Language Change and the Force of Innovation

Roland Mühlenbernd[1] and Jonas David Nick[2]

[1] Eberhard Karls Universität Tübingen, Germany
[2] Eidgenössische Technische Hochschule Zürich, Switzerland

Abstract. Lewis [L1] invented *signaling games* to show that semantic meaning conventions can arise simply from regularities in communicative behavior. The behavioral implementation of such conventions are so-called *signaling systems*. Previous research addressed the emergence of signaling systems by combining signaling games with learning dynamics, and not uncommonly researchers examined the circumstances preventing the emergence of signaling systems. It has been shown that by increasing the number of states, messages and actions for a signaling game, the emergence of signaling becomes increasingly improbable. This paper contributes to the question of how the invention of new messages and extinction of unused messages would change these outcomes. Our results reveal that this innovation mechanism does in fact support the emergence of signaling systems. Furthermore, we analyze circumstances that lead to stable communication structure in large spatial population structures of interacting players.

1 Introduction

Signaling games are a leading model to analyze the evolution of semantic meaning. Researchers in this field use simulations to explore agents' behavior in repeated signaling games. Within this field of study two different research approaches are apparent: first, the simulation of a repeated 2-players signaling game combined with agent-based learning dynamics, in the majority of cases with *reinforcement learning* (e.g. [B1], [BZ1], [S1]); second, evolutionary models of population dynamics, wherein signaling games are usually combined with population-based *replicator dynamics* (e.g. [HH1], [HSRZ1]). To fill the gap between both methods, recent work deals with applying repeated signaling games combined with agent-based dynamics on multi-agent populations, e.g. on social network structures (c.f. [Z1], [W1], [M1], [MF1]). With this paper we want to make a contribution to this line of research.

Barrett [B1] was able to show that the simplest variant of a signaling game, called *Lewis game*, combined with a basic version of the learning dynamics *reinforcement learning*, with 2-players which play the game repeatedly, conventions about meaningful language always emerge. But by extending the domains[1] of

[1] With domains we refer to the number of states, messages and action of a signaling game. It will be introduced in the following section.

M. Colinet et al. (Eds.): ESSLLI 2012/2013, LNCS 8607, pp. 194–213, 2014.

the signaling game, those conventions become more and more improbable. Furthermore, the number of possible different perfect signaling systems increases dramatically. This might be the reason why previous research work basically dealt with very simple variants of signaling games, especially in multi-agent setups, and avoided domain-extended games. If even two players fail to learn a signaling system for a given game, multiple players would not only have this problem, but could ultimately end up in a confusion of tongues, where a lot of different incompatible signaling systems evolve.

With this article we will show that by extending the learning dynamics to allow for innovation we can observe i) an improvement of the probability that signaling systems emerge for domain-extended signaling games and extended population sizes, ii) the emergence of different evolving perfect signaling systems in a spatial population structure with local interaction and iii) the formation of regions of the same signaling system that form a spatial continuum.

This article is divided in the following way: in Section 2 we'll introduce some basic notions of repeated signaling games, reinforcement learning dynamics and multi-agent approaches; in Section 3 we'll take a closer look at the variant of reinforcement dynamics we used - a further development of Bush-Mosteller reinforcement; in Section 4 we show how innovation of new and extinction of unused messages significantly improves the outcome in terms of the emergence of signaling systems; in Section 5 we simulate agents on a two-dimensional toroid lattice to show the emergence of a dialect continuum; we'll finish with a conclusion and some implications of our approach in Section 6.

2 Signaling Games and Learning

A signaling game $SG = \langle \{S, R\}, T, M, A, Pr, U \rangle$ is a game played between a sender S and a receiver R. Initially, nature selects a state $t \in T$ with prior probability[2] $\Pr(t) \in \Delta(T)$, which only the sender observes. Therefore the current state remains a secret to the receiver. S then selects a message $m \in M$, and R responds with a choice of action $a \in A$. For each round of play, players receive utilities depending on the actual state t and the response action a. Here we will be concerned with a common variant of this game, where the number of states is on par with the number of actions ($|T| = |A|$). For each state $t_i \in T$ there is exactly one action $a_j \in A$ that leads to successful communication. This is expressed by the utility function

$$U(t_i, a_j) = \begin{cases} \alpha, & \text{if } i = j \\ -\beta, & \text{otherwise} \end{cases}$$

where $\alpha > 0$ and $\beta \geq 0$. In standard signaling games α is 1 and β is 0. This utility function expresses the particular nature of a signaling game, namely that because successful communication does not depend on the used message, there is no predefined meaning of messages. A signaling game with n states and n messages is called an $n \times n$ game and n is called the *domain* of the game.

[2] $\Delta(X) : X \to \mathbb{R}$ denotes a probability distribution over random variable X.

Fig. 1. Two perfect signaling systems of a 2×2 game, consisting of a pure sender and receiver strategy.

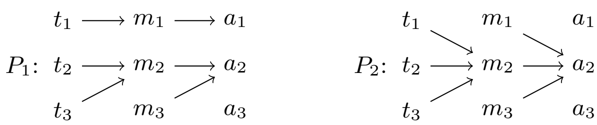


Fig. 2. Two partial pooling systems. P_1 permits an information flow of $^2/_3$, P_2 of $^1/_3$.

2.1 Strategies and Signaling Systems

Although messages are initially meaningless in this game, meaning arises from regularities in behavior. Behavior is defined in terms of strategies. A *behavioral sender strategy* is a function $\sigma : T \rightarrow \Delta(M)$, and a *behavioral receiver strategy* is a function $\rho : M \rightarrow \Delta(A)$. A behavioral strategy can be interpreted as a single agent's probabilistic choice or as a population average. For a 2×2 game exactly two isomorphic strategy profiles constitute a perfect signaling system. In these, strategies are pure (i.e. action choices have probabilities 1 or 0) and messages associate states and actions uniquely, as depicted in Figure 1.

It is easy to see that for an $n \times n$ game the number of perfect signaling systems is $n!$. This means that while for a 2×2 game we get the 2 signaling systems as mentioned above, for a 3×3 game we get 6, for a 4×4 game 24, and for a 8×8 game more than 40,000 perfect signaling systems. Moreover, for $n \times n$ games with $n > 2$ there is a possibility of partial *pooling equilibria*, which transmit information in a fraction of all possible cases. Figure 2 shows different possibilities of partial pooling systems for a 3×3 game.

2.2 Models of Reinforcement Learning

The simplest model of reinforcement learning is *Roth-Erev reinforcement* [RE1] and can be captured by a simple model based on urns, known as *Pólya urns*, which works in the following way: an urn contains balls of different types, each type corresponding to an action choice. Now, drawing a ball means to perform the appropriate action. An action choice can be successful or unsuccessful and in the former case, the number of balls of the appropriate act will be increased by one, such that the probability for this action choice is increased for subsequent draws. All in all, this model ensures that the probability of making a particular decision depends on the number of balls in the urn and therefore on the success of past action choices. This leads to the effect that the more successful an action choice is, the more probable it becomes in following draws.

But Roth-Erev reinforcement has the property that after some time the learning effect[3] slows down: while the number of additional balls for a successful action is a static number α, in the general case $\alpha = 1$, as mentioned above, the overall number of balls in the urn is increasing over time. E.g. if the number of ball in the urn at time τ is n, the number at a later time $\tau + \epsilon$ must be $m \geq n$. Thus the learning effect is changing from α/n to α/m and therefore can only decrease over time.

Bush-Mosteller reinforcement (see [BM1]) is similar to Roth-Erev reinforcement, but without slowing the learning effect down. After a reinforcement step the overall number of balls in an urn is adjusted to a fixed value c, while preserving the ratio of the different balls. Thus the number of balls in the urn at time τ is c and the number at a later time $\tau + \epsilon$ is c and consequently the learning effect stays stable over time at α/c.

A simple yet powerful modification is the adoption of *negative reinforcement*: while in the standard procedures unsuccessful actions have no effect on the urn value, with negative reinforcement an unsuccessful action is punished by decreasing the number of balls that lead to that action.

2.3 Reinforcement Learning and Signaling Games

To apply reinforcement learning to signaling games, sender and receiver both have urns for different states and messages and make their decision by drawing a ball from the appropriate urn. In detail: the sender has an urn $\mho_t$ for each state $t \in T$, which contains balls for different messages $m \in M$. Let $m(\mho_t)$ denote the number of balls of type m in urn $\mho_t$ and $|\mho_t|$ denote the overall number of balls in urn $\mho_t$. If the sender is faced with a state t she draws a ball from urn $\mho_t$ and sends message m, if the ball is of type m. Accordingly, the receiver owns urn $\mho_m$ for each message $m \in M$, which contains balls for different actions $a \in A$. The number of balls of type a in urn $\mho_m$ is denoted as $a(\mho_m)$, the overall number of balls in urn $\mho_m$ as $|\mho_m|$. Upon perceiving message m the receiver draws a ball from urn $\mho_m$ and plays the action a, if the ball is of type a. Thus the sender's behavioral strategy σ and receiver's behavioral strategy ρ can be defined in the following way:

$$\sigma(m|t) = \frac{m(\mho_t)}{|\mho_t|} \quad (1) \qquad \rho(a|m) = \frac{a(\mho_m)}{|\mho_m|} \quad (2)$$

Recently, Franke and Jäger [FJ1] introduced the concept of *lateral inhibition* for reinforcement learning in signaling games in order to lead the system more speedily towards pure strategies. In the next section we will show that lateral inhibition also generally increases the probability that repeated signaling games lead to the emergence of signaling systems (as e.g. depicted in Figure 3).

The concept of lateral inhibition applied on reinforcement learning can basically describes as follows: drawing a successful action not only increases the

[3] The learning effect is the ratio of additional balls for a successful action choice to the overall number of balls.

number of corresponding balls, but also decreases the number of each other type of ball. Likewise, an unsuccessful action decreases its probability, while the probability of competing actions increases. E.g. using Roth-Erev reinforcement with lateral inhibition value $\gamma \in \mathbb{N} \geq 0$ the following update process is executed after each round of play: if communication via t, m and a is successful, the number of balls in the sender's urn $\mho_t$ is increased by $U(t, a) = \alpha \in \mathbb{N} > 0$ balls of type m and reduced by γ balls for each type $m' \neq m$. Similarly, the number of balls in the receiver's urn $\mho_m$ is increased by α balls of type a and reduced by γ balls for each type $a' \neq a$. Furthermore, negative reinforcement also changes urn contents in the case of unsuccessful communication in the following way: if communication via t, m and a is unsuccessful, the number of balls in the sender's urn $\mho_t$ is decreased by $U(t, a) = \beta \in \mathbb{N} \geq 0$ balls of type m and increased by γ balls for each type $m' \neq m$; the number of balls in the receiver's urn $\mho_m$ is decreased by β balls of type a and increased by γ balls for each type $a' \neq a$.

Some further remarks: the lateral inhibition value γ ensures that the probability of an action can become zero and therefore speeds up the learning process. Note that the number of balls can never become a negative value, what is ensured by a lower boundary of 0. Finally, note that in the same way lateral inhibition can be applied on Bush-Mosteller reinforcement.

2.4 Multi-agent Accounts

It is interesting not only to examine the classical 2-players sender-receiver game, but the behavior of agents in a society (e.g. [Z1], [W1], [M1], [MF1]), where more than 2 agents interact with each other and switch between sender and receiver role. In this way an agent can learn both a sender and a receiver strategy. If such a combination forms a signaling system, it is called a *signaling language*. Thus, the number of different possible signaling languages is defined by the number of possible signaling systems and for an $n \times n$ game an agent can learn one of $n!$ different signaling languages. Furthermore, if an agent's combination of sender and receiver strategy forms a pooling system, it is called a *pooling language*. It is easy to see that the number of possible pooling languages exceeds the number of possible signaling languages for any kind of $n \times n$ game.

3 Simulating Bush-Mosteller

Barrett [B1] simulated repeated signaling games with Roth-Erev reinforcement in the classical sender-receiver variant and calculated the *run failure rate* (RFR). The RFR is the proportion of runs not ending with communication via a perfect signaling system. Barrett started 10^5 runs for $n \times n$ games with $n \in \{2, 3, 4, 8\}$. His results show that 100% (RFR = 0) of 2×2 games were successful. But for $n \times n$ games with $n > 2$, the RFR increases rapidly (Figure 3, left).

To compare different dynamics, we started two series of simulation runs for Bush-Mosteller reinforcement in the sender-receiver variant with urn content parameter $c = 20$ and reinforcement value $\alpha = 1$. In the second series we additionally used lateral inhibition with value $\gamma = {}^{1}\!/_{|T|}$. We tested the same games

Game	RFR
2×2	0%
3×3	9.6%
4×4	21.9%
8×8	59.4%

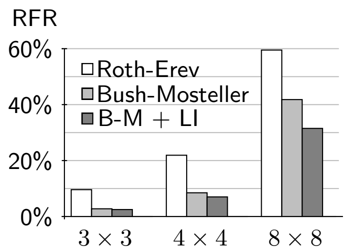


Fig. 3. *Left:* Barrett's results for different $n \times n$ games. *Right:* Comparison of different learning dynamics: Barrett's results of Roth-Erev reinforcement, results for Bush-Mosteller reinforcement without and with lateral inhibition.

as Barrett and correspondingly 10^5 runs per game. In comparison to Barrett's findings, our simulation outcomes i) also resulted in a RFR of 0 for the 2×2 game, but ii) revealed an improvement with Bush-Mosteller reinforcement for the other games, especially in combination with lateral inhibition (see Figure 3, right). Nevertheless, the RFR is never 0 for $n \times n$ games with $n > 2$ and gets worse for increasing n-values, independent of the dynamics.

To analyze the behavior of agents in a multi-agent society, we started experiments with the smallest group of agents in our simulations: three agents arranged in a complete network. In contrast to our first simulations, all agents communicate as both sender and receiver and can learn not only a perfect signaling system, but a signaling language. Furthermore, it was not only recorded if the agents learned a language, but how many agents learned one. With this approach we started between 500 and 1000 simulation runs using Bush-Mosteller reinforcement ($\alpha = 1$, $c = 20$) for $n \times n$ games with $n = 2 \ldots 8$. We stopped a simulation run when each agent in the network learned a signaling or pooling language. We measured the percentage of simulation runs ending with no, one, two or three signaling language learners.

We obtained the following results (Figure 4, left): in 2×2 games, all three agents learned the same signaling language in more than 80% of all simulation runs. But in 3×3 games in less than a third of all runs agents learned a signaling language; in more than 40% of all runs exactly two agents learned a signaling language. And it gets even worse for games with bigger n. E.g. for an 8×8 game in almost 80% of all runs no agents learned a signaling language and it never happened that all agents learned a signaling language.

In addition, we were interested in whether and how the results would change by extending the number of agents. Thus, in another series of experiments we tested the behavior of a complete network of 5 agents in comparison with the results of the 3 agent population. Figure 4 (right) shows the average number of agents who learned a signaling language per run for different $n \times n$ games. As one can see, the percentage of language learners declines rapidly with larger domains and is by and large the same for 3- and 5-agents populations.

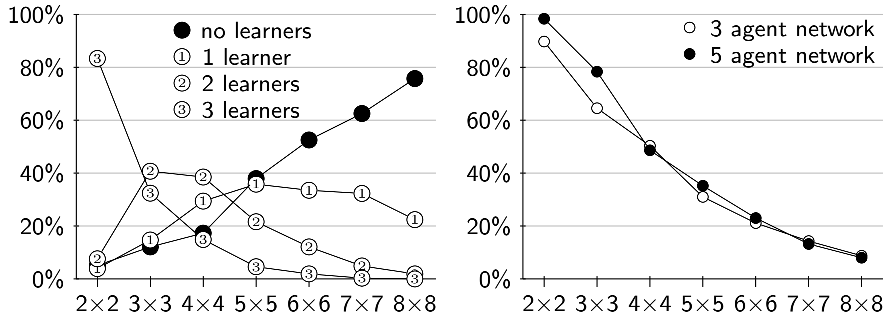


Fig. 4. Left: Percentage of simulation runs ending with a specific number of learners of signaling languages in a network with three agents for different $n \times n$ games with $n = 1 \ldots 8$. Right: Average percentage of agents learning a signaling language over all runs for different $n \times n$ games with $n = 1 \ldots 8$. Comparison of the results of a complete network of 3 agents (white circles) and 5 agents (black circles).

In a nutshell, the results for the classical sender-receiver game reveal that by extending learning dynamics, the probability of the emergence of perfect signaling systems can be improved but it is never one for an $n \times n$ game with $n > 2$. Moreover, the results of the multi-agent network with three agents show that even for the 2×2 game there are cases where not all agents learn a language. And for games with larger domains the results are worse. Furthermore, they don't get better or worse by changing the number of agents, as shown in a multi-agent population with 5 agents. A learning dynamics should be capable of dealing with environments with many states and a lot of interlocutors, because otherwise it does not yield a sufficient explanation for the emergence of many of the signaling systems we find in nature. We show in the next section that by allowing extinction of unused messages and emergence of new messages, perfect signaling systems will emerge with certainty in games with multiple agents and more states.

4 Innovation

The idea of reinforcement learning with innovation is basically as follows: messages can become extinct and new messages can emerge; thus the number of messages during a repeated play can vary, whereas the number of states is fixed. Pioneer work on innovation and extinction for reinforcement learning applied on signaling games stems from Skyrms [S1], further basic experiments with Roth-Erev reinforcement were made by Alexander et al. [ASZ1]. The main contribution of this paper is i) to combine it with Bush-Mosteller reinforcement plus negative reinforcement and ii) to use it for multi-agent accounts.

The process of the emergence of new messages works like this: in addition to the balls for each message type, each sender urn has an amount of *innovative balls* (according to Skyrms we call them *black balls*). If drawing a black ball the sender sends a completely new message. Because the receiver does not have a receiver urn of the new message, he chooses a random action. If action and state matches, the new message is adopted in the *set of known messages* of both interlocutors in the following way: i) both agents get a receiver urn for the new message, wherein the balls for all actions are distributed equiprobably, ii) both agents' sender urns are filled with a predefined amount of balls of the new message and iii) the sender and receiver urn involved in this round are updated according to the learning dynamics. If the newly invented message does not lead to successful communication, the message will be discarded and there will be no change in the agents' strategies.

As mentioned before, messages can go extinct, and that is realized in the following way: because of lateral inhibition, infrequently used or unused messages' value of balls in the sender urns will get lower and lower. At a point when the number of balls of a message is 0 in all sender urns of a particular agent, the message isn't existent in the active use of that agent (i.o.w. she cannot send the message anymore), and will also be removed from the agent's passive use by deleting the appropriate receiver urn. At this point the message isn't in this agent's set of known messages. Some further notes on this model are as follows:

- it is possible that an agent can receive a message that is not in her set of known messages. In this case she adopts the new message like described for the case of innovation. Note that in a multi-agent setup this allows for a spread of new messages
- the black balls are also affected by lateral inhibition. That means that the number of black balls can decrease and increase during runtime; it can especially be zero
- a game with innovation has a dynamic number of messages during a repeated play, but generally ends with $|M| = |T|$. Thus we call an innovation game with n states and n ultimate messages an $n \times n^*$ game

4.1 The Force of Innovation

Since an agent invents a new message if she draws a black ball, the proportion of black balls of an agent's sender urns represents the probability to invent a new message. We call this probability the *force of innovation*, defined as follows:

Definition 1: Given an agent's set of sender urns $\mho = \{\mho_t | t \in T\}$ for a set of states T. An agent's *force of innovation* FOI describes her proportion of black balls over her set of all sender urns:

$$FOI(\mho) = \frac{\sum_{\mho_t \in \mho} \frac{b(\mho_t)}{|\mho_t|}}{|\mho|} \tag{3}$$

where $b(\mho_t)$ is the number of black balls in urn $\mho_t$.

In the following study we investigated the way the force of innovation changes over time in a simulation run. Furthermore we wanted to find out if it correlates

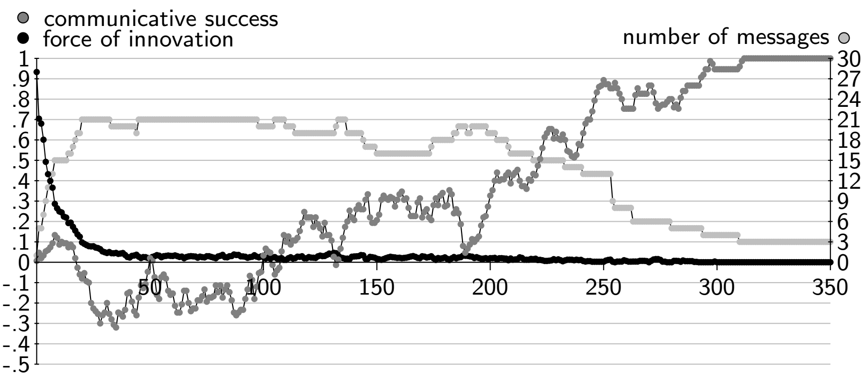


Fig. 5. Simulation run of a $3 \times 3^*$ game with innovation in a 3-agents population: communicative success, force of innovation (averaged over all agents) and the actual number of used messages in the population - alteration over time

with the agents' *communicative success*[4], since we expected a highly negative correlation between it and the force of innovation. We started 100 simulation runs with the following settings:

- network type: complete network with 3 agents
- game type: $3 \times 3^*$ game
- learning dynamics: Bush-Mosteller reinforcement with negative reinforcement, lateral inhibition value ($\alpha = 1$, $\beta = 1$, $\gamma = {}^1\!/_{|T|}$) and innovation
- initial state: every urn of the sender is filled with black balls and the receiver does not have any a priori urn.
- break condition: simulation stops if all agents learned a signaling language

Note that the settings of the learning dynamics implicate that the communicative success value can be between -1 and 1[5], and because of the initial state of the sender urns, the force of innovation of all agents is 1 at the beginning.

The simulation results revealed first of all: all agents learn the same signaling language; and that really quickly: they need maximally 500 simulation steps. Now let's take a closer look at how a $3 \times 3^*$ game played in a 3-agents population develops during a simulation run by analyzing the *communicative success* and the average *force of innovation* of the population, plus the number of messages, used in the population.[6] Figure 5 shows an exemplary course of the resulting values' alteration over time for one of the simulation runs. It shows that in the beginning the agents are very innovative and create a lot of messages, which

[4] The communicative success is measured as the average utility value of all agents' utility value at a given simulation step

[5] Note that the range of an utility value is between $-\beta$ and α.

[6] This is the number of all messages that were i) once invented and ii) of which at least one agent has a non-zero probability to draw at the given simulation step.

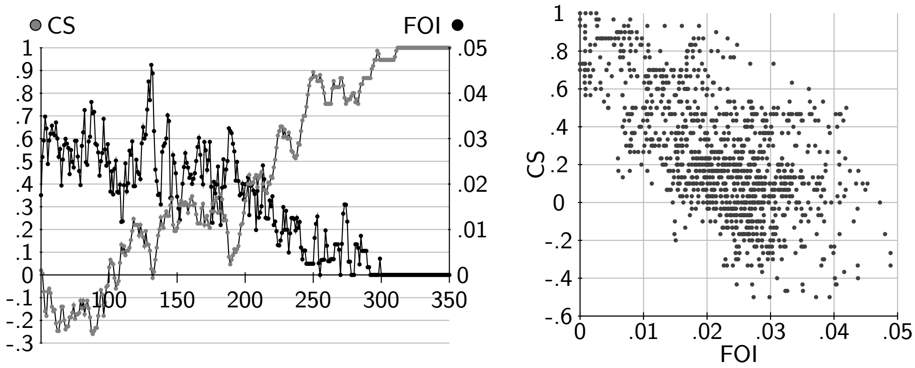


Fig. 6. Left: Simulation run of a $3 \times 3^*$ game with innovation in a 3-agents population, starting with simulation step 50. Comparison of communicative success (CS) and force of innovation (FOI) over time. Right: Data points of 10 simulation runs for FOI values $\leq .05$. CS and FOI reveal a very high negative correlation of $-.6$ for 40,000 data points.

reduces the number of black balls in the urns, because balls for the new messages are added and then the urn content is normalized. Note that for the first communication steps the force of innovation drops rapidly, while the number of messages rises until it reaches 21 messages. Furthermore, the communicative success is below zero at the beginning, since agents use a diversity of different messages and successful communication is less probable than chance. But once there evolved an agreement on which messages might be useful in terms of successful communication, further messages died out, so the number of known messages decreased. Finally, the communicative success reaches a perfect 1 on average, while the number of messages equals the number of states (3) and the force of innovation drops to zero.

By taking a closer look on the data, an interesting interplay between communicative success and force of innovation becomes evident: successful communication lowers the force of innovation, whereas unsuccessful communication raises it. That is not a surprise, since black balls can only change by lateral inhibition: increase in the case of unsuccessful communication and decrease in the case of successful communication. The relationship of both values is better seen in Figure 6 (left) that shows the force of innovation and the communication success of the same simulation runs as already depicted in Figure 5, but this time i) without the initial phase of the first 50 simulation steps and ii) the value of the force of innovation is displayed 20 times more fine-grained. The relationship between both values is clearly recognizable in this figure: one measure's peak is simultaneously the other measure's valley. Admittedly, the mirroring is not perfect, but is clearly reveals a plausible social dynamics: the higher the communicative success, the lover the force of innovation, and vice versa.

To get a more quantitative picture of this relationship, we analyzed the data points' correlation of all 100 simulation runs (about 40,000 data points). It turned out that force of innovation and communicative success reveal a very strong

Table 1. Runtime Table for $n \times n^*$ games with $n = 2 \ldots 8$; for a complete network of 3 agents and 5 agents

Game	$2 \times 2^*$	$3 \times 3^*$	$4 \times 4^*$	$5 \times 5^*$	$6 \times 6^*$	$7 \times 7^*$	$8 \times 8^*$
3 agents	1,052	2,120	4,064	9,640	21,712	136,110	> 500,000
5 agents	2,093	5,080	18,053	192,840	> 500,000	> 500,000	> 500,000

negative correlation: a *Pearson-Correlation* of $-.6$. To get an impression how the data correlate, Figure 5 (right) depicts the data points of ten simulation runs for FOI-values $\leq .05$.[7]

4.2 Learning Languages by Innovation: A Question of Time

In Section 3 we were able to show that the percentage of agents learning a signaling language in a multi-agent context decreases by increasing the domain size of the game. To find out whether innovation can improve these results we started simulation runs for games with different domains. We used the following settings:

- network types: complete network with 3 agents and with 5 agents
- learning dynamics: Bush-Mosteller reinforcement with negative reinforcement and lateral inhibition value ($\alpha = 1$, $\beta = 1$, $\gamma = 1/|T|$) and innovation
- initial state: every urn of the sender is filled with black balls and the receiver does not have any a priori urn.
- experiments: 100 simulation runs per $n \times n^*$ game with $n = 2 \ldots 8$
- break condition: simulation stops if the communicative success of every agent exceeds 99% or the runtime passes the runtime limit of 500,000 communication steps (= runtime)

These simulation runs gave the following results: i) for the 3-agents account in combination with $n \times n^*$ games for $n = 2 \ldots 7$ and the 5-agents in combination with $n \times n^*$ games for $n = 2 \ldots 5$ all agents learned a signaling language in each simulation run and ii) for the remaining account-game combinations all simulation runs exceeded the runtime limit (see Table 1). We expect that for the remaining combination all agents will learn a signaling language as well, but it takes extremely long.

All in all, we were able to show that the integration of innovation and extinction of messages leads to a final situation where all agents learned the same signaling language, if the runtime does not exceed the limit. Nevertheless, we expect the same result for account-game combinations where simulations steps of these runs exceeded our limit for a manageable runtime.

[7] The reason to illustrate only data points with a FOI value $\leq .05$ was to get a better depiction of the data. Note that more that 99% of all data points have a FOI value $\leq .05$ and therefore are depicted here.

4.3 Games with a Limited Message Set

As our previous experiments have shown, increasing the number of agents of the population and/or states of the game has a disastrous impact on the runtime. Especially the dependency of the runtime on the number of agents makes the game inapplicable to experiments with larger populations and network structures. The problem of the current account is as follows: whenever communication does not work well, agents' force of innovation increases and they invent new messages. And the more agents are interacting with each other, the more new messages might arise. Thus, the probability of all agents agreeing on a specific set of messages is virtually zero for a larger population. Of course, the probability is close to zero but non-zero and therefore you just have to wait long enough for an population-wide agreement to happen. But the larger the population, the closer is the probability to zero and the longer is the expected runtime.

A reasonable compromise that allows for innovation while keeping the computational complexity feasible is to limit the maximum number of messages.[8] Thus, we introduce a new signaling game: an $n \times n^m$ game has n states and actions and *maximally* m different messages. In such a game, agents that draw a black ball choose randomly a message from the limited message set, without the restriction that this message must be completely new to the population.

The new game has an intuitive analogy to actual signaling beings. In principle, nature might allow an infinite message set, but living beings are only capable of distinguishing a finite number of messages due to sensory, cognitive or motor imperfection. In this sense, each message represents a particular category of non-distinguishable messages.

By adopting this new feature in our game we made experiments to check the runtime improvement for larger population sizes. In particular, we analyzed a 3×3^m game with a set of 30 messages (3×3^{30} game) in comparison with a $3 \times 3^*$ game by using the following settings:

- network types: complete network with different sizes from 2 up to 9 agents
- learning dynamics: Bush-Mosteller reinforcement with negative reinforcement and lateral inhibition value ($\alpha = 1$, $\beta = 1$, $\gamma = 1/|T|$) and innovation
- initial state: every urn of the sender is filled with black balls and the receiver does not have any a priori urn.
- experiments: 100 simulation runs per network size and for a $3 \times 3^*$ game and 3×3^{30} game as well
- break condition: simulation stops if the communicative success of every agents exceeds 99% or the runtime passes the runtime limit of 500,000 communication steps (= runtime)

The result is depicted in Figure 7: the comparison of runtime behavior of the game with an unlimited message set and a limited message set of 30 messages. As already seen in the experiments of Section 4.2, for a $3 \times 3^*$ game the runtime is

[8] Note that since a new invented message extends the history of all messages ever used by one, the set of possible messages is virtually unlimited.

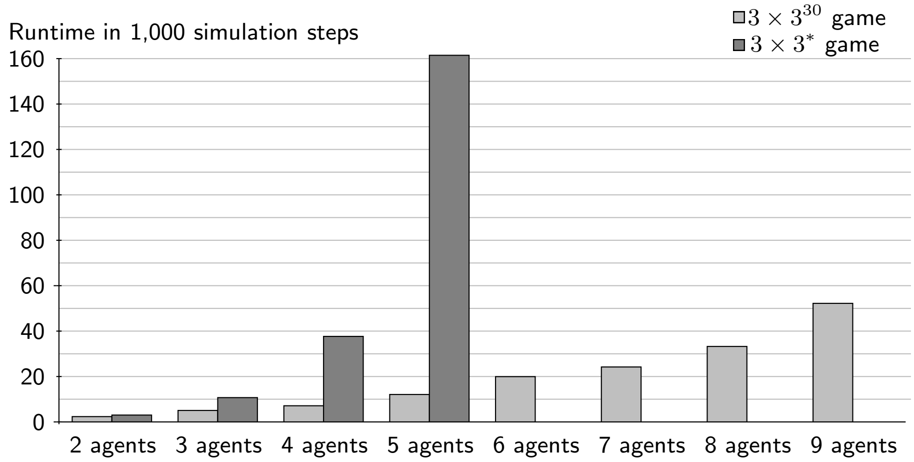


Fig. 7. Runtime comparison of games with limited (3×3^{30}) and unlimited ($3 \times 3^*$) message sets for different population sizes

only manageable for a population up to 5 agents and increases with the number of agents in a strong slope, whereas for a 3×3^{30} game the runtime increases slowly and is for 9 agents still manageable. All in all, the new feature improves the runtime behavior quite well by keeping the innovational nature of the game. This makes it also applicable for larger network structures, as we will show in the next section.

5 Spatial Dynamics

So far we have dealt with fully connected networks of a few agents. However, when modeling natural scenarios encompassing participants of whole populations, it is rather unreasonable that i) the number of population members is that small and ii) all members are connected to each other. To target a more realistic framework, we arranged experiments on a large population with local interaction structure: a *toroid lattice* of 100×100 agents; here each agent can only communicate with her eight direct neighbors (Moore neighborhood).

There are a number of previous studies that addressed signaling games on spatial structures: one of the first studies analyzed a simple 2×2 signaling game on a toroid lattice structure, whereby agents use *imitation* to guide their decisions [Z1]. Some consecutive studies take this analysis up by either changing the dynamics to reinforcement learning [M1] or by changing the interaction structure to small-world networks and by extending the game domains to a 3×3 signaling game [W1]. Another study entails experiments on social network structures as interaction structure plus incorporating reinforcement learning as update dynamics [MF1]. All these studies analyzed a simple variant: a 2×2 or 3×3 game.

The basic result of all these studies was the emergence of regional meaning: the lattice or network structure was split into local language regions.[9]

Note that in all these studies the number of possible signaling systems is quite small. The 2×2 game has only two signaling systems (as depicted in Figure 1) and the 3×3 game has 6 signaling systems. In the upcoming experiments we applied 3×3^{30} games and we also expect regional meaning to emerge. But as opposed to the before-mentioned studies, here not 2 or 6, but 6840 different signaling systems are possible! These prerequisites bring a number of questions about: do stable language regions emerge? And if so, how many different language regions emerge? And how are these regions arranged? Do they depict a specific pattern in terms of arrangements with other language regions? Are they stable? This section addresses these questions.

5.1 Spatial Structure: Dialect Regions

To find answers to the before mentioned questions, we started experiments with the following settings:

- network type: 10,000 agents placed on a 100×100 toroid lattice
- game type: 3×3^{30} game
- learning dynamics: Bush-Mosteller reinforcement with negative reinforcement and lateral inhibition value ($\alpha = 1$, $\beta = 1$, $\gamma = 1/|T|$) and innovation
- break condition: simulation stops after 50,000 simulation steps or every agent has learned a signaling language

Like in the previous experiments, we measured the average communicative success and the force of innovation of the whole population over time. Furthermore, since we were interested in the number of signaling languages that might emerge, we also measured the population-wide number of signaling languages over time. The resulting course of a simulation run for the first 10,000 simulation steps is depicted in Figure 8.

Like for the previous experiments of Section 4.1 (c.f. Figure 5) for a small population of 3 agents, the force of innovation decreases really fast down to (almost) zero, while the communicative success first decreases to a negative value and then increases again. Thus, the initial phase is quite similar. But while for the experiments of Section 4.1, the 3 agent-population quickly agrees on one signaling language and the communicative success reaches a perfect value of 1, here the population of 10,000 agents 'agrees' on more than 600 signaling languages and the communicative success reaches an average value of almost .8 after around 500 simulation steps. From this point on the number of signaling languages slowly decreases, while the value of communicative success slowly increases. Note that Figure 8 only shows the first 10,000 simulation steps. The whole simulation run showed that after 50,000 simulation steps the number of signaling languages has

[9] A language region is defined as a connected sub-network, of which each member has learned the same signaling language L, but each other agent connected to this regions hasn't learned L.

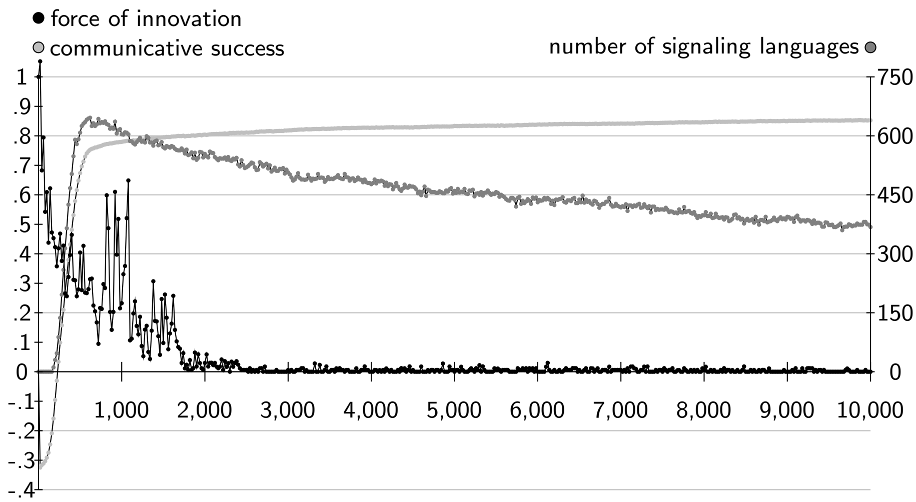


Fig. 8. Simulation run of a 3×3^{30} game in a population of 10,000 agents placed on a 100×100 toroid lattice: average communicative success, force of innovation and the number of society-wide signaling languages over the first 10,000 simulation steps.

decreased to around 170, while the communicative success reached a value of more than .9.

Communication cannot be perfectly successful because the population does not learn one unique, but multiple signaling languages. But how are these different signaling languages spatially arranged? If they would be arbitrarily spread over the whole lattice, we would expect a much lower communicative success value, basically lower than zero. But since the value is around .8, communication works quite well even with such a huge number of different signaling languages. The reason becomes visible if we take a look at the spatial arrangement of the different signaling languages. It turns out that they form what we call *language regions*. A language region is a connected subgraph for which each agent uses the same signaling language. Figure 9 shows the resulting pattern on a 100x100 toroid lattice, the left figure for the pattern after 2,000 simulation steps (more than 500 different language regions), the right figure for the pattern after 50,000 simulation steps (around 170 language regions).

Note that Figure 8 shows that the number of signaling languages slowly decreases over time. Furthermore, Figure 9 shows that each signaling language that evolved forms at least one language regions. Consequently, the number of language regions decreases over time: after 50,000 simulation steps there is only a third of the number of language regions than after 2,000 simulation steps. We haven't analyzed the concrete dynamics that lead to this decline of language regions, but we expect mechanisms like unification, melting, displacement and extinction at the borders of neighboring language regions. The exact dynamics behind this process remains to be analyzed in subsequent studies.

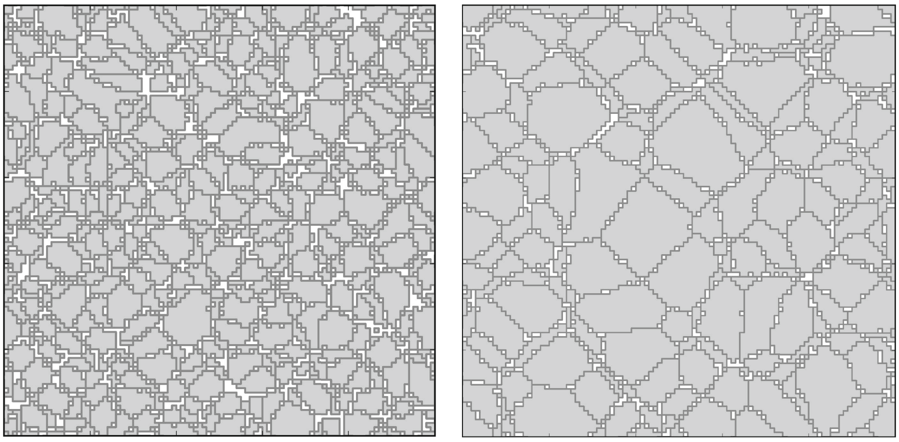


Fig. 9. The allocation of language regions on a 100x100 toroid lattice. A gray cell represents an agent that learned a signaling language. The borders between language regions are marked by darkgray lines. While after 2,000 simulation steps the map is segmented in over 500 language regions (left figure), after 50,000 simulation steps it is only one third of it, around 170 language regions (right figure).

In the upcoming section we present results of the analyses of the spatial relationship between language regions at one point in time to examine if their placement is randomly or follows particular patterns.

5.2 Spatial Relationships

In this section we want to analyze how the different language regions actually relate to each other. We hypothesize that there is an interaction between spatial distance of two language regions and the similarity of their signaling languages $L = \langle s, r \rangle$.[10] For that purpose we define two similarity measures, *lexical similarity* and *mutual intelligibility*, as follows:

Definition 2: *Lexical Similarity* describes the proportionally common items of lexical entries.[11] Thus between two given signaling languages $L_1 = \langle s_1, r_1 \rangle$ and $L_2 = \langle s_2, r_2 \rangle$ the lexical similarity is defined as follows:

$$LS(L_1, L_2) = \frac{|\{m \in M | \exists t \in T : m = s_1(t)\} \cap \{m \in M | \exists t \in T : m = s_2(t)\}|}{|T|} \quad (4)$$

[10] Note that a signaling language is defined as a strategy pair of a *pure* sender and receiver strategy, defined as $s : T \to M$ and $r : M \to A$, respectively. Note: while agents play according to behavioral strategies, once they have learned a signaling language, their behavioral strategy profile represents a pair of pure strategies.

[11] In the case of signaling languages, lexical entries are entailed messages.

Definition 3: *Mutual Intelligibility* describes the expected communicative success for two given signaling languages $L_1 = \langle s_1, r_1 \rangle$ and $L_2 = \langle s_2, r_2 \rangle$ and is defined as follows:

$$MI(L_1, L_2) = \frac{\sum_t (U^x(t, s_1, r_2) + U^x(t, s_2, r_1))}{2 \times |T|} \tag{5}$$

where $U^x(t, s, r)$ is the expected utility[12] for a given state t, a pure sender strategy s and a pure receiver strategy r.

Note that lexical similarity just describes the number of common messages of two signaling languages. In turn, mutual intelligibility also takes the semantics of messages into account: if messages describe the same state/action, mutual intelligibility is higher. But if two signaling languages have common messages for different states/actions, it gives advantage to lexical similarity, but disadvantages mutual intelligibility, since it supports miscommunication.[13]

To give an example of these similarity measures, lets take a look at the lattice distribution after 50,000 simulation steps as depicted in the left picture of Figure 10. There are three language regions that are marked by its signaling languages L_{55}, L_{72} and L_{139}. The concrete signaling languages are depicted in the right picture. It turns out that the close language regions 55 and 72 have a quite high lexical similarity value (.67) and an even higher mutual intelligibility value (.78). The distant language regions 139 and 55 have a low lexical similarity value (.33) and also a low mutual intelligibility value (.22). Similarly, language region 139 and 72 have no lexical similarity and a low mutual intelligibility value (.33).

To compare these similarity measures to spatial distances of language regions in a more systematic way, we introduce the measure *regional distance*, a value that describes the distance between two language regions. In detail, it describes the average distance of all members of one language region to all members of the other language region. It is defined as follows:

Definition 4: *Regional Distance* describes the distance between two connected subgraphs of a connected graph as the average distance over all members $n \in N_1$ of subgraph G_1 and $n \in N_2$ of subgraph G_2, defined as follows:

$$RD(G_1, G_2) = \frac{\sum_{n_i \in N_1} \sum_{n_j \in N_2} SP(n_i, n_j)}{|N_1| \times |N_2|} \tag{6}$$

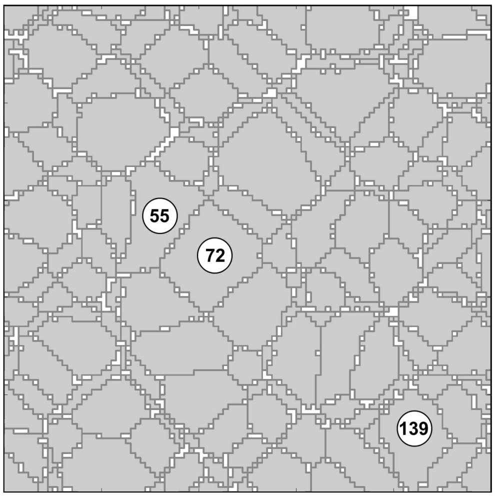
where $SP(n_i, n_j)$ is the *shortest path length*[14] between node n_i and node n_j.

[12] Expected utility $U^x(t, s, r)$ is defined as follows:

$$U^x(t, s, r) = \begin{cases} U(t, r(s(t))) & \text{if } s(t) \in \{m \in M | \exists a \in A : m = r^{-1}(a)\} \\ \frac{\alpha}{|A|} + \frac{\beta \times (|A|-1)}{|A|} & \text{else} \end{cases}$$

[13] It can be shown that two signaling languages can have a high value of lexical similarity as well as a low value of mutual intelligibility, and vice versa.

[14] The shortest path length of two nodes describes the length of a path (= number of edges) between them that has a minimal number of edges.



$$L_{55}: \begin{array}{l} t_1 \longrightarrow m_8 \longrightarrow a_1 \\ t_2 \longrightarrow m_{21} \longrightarrow a_2 \\ t_3 \longrightarrow m_{22} \longrightarrow a_3 \end{array}$$

$$L_{72}: \begin{array}{l} t_1 \longrightarrow m_{23} \longrightarrow a_1 \\ t_2 \longrightarrow m_{21} \longrightarrow a_2 \\ t_3 \longrightarrow m_{22} \longrightarrow a_3 \end{array}$$

$$L_{139}: \begin{array}{l} t_1 \longrightarrow m_7 \longrightarrow a_1 \\ t_2 \longrightarrow m_8 \longrightarrow a_2 \\ t_3 \longrightarrow m_{15} \longrightarrow a_3 \end{array}$$

Fig. 10. Left: The language regions 55, 72 are next to each other, while 139 is far off. Right: strategy profiles of signaling languages L_{55}, L_{72} and L_{139}.

Following this approach we analyzed lexical similarity and mutual intelligibility depending on the distance of each pair of two language regions. The result is depicted in Figure 11: while the number of language regions increases with its distance to a specific language region, the similarity values decrease. Furthermore, both similarity measures follow a curve with falling slope to an expected random value. This result reveals that i) distant language regions seem to have

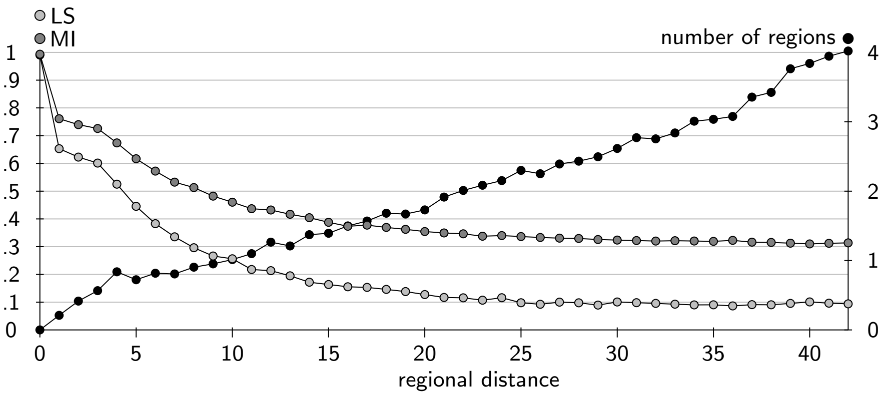


Fig. 11. The grey dots depict the average values of lexical similarity (LS) and mutual intelligibility (MI) between two language regions in dependence of the distance between them. The black dots depict a language region's average number of other language regions in a particular distance to it.

no influence to each other's communication system, since their signaling languages are as similar as randomly chosen ones and ii) spatially close language regions have high similarity values and must strongly influence each other. Both similarity values decrease with the distance.

6 Conclusion and Outlook

In the last few decades, a large body of research has been done to model and analyze the way that stable communication systems emerge among individuals, whereas a popular account in this field is to use repeated signaling games as a model to analyze the circumstances that lead to the emergence to stable communication strategies, so-called signaling systems [BZ1] [FJ1] [HSRZ1]. One premise of most of the work is in accordance with *Occam's razor*: take the simplest model that can explain the phenomenon. The first research results were very promising, since they showed that signaling systems evolve with a very simple learning account: reinforcement learning [HZ1] [S1].

But further studies showed that this result holds basically for simple 2×2 games between two players, but not for more complex games [B1] or larger populations [M1]. Thus we proposed the question: by taking the model of a repeated signaling game in combination with reinforcement learning as starting point, what reasonable additional assumptions are necessary for the emergence of efficient communication (in terms of signaling systems) in complex signaling games played in large populations?

In a first step we extended the learning dynamics with the concept of *innovation* [ASZ1] [S1]. The basic plot is as follows: agents have the ability to occasionally invent new messages. Furthermore, unused messages get automatically lost. We found that these additional concepts enable perfect communication in more complex games and larger populations. But the major drawback of allowing innovation is the exponential computational complexity. It can be shown that extending population size and/or domains of the game even in a moderate magnitude has a tremendous effect on the probability of agents to find a consensus on a common signaling language, what strongly affects the runtime.

By limiting the game's innovation capacity to a limited set of possible messages, we created a new game that keeps the innovative character of the previous game, but solved the problem of computational complexity. With this new account, we varied the network structure, such that agents were arranged on a two-dimensional toroid lattice. This lead to the emergence of language regions, arranged in a particular pattern: the closer the language regions, the more similar their signaling languages.

All in all, we were able to show that, by starting with an account of repeated signaling games and reinforcement learning, one simple extensions is sufficient to realize the emergence of signaling systems for complex signaling games and large populations: innovation of new messages within a limited message set. Furthermore, our experiments showed that a local communication structure leads to local language regions arranged in a continuous way.

Further research should go in multiple directions. First of all, it might be worth to take a closer look at the way language regions change over time, especially at the border regions. Additionally, what remains to be shown is that our results in fact hold for higher numbers of domains of the game. Is our result general, or only true for specific values? It would also be interesting to see what kind of influence more realistic network-types (c.f. small-world networks) would have on the outcome. These are but a few extensions, as a multitude of further experiments addressing factors that might influence the way language regions emerge and interact readily suggest themselves.

References

[ASZ1] Alexander, J.M., Skyrms, B., Zabell, S.L.: Inventing new signals. Dynamic Games and Applications 2(1), 129–145 (2012)

[B1] Barret, J.A.: The Evolution of Coding in Signaling Games. Theory and Decision 67, 223–237 (2009)

[BZ1] Barret, J.A., Zollman, K.J.S.: The Role of Forgetting in the Evolution and Learning of Language. Journal of Experimental and Theoretical Artificial Intelligence 21(4), 293–309 (2009)

[B2] Bloomfield, L.: Language. George Allen & Unwin, London (1935)

[BM1] Bush, R., Mosteller, F.: Stochastic Models of Learning. John Wiley & Sons, New York (1955)

[FJ1] Franke, M., Jäger, G.: Bidirectional optimization from reasoning and learning in games. Journal of Logic, Language and Information 21(1), 117–139 (2012)

[HH1] Hofbauer, J., Huttegger, S.M.: Feasibility of communication in binary signaling games. Journal of Theoretical Biology 254(4), 843–849 (2008)

[HSRZ1] Huttegger, S.M., Skyrms, B., Smead, R., Zollman, K.J.: Evolutionary dynamics of Lewis signaling games: signaling systems vs. partial pooling. Synthese 172(1), 177–191 (2010)

[HZ1] Huttegger, S.M., Zollman, K.J.S.: Signaling Games: Dynamics of Evolution and Learning. In: Benz, A., Ebert, C., Jäger, G., van Rooij, R. (eds.) Language, Games, and Evolution. LNCS (LNAI), vol. 6207, pp. 160–176. Springer, Heidelberg (2011)

[L1] Lewis, D.: Convention. Harvard University Press, Cambridge (1969)

[M1] Mühlenbernd, R.: Learning with Neighbours. Synthese 183(S1), 87–109 (2011)

[MF1] Mühlenbernd, R., Franke, M.: Signaling Conventions: Who Learns What Where and When in a Social Network. In: Proceedings of EvoLang IX, pp. 242–249 (2011)

[RE1] Roth, A., Erev, I.: Learning in extensive-form games: experimental data and simple dynamic models in the intermediate term. Games and Economic Behaviour 8, 164–212 (1995)

[S1] Skyrms, B.: Signals: Evolution, Learning & Information. Oxford University Press, Oxford (2010)

[W1] Wagner, E.: Communication and Structured Correlation. Erkenntnis 71(3), 377–393 (2009)

[Z1] Zollman, K.J.S.: Talking to neighbors: The evolution of regional meaning. Philosophy of Science 72(1), 69–85 (2005)