qquad \text{and} \qquad M = \{m_{\mathrm{A}}, m_{\mathrm{B}}, m_{\mathrm{C}}, m_{\mathrm{AB}}, m_{\mathrm{AC}}, m_{\mathrm{BC}}\}$$

we need to give a similarity measure on T and one on M. Not to make any stronger commitments about the nature of complex states and messages than necessary, let's simply assume, on top of the natural constraints of similarity, that any two different simple states/messages are maximally dissimilar and that complex states have parametrized similarity $s \in [0; 1]$ to their component parts. As a consequence of the former assumption, any two complex signals/messages will also be maximally discriminable (as they contain a clearly distinguished element). To keep matters simple, I will do as if the same level of similarity s governs state similarity and signal similarity. It will transpire that dispensing with this simplifying assumption is not difficult, but also not important. Together, this gives the following table of similarities, that applies to both states and messages:

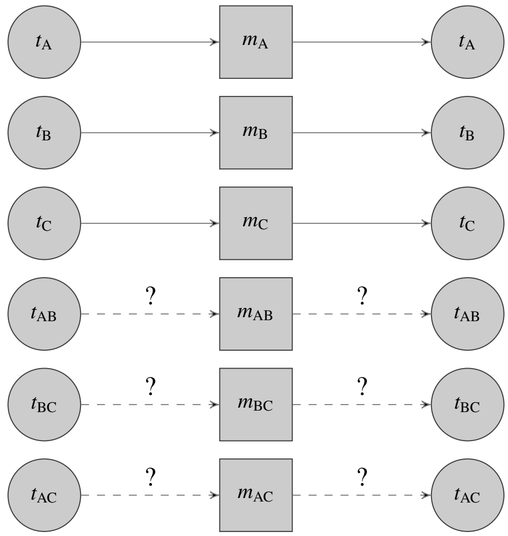


Figure 2: When playing the game with only simple states and messages, we are interested in whether agents can implicitly acquire a behavioral disposition that favors a compositional-like mapping of complex states and meaning.

	A	B	C	AB	AC	BC
A	1	0	0	s	s	0
B	0	1	0	s	0	s
C	0	0	1	0	s	s
AB	s	s	0	1	0	0
AC	s	0	s	0	1	0
BC	0	s	s	0	0	1

What we are after is an account of creative compositionality. So in order to make sure that we can test agents' dispositions to act in novel situations, we do need to tinker a little bit with the normal flow of playing and learning in a signaling game. In particular, we need to have agents play some restricted game, that does not make all of the states and messages available from the start. We are then interested in the agents' acquired dispositions in choice points that they have never seen before, and their acquired disposition to use acts that so far were not available to them by modeller's fiat. The most natural way of restricting the signaling game at hand, is to have agents play with only the simple states and messages and to check their evolving implicit dispositions to use and interpret complex states and messages (see Figure 2). The labeling of complex states and messages, which informs the assumed similarity structure, then defines which association of complex states and messages would count as a systematic compositional mapping.

Notice that under the assumed similarity structure, the game restricted to simple states and messages only involves maximally dissimilar contingencies (by assumption), and so we know from Barrett's simulation results that the emergence of a perfectly successful communicative behavior is almost certain. In order to give an analytical argument why spill-over RL can give us creative compositionality, let's therefore suppose that the agents have in fact acquired a perfect code to communicate about simple states with simple messages. Modulo renaming, we are in the situation depicted in Figure 2. The question to be asked then is what are the implicitly acquired behavioral dispositions involving complex states and messages?

It suffices to analyze the case for the sender, because the receiver part is parallel. Likewise, the argument for all three complex states is identical, so we can focus on the sender's behavioral dispositions at choice point t_{AB}. The following table lists the number of balls corresponding to the messages in the columns that are added to urn t_{AB} when the state-message pairs in the rows are used (by assumption: with the effect of successful information transmission):

	m_A	m_B	m_C	m_{AB}	m_{BC}	m_{AC}
$\langle t_A, m_A \rangle$	s	0	0	s^2	0	s^2
$\langle t_B, m_B \rangle$	0	s	0	s^2	s^2	0
$\langle t_C, m_C \rangle$	0	0	0	0	0	0

If states occur equally frequently, then, averaging out all stochastic fluctuations, we get the expected non-normalized proportions of balls in the t_{AB}-urn as the sum of each column. Can the creatively compositional-like choice of m_{AB} in t_{AB} be the most likely choice? Yes, it can. It is whenever $s < 2s^2$, i.e., whenever $s > .5$. In other words, there are parameter values for which a spontaneous compositional-looking choice of complex signals can be the most likely event in an hitherto unseen choice point. This is so, even though the agent is not extrapolating a rule from his use of signals, or reasoning about the most likely interpretations that different signals will receive by an anticipated listener (as under the obverter strategy). In sum, this limit argument suggests that creative compositionality *can* arise as the most likely event for favorable parameter constellations under spill-over RL.

The critical reader will have many concerns at this moment. I will focus here on a few obvious worries of a more technical nature and postpone any deeper conceptual issues until Section 6. Firstly, that the creative compositional choice is the *most* likely does not mean that it is likely. True, but it does not have to be for the argument that matters, namely that spill-over RL can pave the way for an immediate adaptive benefit of a mutation that exploits the stochastic prevalence of compositional-like regularity with greater fidelity. Secondly, the argument given presupposes that the "base language" is in securely in place, and used over a very long time. Otherwise stochastic noise cannot be satisfactorily excluded, and the creative compositional choice might not be the most likely after all. As I will show presently with numerical simulations, this worry is legitimate but not fatal. Even if the "base language" is learned imperfectly over a relatively small period, the creative compositional choice can still trump all others. Thirdly, my argument only shows that *there are* parameter values that give the desired result, but not that the desired result is necessary under spill-over RL. This is true, and I believe that this is as it should be. If compositional-like signaling was a near certainty under a low-level learning account like spill-over RL, we would presumably have to reject the approach for predicting wrongly that compositional-like signaling should be much more widespread in nature than it appears to be the case.

A further technical worry is that *lateral inhibition* might spoil the picture. Lateral inhibition is sometimes used in applications of RL to signaling games (e.g. Mühlenbernd, 2011; Franke and Jäger, 2012), and frequently in agent-based approaches, such as the *naming game* (e.g. Loreto, Baronchelli, and Puglisi, 2010). In a setting that allows for perfect discriminability of choice points and acts, lateral inhibition kicks in after a successful trial. If that involved state t^* and m^*, lateral inhibition removes balls of type $m \neq m^*$ from the t^*-urn, and m^* balls from all other urns. Similarly for the receiver. But this presupposes perfect discriminability of choice points and acts. Under spill-over RL, the assumption that choice points and acts are imperfectly discriminably would also have to affect lateral inhibition. It could be suspected that lateral inhibition might therefore counteract the effects of spill-over RL and hamper the emergence of creative compositionality.

There is only one way to address this challenge: define spill-over RL with lateral inhibition and check whether creative compositional-like choices can still trump all others. If t^*, m^* and a^* have been used in the last round of play for payoff $\mathrm{U}(t^*, a^*) = u^*$, then spill-over RL with lateral inhibition has the sender update his urns like so:

$$\begin{aligned} B'(t, m) = \max(0, B(t, m) &+ \mathrm{Sim}(t, t^*) \cdot \mathrm{Sim}(m, m^*) \cdot u^* \\ &- (1 - \mathrm{Sim}(t, t^*) \cdot \mathrm{Sim}(m, m^*)) \cdot u^* \cdot i)\,. \end{aligned} \tag{3}$$

Here, $i \in [0; 1]$ is another parameter that captures the strength of lateral inhibition. For $i = 0$, the definition above reduces to the previous Equation (2). If $i > 0$ and $u^* > 0$, lateral inhibition removes m-balls from t-urns proportionally to how similar m and t are to m^* and t^*. Removal of balls can never lead to a negative number of balls in an urn. Notice also that no changes occur if $u^* = 0$, so that lateral inhibition does not introduce negative reinforcement through the backdoor.

Here is another mean field argument parallel to the previous one, but now also involving possible lateral inhibition. As before, the following table lists the number of balls corresponding to the messages in the columns that are added to the t_{AB}-urn when the state-message pairs in the rows are used, if that number is non-negative:

	m_{A}	m_{B}	m_{C}	m_{AB}	m_{BC}	m_{AC}
$\langle t_{\mathrm{A}}, m_{\mathrm{A}} \rangle$	$s - (1 - s)i$	0	0	$s^2 - (1 - s^2)i$	0	$s^2 - (1 - s^2)i$
$\langle t_{\mathrm{B}}, m_{\mathrm{B}} \rangle$	0	$s - (1 - s)i$	0	$s^2 - (1 - s^2)i$	$s^2 - (1 - s^2)i$	0
$\langle t_{\mathrm{C}}, m_{\mathrm{C}} \rangle$	0	0	0	0	0	0

Again, the non-normalized proportions of ball types in the noise-free long run are given by the sums of each column. A spontaneous compositional-like choice of m_{AB} is the modal outcome under the mean field expectation if:

$$i < -\frac{2s^2 - s}{2s^2 - s - 1}\,.$$

Figure 3 pictures the area of the two-dimensional parameter space for which that condition is met. This means that lateral inhibition *can* counteract the effect of similarity spill-over in the emergence of creative compositionality, but it does not undermine the effect entirely. If $s > .5$, there is always some upper bound on i below which creative compositionality is the most likely outcome.

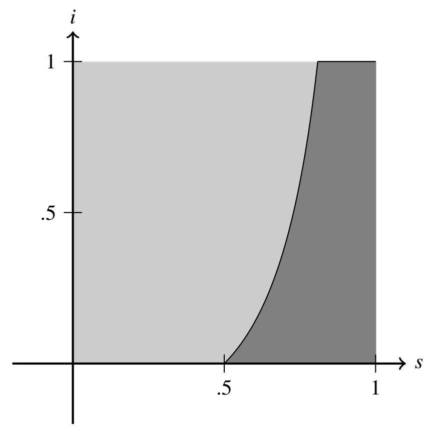


Figure 3: The dark gray region is the subset of parameter values for spill-over RL with lateral inhibition for which a spontaneous compositional-like choice is the most likely outcome in the mean field expectation.

Mean field arguments inform us about idealized long-term behavior. It remains to look at short-term dynamics. This can be done by numerical simulations. Figure 4 shows the results of a modest parameter sweep. For all pairs of parameter values $s \in \{.5, .65, .8\}$ and $i \in \{0, .2, .4\}$, 40 trials of spill-over RL were simulated. For each of these, all sender and receiver urns had one ball of each type in the beginning. Strategies then evolved by actually playing the game and updating urns, as described by Equation (3). Strategies were recorded after a fixed number of rounds. During the first 10,000 rounds, the game was restricted to the simple states and messages, but implicit dispositions involving complex states and messages were tracked. Only the results of the sender strategies are reported here, because the receiver side is entirely parallel. For each trial and round, the sender's evolved strategy was classified as belonging to exactly one of the following disjoint and exhaustive categories: (i) non-separating, (ii) separating, but non-compositional, and (iii) compositional. A sender strategy is *separating* if for each simple state there is a unique simple message, such that the message is the most likely choice for the state. A sender strategy is *compositional* if it is separating and the sender's most likely choices in complex states are exactly the complex messages that would count as compositional-like signaling under the mapping of simple states to simple messages, suggested by the strategy being separating. (When a sender strategy is not separating, it is moot to ask whether it is compositional, because we would not know what the systematic compositional-like mapping of complex states to complex messages should be.) Figure 4 then gives the proportions of trials whose sender strategy falls into the three relevant categories.

Ideally, we would like to see parameter values for which most or all trials are separating and compositional. The worst case, for our purposes, is the second category: a separating but non-compositional signal use. In other words, we would like to see little white and a lot of dark gray in the plots in Figure 4, for at least some parameter values.

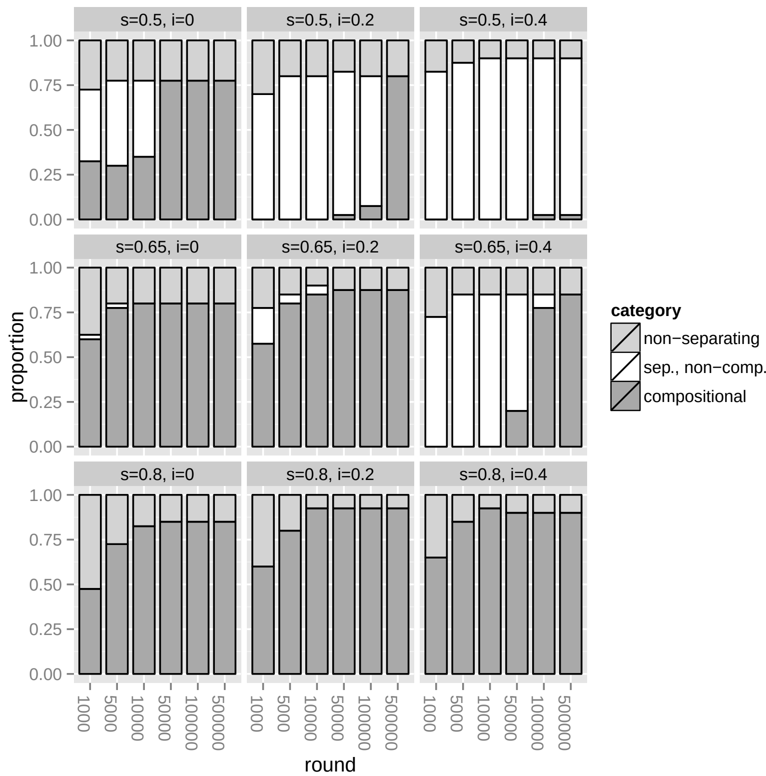


Figure 4: For each combination of tested parameter values, the plots show the proportion of trials with a sender strategy that fell into each of the three categories (see main text for definitions).

And this is indeed what we find. There are pairs of parameter values that almost guarantee creative compositionality already at round 10,000, where senders have never before seen complex states or used complex messages. This is the case for high enough spill-over paired with low enough lateral inhibition: the picture that ensues from numerical simulation is perfectly consistent with the previous mean field analysis. This suggests that there are parameter values for which spontaneous compositional-like signaling can emerge, even if agents had only little time to acquire a “base language” from scratch. (Obviously, the restriction to an initial “training phase” to acquire a “base language” is artificial, but we need some such restriction in order to test whether agents have evolved relevant dispositions involving previously unseen choice points and previously untried acts.)

To conclude, spill-over RL does not guarantee the emergence of compositional-like signaling behavior, but almost guarantees it for certain values of spill-over and lateral inhibition. Spill-over RL does not presuppose that agents use or interpret signals in a smart, introspective or strategic manner. The compositional-like dispositions can be seen as a concomitant of an inability to sharply distinguish contingencies. The relative prevalence of compositional-like signaling can therefore create an environment in which any mutation that exploits the systematicity in form-meaning mappings more rigorously would be immediately adaptive. Something like spill-over RL, i.e., any horizontal learning dynamic, which can produce rudimentary systematicity in form-meaning mappings by chance alone, therefore helps explain how compositionality could have evolved gradually. Other low-level learning dynamics could deliver similar results. No argument for the supremacy of spill-over RL over potential alternatives is implied. The point of this paper is merely to give a proof of existence.

5 Order-sensitivity: towards syntactic signaling

The kind and amount of compositionality that spill-over learning produces is very basic. Although I would maintain that this is exactly as it should be for a gradualist account of evolving compositionality, there is a sense in which something of substance in still missing. As long as signals are unstructured sets of signals, like bags of words, the only meaning compositions that can be accounted for with spill-over RL are commutative meaning compositions, like meaning intersection. To pick up the previous example, if t_{A} is a state in which birds are present, and if t_{B} is a state in which water is present, then order-insensitive t_{AB} could be a state in which birds and water are present, in whatever way. This is a commutative meaning combination, and it was shown to pair, possibly compositional-like, with order-insensitive emission of signal pairs.

But there are ways of combining meanings that are not commutative. In English, the compound noun *water bird* denotes a type of bird, and the compound noun *bird water* denotes a type of water: quite different things. It would be speculative to say that commutative meaning compositions are more basic than non-communitative ones. At the same time, what is most interesting about compositionality in human languages is that the systematicity in form-meaning mapping is governed in large extent by the *syntactic structure* of the form. The story told so far falls short of touching on anything like structured signaling. These considerations suggest that it would be desirable to see whether the present approach, using spill-over RL in signaling games, could also possibly extend to something like an explanation of the evolution of a proto-syntax that informs semantic interpretation: if not fully structured signals, then at least order-sensitive signals that are mapped, as if a compositional syntax-semantic rule is followed, onto non-commutative meaning compositions.

The following paragraphs offer a few preliminary ideas on how a gradual evolution of "proto-syntactic compositionality" could be explained in terms of a low-level learning account like spill-over RL. The logic of the overal argument is basically the same as before. If spill-over RL can produce, as the most likely choices, spontaneous behavior that looks like a generalization of a syntax-driven form-meaning mapping, then this helps explain the gradual emergence of such a phenomenon, because in an environment where such systematicity in signaling exists, mutations that exploit the systematicity more rigorously can be immediately adaptive, thus alleviating the threshold problem. The amount of syntactic regularity that agents should implicitly track should ideally be small, at least to start with. The following therefore explores a case in which the kind of syntax-driven form-meaning mapping is really very rudimentary: I would like to show that order-sensitive compositional-like form-meaning mappings can spontaneously evolve under spill-over RL.

The simplest example in which this could occur is just an extension of the game that we looked at in the previous section, only that we now need to distinguish complex states and messages with respect to the order in which their simple parts occur. So, we now have a game with:

$$T = \{t_A, t_B, t_C, t_{AB}, t_{BA}, t_{AC}, t_{CA}, t_{BC}, t_{CB}\} \quad \text{and}$$
$$M = \{m_A, m_B, m_C, m_{AB}, m_{BA}, m_{AC}, m_{CA}, m_{BC}, m_{CB}\} \ .$$

For instance, m_{AB} is a sequence of messages where first m_A is produced and then m_B, whereas m_{BA} is a sequence of messages where first m_B is produced and then m_A. Similarly, t_{AB} is a state that differs from t_{BA}, for instance, in that either the first or the second occurring state is the main meaning component, to which the other meaning component is added as a modifier. Again, I would like to postpone until later a more in-depth reflection on what states and messages could represent, and why these and only these are taken into account here.

In order to apply spill-over RL to this case, measures of similarity are necessary. For continuity, it is best to keep all previous assumptions in place, except to make room also for similarity between states and messages that share a component part in the same position. Barring a better motivated choice, I propose to use a parameter $r \in [0; 1]$ that captures similarity between order-sensitive complex pairs as inverse Hamming distance, which gives us:

	A	B	C	AB	BA	AC	CA	BC	CB
A	1	0	0	s	s	s	s	0	0
B	0	1	0	s	s	0	0	s	s
C	0	0	1	0	0	s	s	s	s
AB	s	s	0	1	0	r	0	0	r
BA	s	s	0	0	1	0	r	r	0
AC	s	0	s	r	0	1	0	r	0
CA	s	0	s	0	r	0	1	0	r
BC	0	s	s	0	r	r	0	1	0
CB	0	s	s	r	0	0	r	0	1

A minimal situation in which agents could show behavior that looks like they are exploiting an acquired order-sensitive form-meaning mapping is shown in Figure 5. The setup assumes that a "base language" is in place, and that also a pair of form-meaning mappings is in use that determines what would count as an order-sensitive compositional form-meaning mapping. The question then is: will spill-over RL produce the systematic order-sensitive compositional-like choice for unseen choice points?

A mean field argument, similar to the previous one, suggests that it can, at least for some values of parameters s, i and r. With the form-meaning mapping that is assumed

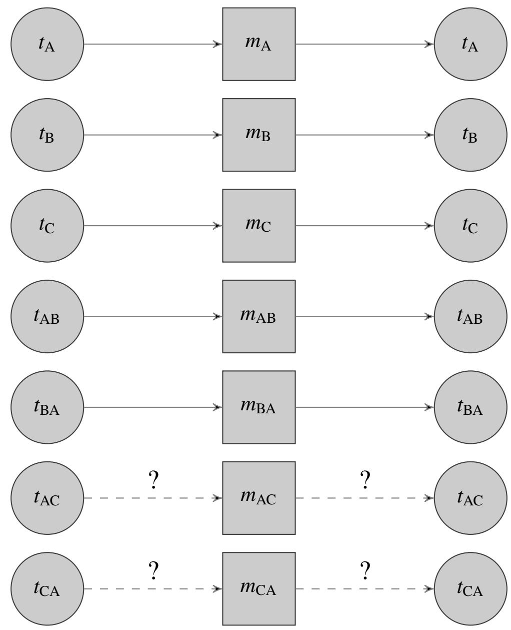


Figure 5: If a "base language" is in place, and if at least one pair of complex states and messages fixes the potential order-sensitivity, can spill-over RL bestow implicit dispositions on agents that looks like they have generalized the order-sensitive compositional form-meaning mapping?

to be in place, the changes in the urns for the state t_{AC} are (with the implicit proviso that entries in this table are never negative):

	m_A	m_B	m_C	m_{AB}	m_{BA}
$\langle t_A, m_A\rangle$	$s-(1-s)i$	0	0	$s^2-(1-s^2)i$	$s^2-(1-s^2)i$
$\langle t_B, m_B\rangle$	0	0	0	0	0
$\langle t_C, m_C\rangle$	0	0	$s-(1-s)i$	0	0
$\langle t_{AB}, m_{AB}\rangle$	$rs-(1-rs)i$	$rs-(1-rs)i$	0	$r-(1-r)i$	0
$\langle t_{BA}, m_{BA}\rangle$	0	0	0	0	0

	m_{AC}	m_{CA}	m_{BC}	m_{CB}
$\langle t_A, m_A\rangle$	$s^2-(1-s^2)i$	$s^2-(1-s^2)i$	0	0
$\langle t_B, m_B\rangle$	0	0	0	0
$\langle t_C, m_C\rangle$	$s^2-(1-s^2)i$	$s^2-(1-s^2)i$	$s^2-(1-s^2)i$	$s^2-(1-s^2)i$
$\langle t_{AB}, m_{AB}\rangle$	$r^2-(1-r^2)i$	0	0	$r^2-(1-r^2)i$
$\langle t_{BA}, m_{BA}\rangle$	0	0	0	0

Under the (certainly debatable) assumption that all five of the available states are equiprobable, the expected number of balls of each type in the urn for t_{AC} is, again, simply given by the sum of each column. We see immediately that a choice of m_{AC} is always at least as probable as a choice of m_{BA}, m_{CA}, m_{BC} and m_{CB}. This means that a choice of a signal that would look like the order-sensitive form-meaning rule had been applied in reverse is never the most likely. Moreover, there are parameter triples for

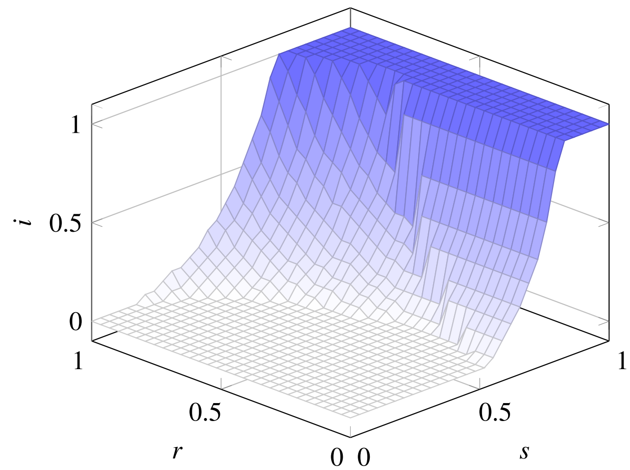


Figure 6: The area below the surface is the part of the parameter space for which the sequential messages m_{AC} is the most likely choice in state m_{AC} in the mean field of playing with the restricted base language and two sequentially consistent compositional state-message pairs.

which m_{AC} is indeed the most likely choice of signal. Figure 6 shows the subspace of the three-dimensional parameter space in which rule-like order-sensitive signaling is the modal response as the area under the plane. As before, spill-over parameters s and r need to be sufficiently high, especially $s > .5$, and lateral inhibition i must not be too high. If we look at the special case of $r = s$, the case actually reduces to the previous one, and the subplane of the parameter space where rule-like order-sensitive signaling is the modal choice is the dark gray area in Figure 3, as before.

This shows that it is possible to see the beginnings of something like order-sensitive compositionality arise spontaneously from spill-over RL. Admittedly, there are many open issues. The following section tends to some of the most pressing.

6 Reflection

Let us take a step back and ask why the models presented here evolve compositional-like behavioral dispositions, at least for some parameter values. This is important in order to deflect the potential criticism that a disposition towards compositional-like signaling was deliberately engineered-in.

It is obvious that the key element for emerging dispositions for spontaneous compositional-like signaling is spill-over. Lateral inhibition is not responsible for the evolution of creative compositionality. It even slightly hampers its emergence. Spill-over was defined in terms of two similarity measures: one between states and one between messages. Reinforcements then percolate, so to speak, to similar states and similar messages. Agents are likely to choose acts in choice points that they have never used or seen before if these acts are similar to acts that have proven effective at similar choice points. This is a simple and general principle of learning under imprecise delineation of contingencies. It is not an unmotivated design to make agents acquire goal-oriented language-like behavior.

For contrast, consider what spill-over RL is *not* doing. It is not looking at a stock of observed form-meaning pairs, and extrapolating a general rule or pattern from it. It is not inspecting the similarity between a state and a signal directly. It is also not looking at the similarity between any state-meaning pair that it has learned to associate and another state-meaning pair that is has not yet learned to associate. Decisions under spill-over RL are not based on reasoning about the effects that different choices could have (as in the obverter strategy), thus presupposing that agents know that they are playing a game with another player and what the goal of the game is.

Although this is not necessary, spill-over of behavioral reinforcements can be seen as something smart too. Particularly the order-sensitive parameter r from the model of Section 5 is arguably less innocuous than the simple spill-over parameter s. This raises an interesting question about the evolution of different abilities to track similarities.We should consider evolutionary competition among different individual learning rules, and ask whether certain forms of similarity-tracking are likely to evolve. Exploring this fascinating and important issue must unfortunately be left to another occasion (but see Zollman and Smead, 2010; Smead and Zollman, 2013, for related work).

The models I presented could be criticized for their abstractness, in particular for the lack of a more specific commitment about the interpretation of simple and complex states and signals. To a certain extent, I have committed myself to an interpretation already. Simple messages are single emissions of independently individuated sounds, gestures or the like. Complex signals are pairs of such signals either order-insensitive or order-sensitive. This is rather concrete. It gets fuzzier on the meaning side, as is to be expected. Complex states are states in which the individuating elements of simple states are combined, either commutatively or non-commutatively. This is vague and allows for all sorts of interpretations of simple states and meaning operations. But at the same time, any more specific interpretation would suggest that this interpretation was most relevant or prevalent in the course of evolution of compositionality. I would not know presently how to defend any such claim, and therefore must leave this unspecified.

Nevertheless the interpretation of the game models raises further interesting and critical concerns. Firstly, it is debatable as to whether compositionality evolved from combining more basic simple elements creatively into bigger chunks, or whether it was rather the reverse, i.e., that there were complex signals first associated with complex states of affairs, whose component parts were then gradually deemed more and more meaningful in their own right (thereby changing the complex-complex mappings as well).Actually, the model presented in Section 3 is compatible with both views, because of the properties of the similarity relation. In spite of this, I feel more comfortable with the first interpretation, because it seems more natural under the assumption that agents have minimal information processing powers. To assume that agents use long signals that are individuated by their parts and their composition seems to presuppose more cognitive ability.

These latter considerations also raise a more foundational issue concerning the individuation of contingencies in a game model. This problem has many faces. Here is a very tangible one:why does the model of Section 5 where signal pairs are order-sensitive also include repetitions of simple signals, such as m_{AA}? The more general formulation of this problem is: how do we make sure that the contingencies in the game model (choice points and acts) are reasonable descriptions of the way that the agents themselves represent the situation? (Here "representation" does not mean "conscious representation"; even basic RL presupposes that choice points and acts are represented somehow.) To answer this problem in its entirety is way beyond the scope of this article. The problem is shared by many if not almost all other models of language evolution. Conceived in this way, we are brought back to the previous issue of studying the evolution of similarity measures. Even more generally we would ideally like to account for the co-evolution of perceptual categories and signaling behavior. Since so many factors will play a role in such an extensive model, this is a giant's task. Enough reason to stick to something more workable yet insightful first, like the modest models presented here.

There are other recent approaches that nicely complement the material presented here. Steinert-Threlkeld (2014) discusses a model where basic RL can assign a functional meaning component to part of a complex signal. More concretely, a special signal can be sent and combined with another independently meaningful signal. Given Steinert-Threlkeld's set-up, sender and receiver behavior can evolve that assigns the special signal the function of negating the complement signal. This is very interesting but calls for further exploration, because the set of available interpretations of the special signal is as of yet rather restricted and geared towards an interpretation as a function word. Alternatively, Brinkhorst (2014) introduces a further extension of urn-based reinforcement learning in which senders can learn to send arbitrarily complex structured signals. In Brinkhorst's model the sender's signaling behavior is defined by an extensible set of urns, which variably determine which if any signal to send next and which urn to visit subsequently. This way, senders essentially evolve a probabilistic finite-state automaton that generates structured signals whose complexity is upper-bounded only by the receiver's co-evolving ability to "parse".

References

Argiento, Raffaele et al. (2009). "Learning to Signal: Analysis of a Micro-Level Reinforcement Model". In: *Stochastic Processes and their Applications* 119, pp. 373–390.

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

— (2009). "The Evolution of Coding in Signaling Games". In: *Theory and Decision* 67, pp. 223–237.

Barrett, Jeffrey A. and Kevin J. S. Zollman (2009). "The Role of Forgetting in the Evolution and Learning of Language". In: *Journal of Experimental and Theoretical Artificial Intelligence* 21.4, pp. 293–309.

Batali, John (1998). “Computational Simulations of the Emergence of Grammar”. In: *Evolution of Language: Social and Cognitive Bases*. Ed. by James R. Hurford, Michael Studdert-Kennedy, and Chris Knight. Cambridge, UK: Cambridge University Press.

Blume, Andreas, Yong-Gwan Kim, and Joel Sobel (1993). “Evolutionary Stability in Games of Communication”. In: *Games and Economic Behavior* 5.5, pp. 547–575.

Brinkhorst, Sanne (2014). “On Evolution of Compositionality”. MA thesis. Universiteit van Amsterdam.

Franke, Michael and Gerhard Jäger (2012). “Bidirectional Optimization from Reasoning and Learning in Games”. In: *Journal of Logic, Language and Information* 21.1, pp. 117–139.

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

Frege, F. L. Gottlob (1923). “Gedankengefüge”. In: *Beiträge zur Philosophie des deutschen Idealismus. Band III*. Deutsche Philosophische Gesellschaft, pp. 36–51.

Gong, Tao (2007). *Language Evolution from a Simulation Perspective: On the Coevolution of Compositionality and Regularity*. Chinese University of Hong Kong.

Hurford, James R. (1989). “Biological Evolution of the Saussurean Sign as a Component of the Language Acquisition Device”. In: *Lingua* 77.2, pp. 187–222.

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

Huttegger, Simon M. and Kevin J. S. Zollman (2011). “Signaling Games: Dynamics of Evolution and Learning”. In: *Language, Games, and Evolution*. Ed. by Anton Benz et al. LNAI 6207. Springer, pp. 160–176.

Huttegger, Simon M. et al. (2010). “Evolutionary Dynamics of Lewis Signaling Games: Signaling Systems vs. Partial Pooling”. In: *Synthese* 172.1, pp. 177–191.

Kirby, Simon (2002). “Learning, Bottlenecks and the Evolution of Recursive Syntax”. In: *Linguistic Evolution through Language Acquisition: Formal and Computational Models*. Ed. by T. Briscoe. Cambridge University Press, pp. 173–204.

Kirby, Simon and James R. Hurford (2002). “The Emergence of Linguistic Structure: An Overview of the Iterated Learning Model”. In: *Simulating the Evolution of Language*. Ed. by A. Cangelosi and D. Parisi. Springer, pp. 121–148.

Lewis, David (1969). *Convention. A Philosophical Study*. Cambridge, MA: Harvard University Press.

— (1988). “Relevant Implication”. In: *Theoria* 54, pp. 162–174.

Loreto, Vittorio, Andrea Baronchelli, and Andrea Puglisi (2010). “Mathematical Modeling of Language Games”. In: *Evolution of Communication and Language in Embodied Agents*. Ed. by Stefano Nolfi and Marco Mirolli. Springer-Verlag. Chap. 15, pp. 263–281.

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

Nowak, Martin A. and David C. Krakauer (1999). “The Evolution of Language”. In: *PNAS* 96, pp. 8028–8033.

Nowak, Martin A., Joshua B. Plotkin, and Vincent A. A. Jansen (2000). “The Evolution of Syntactic Communication”. In: *Nature* 404.30.

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

Oliphant, Michael and John Batali (1997). "Learning and the Emergence of Coordinated Communication". In: *Center for Research on Language Newsletter* 11.1.

Pagin, Peter and Dag Westerståhl (2010a). "Compositionality I: Definitions and Variants". In: *Philosophy Compass* 5.3, pp. 250–264.

— (2010b). "Compositionality II: Arguments and Problems". In: *Philosophy Compass* 5.3, pp. 265–282.

Pawlowitsch, Christina (2008). "Why Evolution does not Always Lead to an Optimal Signaling System". In: *Games and Economic Behavior* 63.1, pp. 203–226.

Roth, Alvin E. and Ido Erev (1995). "Learning in Extensive Form Games: Experimental Data and Simple Dynamical Models in the Intermediate Term". In: *Games and Economic Behavior* 8.1, pp. 164–212.

Selten, Reinhard (1980). "A Note On Evolutionarily Stable Strategies in Asymmetric Animal Conflicts". In: *Journal of Theoretical Biology* 84, pp. 93–101.

Skyrms, Brian (2010). *Signals: Evolution, Learning, and Information*. Oxford: Oxford University Press.

Smead, Rory and Kevin J. S. Zollman (2013). "The Stability of Strategic Plasticity". Unpublished manuscript.

Smith, Kenny, Simon Kirby, and Henry Brighton (2003). "Iterated Learning: A Framework for the Emergence of Language". In: *Artificial Life* 9, pp. 371–386.

Steinert-Threlkeld, Shane (2014). "Learning to Use Function Words in Signaling Games". In: *Proceedings of* Information Dynamics in Artificial Societies IDAS-14. Ed. by Emiliano Lorini and Laurent Perrussel.

Tria, Francesca, Bruno Galantucci, and Vittorio Loreto (2012). "Naming a Structured World: A Cultural Route to Duality of Patterning". In: *PLoS ONE* 7.6, e37744.

Wärneryd, Karl (1993). "Cheap Talk, Coordination, and Evolutionary Stability". In: *Games and Economic Behavior* 5.4, pp. 532–546.

Zollman, Kevin J. S. and Rory Smead (2010). "Plasticity and Language: An Example of the Baldwin Effect?" In: *Philosophical Studies* 147.1, pp. 7–21.