Where Innovation Meets Evolution
Signaling Game Dynamics and the Vowel Space

Roland Mühlenbernd & Jonas David Nick

Abstract

Jäger (2008) used repeated signaling games with a metric message space applied on multi-agent populations to simulate the emergence of categorization systems of the human vowel space. His implementation can be seen as a form of exemplar dynamics. Inspired by this idea we also used repeated signaling games with a metric message space to simulate exactly the same phenomenon, but our implementation can rather be seen as a form of agreement dynamics, as first introduced through the *Naming game* by Steels (1998). As opposed to Jäger's results our experiments show that the segments of the vowel space that constitute phoneme categories cannot emerge simultaneously, but incrementally by reorganizing the segmentation structure. Furthermore it reveals that in some cases only particular n-vowel system can be predecessors of extended $n+1$-vowel systems.

1 Introduction

With the objective to explore the evolution of language or at least signaling conventions, especially regarding semantic meaning, signaling games (Lewis, 1969) recently became a leading model for this purpose. In line with this trend researchers used simulations to explore the behavior of agents playing repeated signaling games. Within this field of study two different lines of research are apparent: i) the simulation of a repeated 2-players signaling game combined with agent-based learning dynamics, in the majority of cases with the dynamics *reinforcement learning* (c.f. Barrett, 2009; Barrett and Zollman, 2009; Skyrms, 2010) and ii) evolutionary models by simulating population behavior, wherein signaling games are usually combined with the population-based *replicator dynamics* (c.f. Hofbauer and Huttegger, 2008; Huttegger, 2010). To fill the gap between both accounts, recent work deals with applying repeated signaling games combined with agent-based dynamics on social network structures or at least multi-agents accounts (c.f. Zollman, 2005; Wagner, 2009; Mühlenbernd, 2011; Mühlenbernd and Franke, 2011).

With this paper we want to make a contribution to this line of research by extending the account of signaling games and reinforcement learning with the possibility of innovation, i.e. agents can invent new messages and furthermore unused messages can get extinct. This new account constitutes a form of agreement dynamics (c.f. Steels, 1998; Barrat et al., 2007). We'll show that with the property of innovation the emergence of society-wide signaling systems is guaranteed even if the game has a large number of states and messages, i.e. all agents agree on a specific convention for communication. But our simulation results revealed that by extending this account with a metric message space to simulate the emergence of categorization systems of the human vowel space, agents

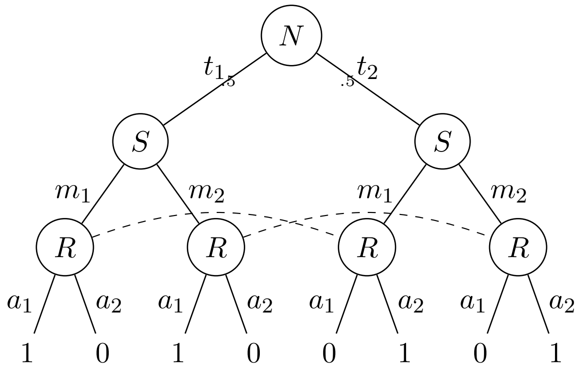


Figure 1: Extensive form game for the 2×2- game. Dashed lines denote states the receiver cannot distinguish.

won't find an agreement, if the number of states is more than 3 from the beginning. But we can show i) that by incrementally increasing the number of states, agents will find an agreement on a specific vowel system and ii) how different vowel systems are probably originally related.

This article is divided in the following way: in Section 2 we'll introduce some basic notions of repeated signaling games. In Section 3 we'll introduce different types of reinforcement learning dynamics, especially for multi-agent accounts; furthermore we'll show some simulation results for a variant of reinforcement dynamics and illustrate the obstacles which prohibit a population of agents to find an agreement; in Section 4 we'll show that by extending the mentioned variant with the possibility of innovation of new and extinction of unused messages, this modification significantly improves the probability of the emergence of society wide agreement on one signaling system; in Section 5 we'll show that by considering a metric message space and distance-dependent update dynamics to simulate the emergence of human vowel systems, a society-wide agreement on compatible signaling systems can only emerge, if the signaling game's number of states is low and will be increased incrementally. We'll also got further insights of the relationships of different vowel systems. We'll finish with some implications of our approach in Section 6.

2 Signaling Games

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 $\Pr(t) \in \Delta(T)$[1], which the sender observes, but the receiver doesn't. 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 state t and action a. We will only be concerned with a variant, where the number of states is on par with the number of actions ($|T| = |A|$) and for each state $t \in T$ there is exactly one action $a \in A$ that leads to successful communication. This is expressed by the utility function $U(t_i, a_j) = 1$ if $i = j$ and 0 otherwise. Figure 1 depicts the *extensive form game* for the simplest game ($|T| = |M| = 2$): a tree depicting the sequences of possible move combinations with the resulting utility value at the leaves.

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

$$L_1: \begin{tikzcd} t_1 \arrow[r] & m_1 \arrow[r] & a_1 \\ t_2 \arrow[r] & m_2 \arrow[r] & a_2 \end{tikzcd} \qquad L_2: \begin{tikzcd} t_1 \arrow[dr] & & m_1 \arrow[dr] & \\ & & & \\ t_2 \arrow[ur] & & m_2 \arrow[ur] & \end{tikzcd}$$

$$L_2: \begin{tikzcd} t_1 \arrow[r] & m_2 & m_1 \arrow[r] & a_2 \\ t_2 \arrow[r] & m_1 & m_2 \arrow[r] & a_1 \end{tikzcd}$$

Figure 2: The two perfect signaling systems of a 2×2-game.

$$P_1: \begin{tikzcd} t_1 \arrow[r] & m_1 \arrow[r] & a_1 \\ t_2 \arrow[r] & m_2 \arrow[r] & a_2 \\ t_3 \arrow[ur] & m_3 \arrow[ur] & a_3 \end{tikzcd} \qquad P_2: \begin{tikzcd} t_1 \arrow[dr] & m_1 \arrow[dr] & a_1 \\ t_2 \arrow[r] & m_2 \arrow[r] & a_2 \\ t_3 \arrow[ur] & m_3 \arrow[ur] & a_3 \end{tikzcd}$$

Figure 3: Two partial pooling systems. P_1 has an information flow of $2/3$, P_2 of $1/3$.

Note that the utility function expresses the particular nature of a signaling game, namely that because successful communication doesn't 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, whereby n is called the *domain* of the game.

2.1 Strategies and Signaling Systems

Although messages are initially meaningless in this game, meaningfulness arises from regularities in behavior. Behavior is defined in terms of strategies. A *behavioral sender strategy* is a function $\sigma : T \to \Delta(M)$, and a *behavioral receiver strategy* is a function $\rho : M \to \Delta(A)$. 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 2.

It is easy to show 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 3 shows different possibilities of partial pooling systems for a 3×3-game.

3 Reinforcement Learning

The simplest model of reinforcement learning is *Roth-Erev reinforcement* (c.f. Roth and Erev, 1995) 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 to be elected in following draws.

But Roth-Erev reinforcement has the property that after a while the learning effect[2] 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 (c.f. Bush and Mosteller, 1955) 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 further modification is the adaption of *negative reinforcement*: while in the standard account unsuccessful actions have no effect on the urn value, with negative reinforcement unsuccessful communication is punished by decreasing the number of balls leading to an unsuccessful action. The effect of reinforcement can also be improved by the concept of *lateral inhibition*. In particular, a successful action will not only increase its probability, but also decrease the probability of competing actions. In our account lateral inhibition applies to negative reinforcement as well: for an unsuccessful action the number of the appropriate balls will be decreased, while the number of each other type of ball will be increased.

3.1 Applying Reinforcement Learning on Repeated Signaling Games

To apply reinforcement learning on 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. We assume that the states are equally distributed. The sender has an urn $\mho_t$ for each state $t \in T$, which contains balls for different messages $m \in M$. The number of balls of type m in urn $\mho_t$ is designated as $m(\mho_t)$, the overall number of balls in urn $\mho_t$ as $|\mho_t|$. 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 has an urn $\mho_m$ for each message $m \in M$, which contains balls for different actions $a \in A$, whereby the number of balls of type a in urn $\mho_m$ is designated as $a(\mho_m)$, the overall number of balls in urn $\mho_m$ as $|\mho_m|$. For a received 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 the receiver's behavioral strategy ρ can be defined in the following way:

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

The learning dynamics are realized by changing the urn content dependent on the *communicative success*. For a Roth-Erev reinforcement account with a positive update value $\alpha \in \mathbb{N} > 0$ and a lateral inhibition value $\gamma \in \mathbb{N} \geq 0$ the following update process

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

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 α balls of type m and reduced by γ balls of 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 of type $a' \neq a$.

Furthermore for an account with negative reinforcement urn contents also change in the case of unsuccessful communication for the negative update value $\beta \in \mathbb{N} \geq 0$ 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 β balls of type m and and increased by γ balls of 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 of type $a' \neq a$. The lateral inhibition value γ ensures that the probability of an action can get zero and it speeds up the learning process.

For Bush-Mosteller reinforcement the content of the appropriate sender and receiver urns will be adjusted to a predefined value in the following way: for the given value c of fixed urn content it is assumed that before a round of play the urn content of all sender and receiver urns $|\mho| = c$. After a round of play it may be the case that the urn content $|\mho| = d \neq c$. Now the number n_i of each type of ball i is multiplied by c/d.[3]

3.2 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 (c.f. Zollman, 2005; Wagner, 2009; Mühlenbernd, 2011; Mühlenbernd and Franke, 2011): more than two agents interact with each other and switch between sender and receiver role. In this way an agent can learn a sender and a receiver strategy as well. Now if such a combination forms a signaling system, it is called a *signaling language* and the corresponding agent is called a *learner*. Thus the number of different possible signaling languages is defined by the number of possible signaling systems and therefore for a $n \times n$-game there are $n!$ different languages an agent can learn. 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 show that the number of possible pooling languages outvalues the number of possible signaling languages for any kind of $n \times n$-game.

3.3 Simulating Bush-Mosteller

Barrett (2009) could show that i) for the simplest variant of a 2×2-game combined with a basic version of reinforcement learning in a 2-players repeated game conventions of meaningful language use emerge in any case, but ii) by extending the domains of the signaling game those conventions become more and more improbable. Furthermore the number of possible perfect signaling systems increases dramatically. This let surmise the motive that up to now researchers applied only the simple variant 2×2-game on populations and keep the hands off domain-extended signaling games. Because if even two players fail to learn perfect signaling from time to time, multiple players will not only have this problem, but also be confronted with an environment evolving to Babylon, where a great many of different signaling systems may evolve.

In his work Barrett simulated repeated signaling games with Roth-Erev reinforcement in the classical sender-receiver variant and computed the *run failure rate* (RFR). The RFR is the proportion of runs not ending with communication via a perfect signaling system.

[3] In this account urn contents and numbers of balls are real numbers

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

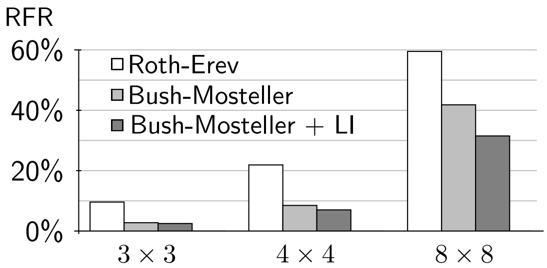


Figure 4: *Left:* Barrett's results for different $n \times n$ games. *Right:* Comparison of different learning dynamics: Barrett's results of Roth-Erev reinforcement in comparison with our results for Bush-Mosteller reinforcement without and with lateral inhibition.

Barrett started 10^5 runs for $n \times n$-games with $n \in \{2, 3, 4, 8\}$. His results show that 100% of 2×2-games were successful (RFR = 0%). But for $n \times n$-games with $n > 2$, the RFR increases rapidly (Figure 4, left).

To compare different dynamics, we started two lines of simulation runs for Bush-Mosteller reinforcement in the sender-receiver variant with urn content parameter $c = 20$ and reinforcement value $\alpha = 1$. For the second line we additionally used lateral inhibition with value $\gamma = {}^1\!/_{|T|}$. We tested the same $n \times n$-games like Barrett for correspondingly 10^5 runs per game. In comparison with Barrett's findings our simulation outcomes i) resulted also 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 4, 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 account, we started 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 sender and as receiver as well and have the opportunity not only to compose a perfect signaling system with a communication partner, but also *learn* a signaling language. As mentioned before: we say that an agent has learned a signaling language if her own sender and receiver strategy forms a signaling system. Furthermore an agent has learned a pooling language if her own sender and receiver strategy form a pooling equilibrium. With this definition we wanted to examine how many signaling languages possibly emerge in a population and how many agents have learned a specific signaling language. For this 3-agents account we started between 500 and 1000 simulation runs with Bush-Mosteller reinforcement ($\alpha = 1$, $c = 20$) for $n \times n$-games with $n = 2 \ldots 8$. Each simulation run stopped, when each agent in the network has either learned a signaling language or a pooling language. We measured the percentage of simulation runs ending with no, one, two or three agents, which have learned a signaling language.

We got the following results: for a 2×2-game, all three agents have learned the same signaling language in more than 80% of all simulation runs. But for a 3×3-game in less than a third of all runs all three agents have learned a signaling language; in more than 40% of all runs two agents have learned a signaling language and the third one a pooling language. And it gets even worse for higher $n \times n$-games. E.g. for an 8×8-game in almost 80% of all runs no agent has learned a signaling language and never have all agents learned a signaling language. Figure 5 depicts the distribution of how many agents have learned

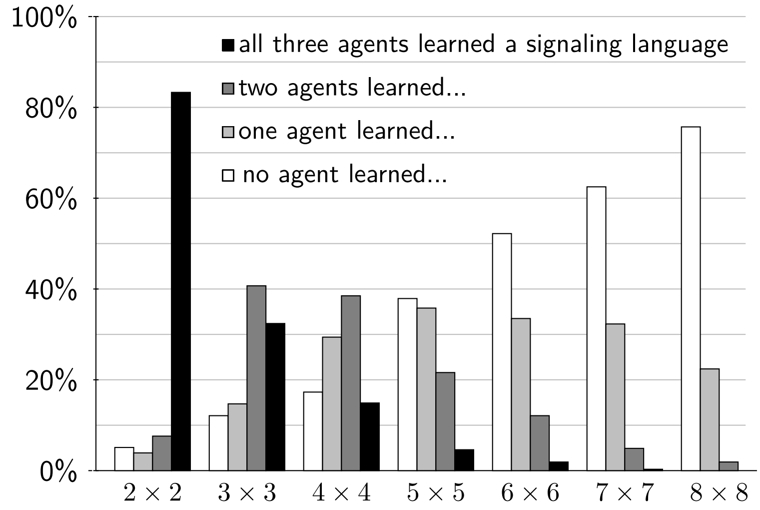


Figure 5: Percentage of simulation runs ending with a specific number of learners of signaling languages in a network with three agents for different $n \times n$-games with $n = 1 \ldots 8$.

a signaling language (no learner, only one learner, two learners or all three agents are learners of a signaling language) for $n \times n$-games for $n = 2 \ldots 8$.[4]

In addition we were interested in whether and how the results would change by extending the number of agents. Thus in another line of experiments we tested the behavior of a complete network of 5 agents for comparison with the results of the 3-agents account. Figure 6 shows the average number of agents who learned a signaling language per run for different $n \times n$-games. As you can see for 2×2-games and 3×3-games the enhancement of population size leads to a higher average percentage of agents learning a signaling language. But for games with larger domains the results are by and large the same.

The results for the classical sender-receiver game reveal that by extending learning accounts the probability of the emergence of perfect signaling systems can be improved but nevertheless is never one for an $n \times n$-game, if n is large enough. Furthermore the results for the multi-agent account with only three agents show that even for a 2×2-game not in any case all agents learn a language. And for games with larger domains results get worse. Furthermore results don't get better or worse by changing the number of agents, as shown in a multi-agent account with 5 agents. But how could natural languages or at least signaling conventions arise by assuming them having emerged from $n \times n$-games with a huge n-value and in a society of much more interlocutors? We'll show that by allowing for the extinction of unused messages and the emergence of new messages, perfect signaling systems emerge for huge n-values and multiple agents in any case. In other words, we'll show that stabilization needs innovation.

[4]Note: further tests with Bush-Mosteller reinforcement in combination with negative reinforcement and/or lateral inhibition revealed that in the same cases the results could be improved for 2×2-games, but were in any case worse for all other games with larger domains.

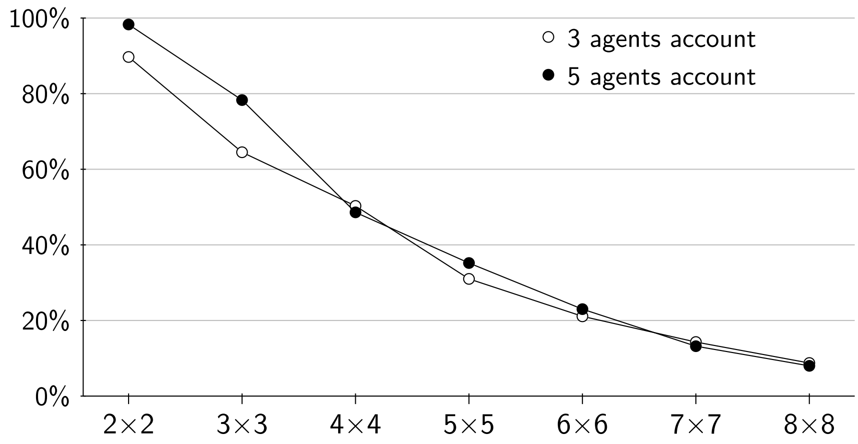


Figure 6: Average percentage of agents learning a signaling language over all runs for different $n \times n$-games with $n = 1 \dots 8$. Comparison of the results of a complete network of 3 agents (white circles) and 5 agents (black circles).

4 Innovation

The idea of innovation in our account is that 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. The idea of innovation and extinction for reinforcement learning applied on signaling games stems from Skyrms (2010), whereby to our knowledge it is completely new i) to combine it with Bush-Mosteller reinforcement plus negative reinforcement and ii) to use it for multi-agent accounts.

4.1 Black Balls and New Urns

The process of the emergence of new messages works like this: additionally to the balls for each message type each sender urn has an amount of *innovative balls* (according to Skyrms (2010) we call them *black balls*). If drawing a black ball the sender sends a completely new message, not ever used by any agent of the population. Because the receiver hasn't a receiver urn of the new message, she 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 equiprobable distributed, 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 dynamic. If the newly invented message doesn't lead to successful communication, the message will be discarded and there won't be a change in the agents' strategies.

As mentioned before, messages can become extinct, and that happens 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 for all sender urns, the message isn't existent in the active use of the 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. Besides, there is no other interference between sender and receiver urn of one agent.

Obviously 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. Furthermore 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. Finally all our results revealed that while during a simulation run the number of messages in the population may varying but each run ends with a number of messages $|M| = |T|$, no matter with what number of initial messages a run started. Thus we call an innovation game with n states and n ultimate messages an $n \times n^*$-game.

4.2 The Force of Innovation

The total number of black balls of an agent's sender urns describes her personal *force of innovation*. Note that black balls can only increase by lateral inhibition in the case of unsuccessful communication and decrease by lateral inhibition in the case of successful communication. This interrelationship leads to the following dynamics: successful communication lowers the personal force of innovation, whereas unsuccessful communication raises the personal force of innovation. If we define the global force of innovation for a group of connected agents X as the average personal force of innovation over all $x \in X$, then the following holds: the better the communication between agents in a group X, the lower the global force of innovation of this group and vice versa. In other words, this account realizes a plausible social dynamics: if communication works, then there is no need to change and therefore the value of the force of innovation is low (or zero, if the society has learned one unique signaling language), whereby if communication doesn't work, the force of innovation increases.

4.3 Learning Languages by Innovation: A Question of Time

We could show in Section 3 that the percentage of agents learning a signaling language in a multi-agent context is being decreased by increasing the domain size of the game. To find out whether innovation can improve these results we started simulation runs with 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 ($\alpha = 1$, $\beta = 1$, $\gamma = 1/|T|$) plus innovation
- initial state: every urn of the sender is filled with black balls, no receiver urns
- experiments: 100 simulation runs per $n \times n^*$-game with $n = 2 \ldots 8$
- 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

The 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 have learned a unique population-wide signaling language in each

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

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

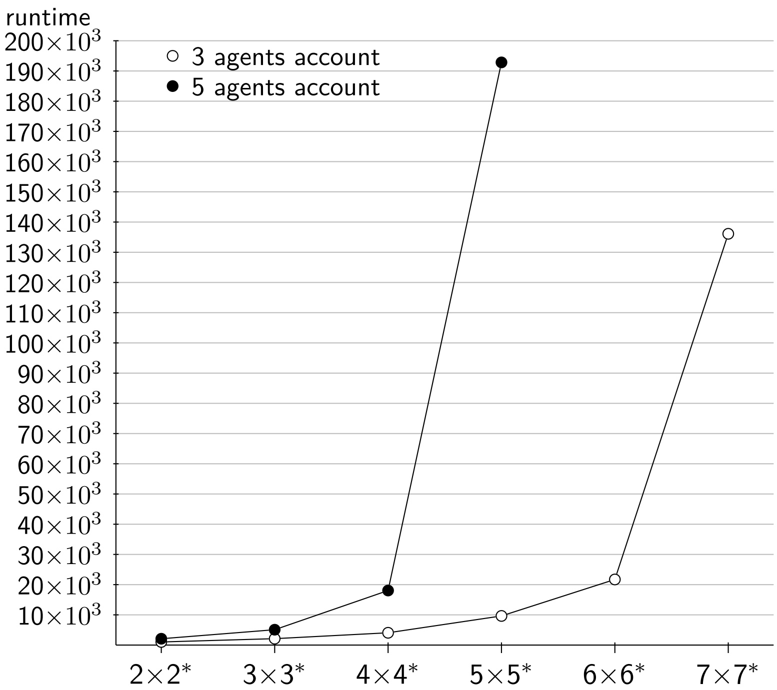


Figure 7: Average runtime over 30 runs for the emergence of a society of learners for $n \times n^*$-games with $n = 2 \ldots 7$; for a complete network of 3 agents and 5 agents.

simulation run and ii) for the remaining account-game combinations all simulation runs exceeded the runtime limit. Table 1 depicts the average runtime for all different games and both population sizes. We expect that for the remaining combination exceeding the runtime limit that all agents will learn a signaling language as well, but it takes extremely long. Figure 7 depicts the Table 1 data. As you can see the runtime increases roughly exponentially in dependency of n for a $n \times n^*$-game. Furthermore the slope of a 5-agents account is much greater than of the 3-agents account. All in all we could show that the integration of innovation leads to a final situation where all agents have learned the same signaling language, if the runtime doesn't exceed the limit. And we expect the same result for account-game combinations where the simulations steps of runs exceeded our limit for a manageable runtime.

4.4 The Course of Learning with Innovation

As our experiments of the last section showed, by applying Bush-Mosteller reinforcement learning with innovation all agents learn the same signaling language for a small group of

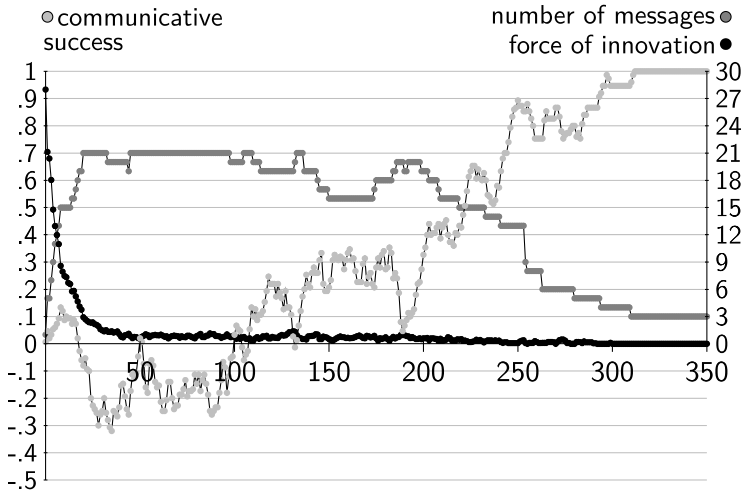


Figure 8: Simulation run of a $3 \times 3^*$-game with innovation in a 3-agents population. Communicative success, number of used messages and force of innovation of all agents in the population; number of simulation steps at x-axis.

agents and any $n \times n^*$-game with $n = 2 \ldots 7$. Let's take a closer look at how a $3 \times 3^*$-game develops during a simulation run by analyzing global properties of the whole population. Three such properties are of special interest here:

· *communication success:* the utility value an agents gained in communication, averaged over the last 20 communication steps and averaged over all agents in the population

· *number of messages in use:* the size of the set of actually used messages in the whole population

· *force of innovation:* absolute number of black balls averaged over all agents

Figure 8 shows the resulting values for the whole population: in the beginning agents are explorative and try out a lot of messages, which 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 here. Then the agents progressively approximate a perfect agreement by sorting out more unsuccessful messages, thus the number of messages decreases, while the communication success increases. And communication success gives lateral inhibition the role for a killer of the force of innovation, which decreases steadily. At the end all agents agreed on 3 messages and interact with perfect communicative success, what signifies that they all use the same signaling language. Finally the force of innovation is vanished, so no urn contains black balls any more.

Admittedly a strong interrelationship of the force of innovation and communication success is not well visible in Figure 8, because of the coarse-grained scaling of the force of innovation value along the y-axis. Fore a more detailed view Figure 9 shows the force of innovation and the communication success between step 50 and 350 of the simulation run,

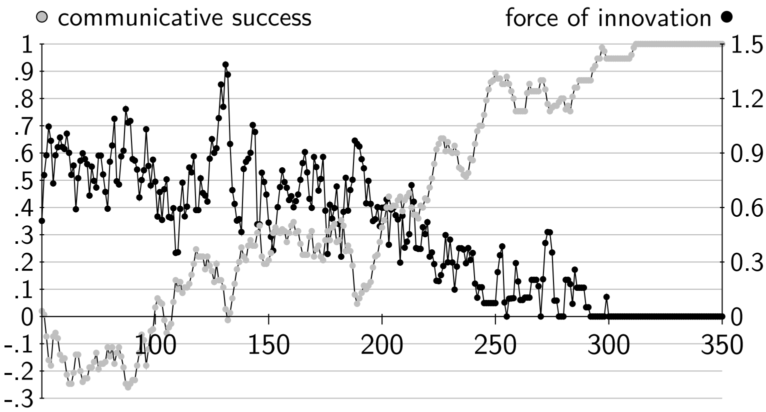


Figure 9: Simulation run of a $3 \times 3^*$-game with innovation in a 3-agents population. Comparison of communicative success and force of innovation; number of simulation steps at x-axis.

as already depicted in Figure 8, but with the force of innovation value that is 20 times more fine-grained. Here the interrelationship between both values is clearly recognizable, one measure's peak is simultaneously the other measure's valley. Admittedly the mirroring is not perfect, but it can be shown that it improves by increasing the number of agents.

Thus what we could show with this experiments is that innovation leads to perfect signaling even for $n \times n^*$-games with a large domain and even in multi-agent populations. Here in the end all agents learn the same global signaling language, what resembles a type of agreement dynamics, similar to the *naming game* account, introduces by Steels (1998) and e.g. applied on multi-agent accounts by Barrat et al. (2007). Furthermore a very interesting and plausible social dynamics is observable: a contrasting interrelationship of communication success and force of innovation. That makes realistically sense, because while poor communication begs for innovation, perfect communication doesn't desire changes.

5 The Evolution of Human Vowel Systems

In this section we extend the hitherto developed account by a metric message space to simulate the evolution of human vowel spaces. Previously Jäger (2008) used signaling games to simulate the evolutionary process leading to the phoneme segmentation of the human vowel space. His idea goes back to an account by Nowak et al. (1999), where the signals are located in some metric space instead of being just atomic entities. In Jäger's experiments, this metric space is the two-dimensional vowel space, which is defined by the first two formants and has roughly the shape of a triangle (for a precise numerical model see Liljencrants and Lindblom (1972)). Considering the fact that most human vowel systems consist of between 3 and 9 vowel categories, Jäger started his experiments with a number of categories between 3 and 9 as a fixed parameter, i.e. it cannot change during one simulation run. In each simulation step 1000 production and 1000 perception events occur, whereby a single round looks like this:

1. pick randomly a vowel category t'

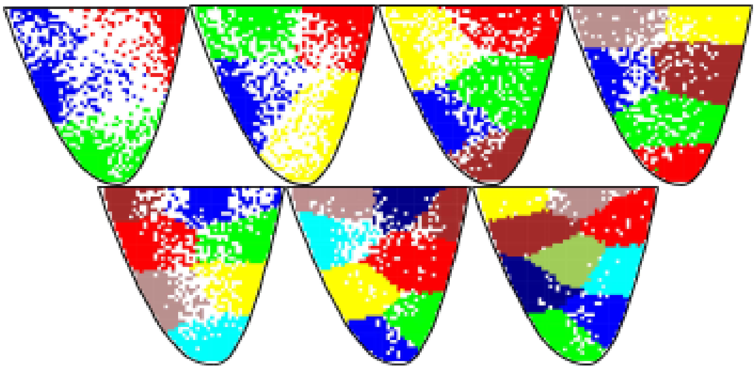


Figure 10: The result of Jäger's simulation experiments for 3 up to nine categories.

2. among all production events in memory having expressed t' choose randomly the signal m' to produce
3. to simulate noise a normally distributed variable is added to change m' to signal m''
4. among all perception events $\langle m, t\rangle$ the one is chosen minimizing the distance between m and m''
5. if $t' = t$ the pair $\langle t', m'\rangle$ is added to the memory of production events and the pair $\langle m'', t'\rangle$ is added to the memory of perception events, otherwise memory remains unchanged

In Jäger's account the population consists of phonetic events instead of language users and can also be interpreted as a realization of *exemplar dynamics* (see e.g. Pierrehumbert (2001)). Jäger's simulation runs stopped after 300,000 iteration steps, exemplary results are depicted in Figure 10, where each color represents a vowel category. These results resemble realistic patterns of human languages. Figure 11 shows a survey of vowel systems across 264 different languages, taken from Schwartz et al. (1997). As Jäger's results show, 6 of the 7 resulting systems corresponds to one of the system in this table.

Reconsider that a simulation step in Jäger's model constitutes roughly a round of play of a signaling game with distance related choice of message production and interpretation and a memory access to guide the decisions. Production events are sender moves (steps 1-2), perception events receiver moves (step 4). Successful communication is rewarded by adding events to the memory what makes this category-signal mapping more probable to be chosen again (step 5). This is basically the way reinforcement learning works. Thus Jäger's model is in some senses strongly related to signaling games with reinforcement learning account like ours. Nevertheless our account is different in some senses, because among other things there is the force of innovation and the relatedness to agreement dynamics. But this is exactly the motivation to adapt our model to make it applicable for experiments to analyze the emergence of vowel space categorization systems like Jäger did. Do see if we get similar and possible additional results? The next subsection describes the extensions necessary for using signaling games with RL dynamics and innovation to apply it on a metric message space.

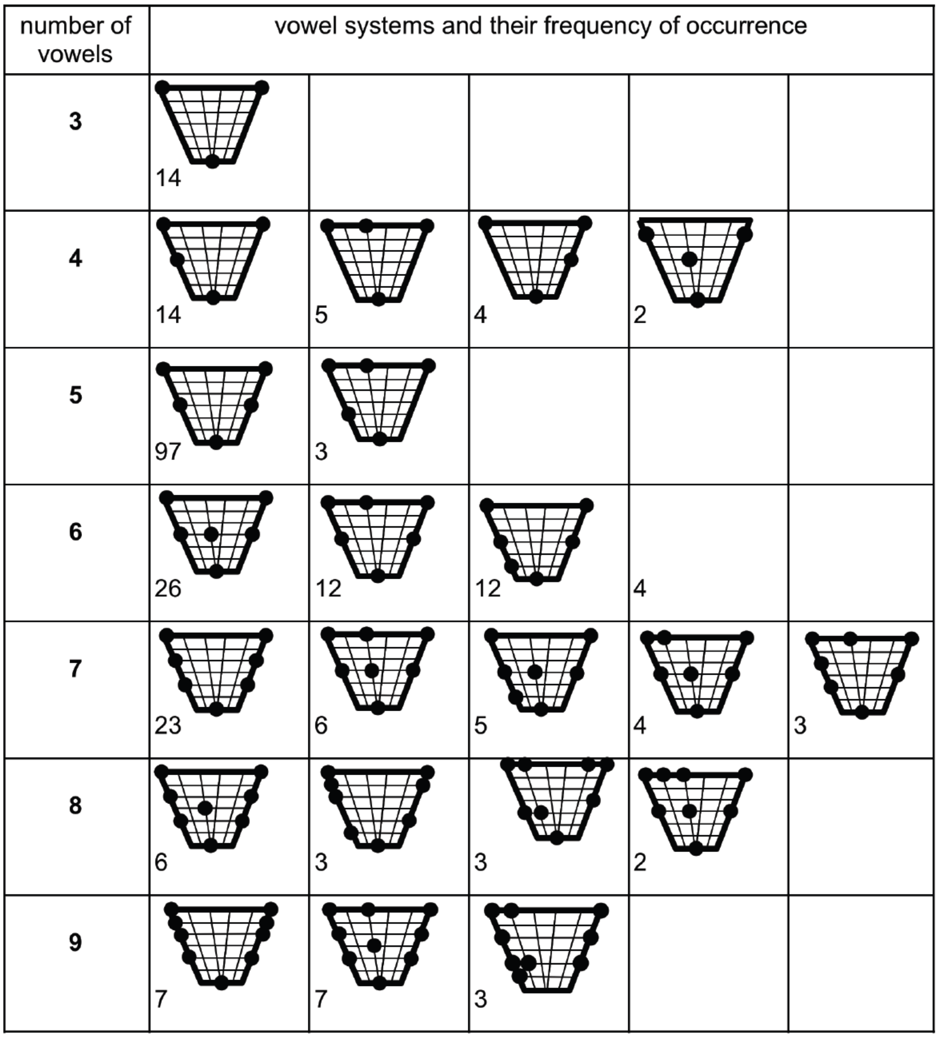


Figure 11: Survey table of vowel systems from Schwartz et al. (1997), published in Journal of Phonetics 25, The dispersion-vocalization theory of vowel systems, 255-286, ©Elsevier.

5.1 Simulating Evolution by Innovation

The basic idea is the following: the metric message space is partitioned into discrete units, of which each one is a possible message. In our experiments we constituted a two-dimensional space that pattern a human vowel space according the first two formants and partitioned it in 141 discrete units, as depicted in Figure 12 (left). Then we let population of agents play a repeated signaling game with reinforcement learning plus innovation as update dynamics, where the set of states T constitutes the vowel categories. Thus by letting them play a $n \times n^*$-game, they have to find an agreement which messages to use for the n vowel categories. Because of innovation the number of messages is flexible, but as we showed in the last section, in the end they will agree on three messages which are then the concrete denotation for the categories. The important extension in comparison to our basic model is that the agents should reconsider the distance between messages in the

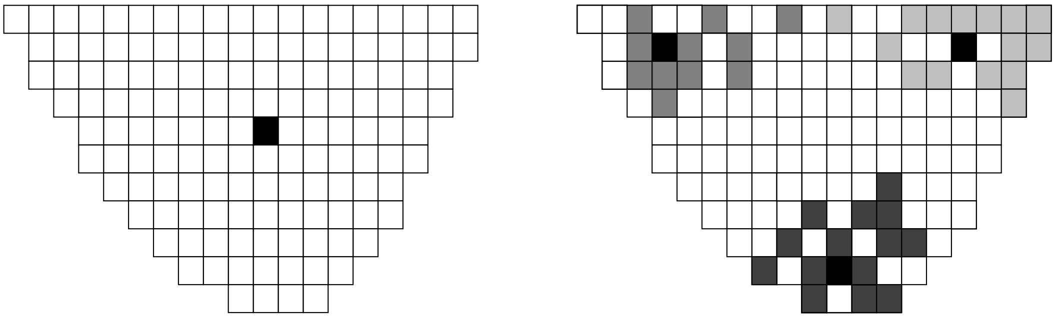


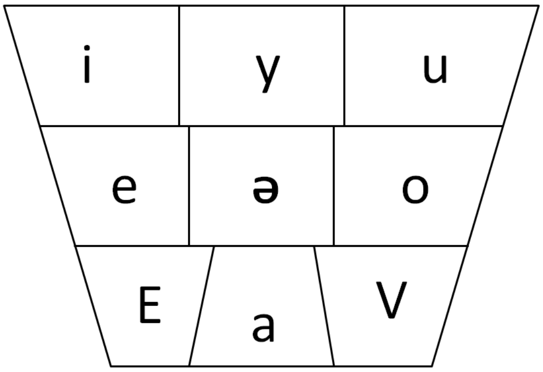
Figure 12: The message space in shape of a vowel space, partitioned in 141 possible messages. *Left:* The initial setting with only one known message for all agents: the black labeled message in the center of the vowel space. *Right:* An exemplary final and stabilized result for 50 agents playing a $3 \times 3^*$ game. The darkgray, mediumgray and lightgray squares depicts all different messages used to communicate t_1, t_2 and t_3, appropriately. The black squares represent the average coordinate of all messages in the appropriate region.

message space for interpretation. And here is the crucial modification: it is possible that an agent can receive a message m' that is not in her set of known messages. In our basic model she adopts it as a new message, thus creates a new receiver urn and appropriately puts balls in the sender urns (see Chapter 4.1). In this modified version the receiver checks if she has a message m'' in her set of known messages, which metric distance to m' is below a predefined threshold ϵ and which has also minimal distance to m'.

To put it formal: given is the situation that an agent x receives an unknown message m'. Let's say that $M_\epsilon(m')$ is the set of all messages in message space M, whose distance is below a given threshold ϵ, thus $M_\epsilon(m') = \{m | m \in M \wedge dist(m, m') < \epsilon\}$. Furthermore $M_{k(x)}$ is the set of known messages of agent x. Now if $M_\epsilon(m') \cap M_{k(x)} \neq \emptyset$, then the agent construes m' as $m'' = \min_{dist(m',m)} m \in M_\epsilon(m') \cap M_{k(x)}$. It is possible that the agent's set of known messages doesn't contain a message close enough to m' ($M_\epsilon(m') \cap M_{k(x)} = \emptyset$). In this case, the agent invents a new message, randomly chosen from the set of unknown messages $M_{unknown} = M - M_{k(x)}$ of the message space according to the basic account. This whole modification considers the fact that receivers cannot necessarily distinguish between two too similar sounds and would construe it with the closest exemplar in their repertoire, if there is one close enough.

We started the first line of experiments with this account for a $3 \times 3^*$-game with 50 agents. Furthermore the starting conditions were as such that all agents had only one message in their set of known messages, namely the most central sound, depicted in Figure 12 (left) as the black square. An exemplary resulting stabilized pattern of such a game is depicted in Figure 12 (right). The darkgray, mediumgray and lightgray squares depicts all different messages used to communicate t_1, t_2 and t_3, appropriately. The black squares represent the average coordinates of each region.

We started 100 simulation runs with these settings to examine the permanence of this exemplary pattern. To give the final pattern a specific name that characterizes a categorization system, we divided the message space in 9 bigger subspaces with names for the prototypical sound for this subspace, as depicted in Figure 13 (left). Then we labeled a specific resulting pattern as $[x\ y\ z]$, if the average coordinates of each region are located



system	share
[*i a u*]	40%
[*i V u*]	9%
[*i* ə *u*]	8%
[*i E u*]	5%
[*e a u*]	4%
[*i E o*]	4%
[*e V u*]	4%
other	26%

Figure 13: The categorization systems for 3 categories.

in the subspaces x, y and z. The first basic result was that for 3 vowel categories agents agree after a while with the following result as depicted in the table of Figure 13: in 40% of all cases the expected system [*i a u*] emerged, but also the other system are close to this system with only one slightly shifted category (e.g. [*i V u*], [*i* ə *u*], [*i E u*] or [*e a u*]) or two in the same direction shifted categories ([*i E o*] or [*e V u*] and all the others). All in all the basic result is that in 40% of all runs the resulting categorization is [*i a u*] and in all other cases it is really close to it and only slightly shifted. This result shows that the trend goes to the expected categorization system of the survey of Figure 11 and there is also some latitude for the emergence of similar systems.

The second basic result is that if the agents play $n \times n^*$-games for $n \geq 4$ and therefore for four or more categories the population could never come to an agreement at least not in the range of number of maximal simulation steps. This fact throws the question if agents in fact will never find an agreement i) because of circular attraction and rejection dynamics accused by the distance-related assessment of message similarities or ii) simply because the runtime of 50 agents agreeing in a $4 \times 4^*$-game lasts verbatim an eternity, what is possibly not far from the truth by taking a look on the slopes of Figure 7 for experiments with only 3 and 5 agents. But to find the answer of this question goes beyond this study. However, how realistic may it be that 4 or more vowel categories evolved simultaneously out of nothing? Shouldn't we rather assume an incremental process?

By thinking about it, it seems realistic to us that in the evolution of vowel categorization humans didn't use all phonemes at once, but maybe had first two or three phonemes and then rearranged the space to adopt additional ones. To reconsider this idea we executed experiments in the following way: we started with an $n \times n^*$-game with a low domain (in this case $n = 3$) and let it run until agents agreed on the n categories. Then during runtime we changed the signaling game by extending it with a further state, so that agents had to rearrange the space to agree on $n + 1$ categories. And in fact this worked: agents found an agreement all the time. With this experiment settings it was possible to draw a *directed acyclic graph* (see Figure 14 for $n \in \{3, 4, 5\}$ and starting node [*i a u*]). The nodes are the resulting vowel systems, the directed edges $A \rightarrow B$ constitute the following relation: vowel system B is the result of vowel system A, if the category is

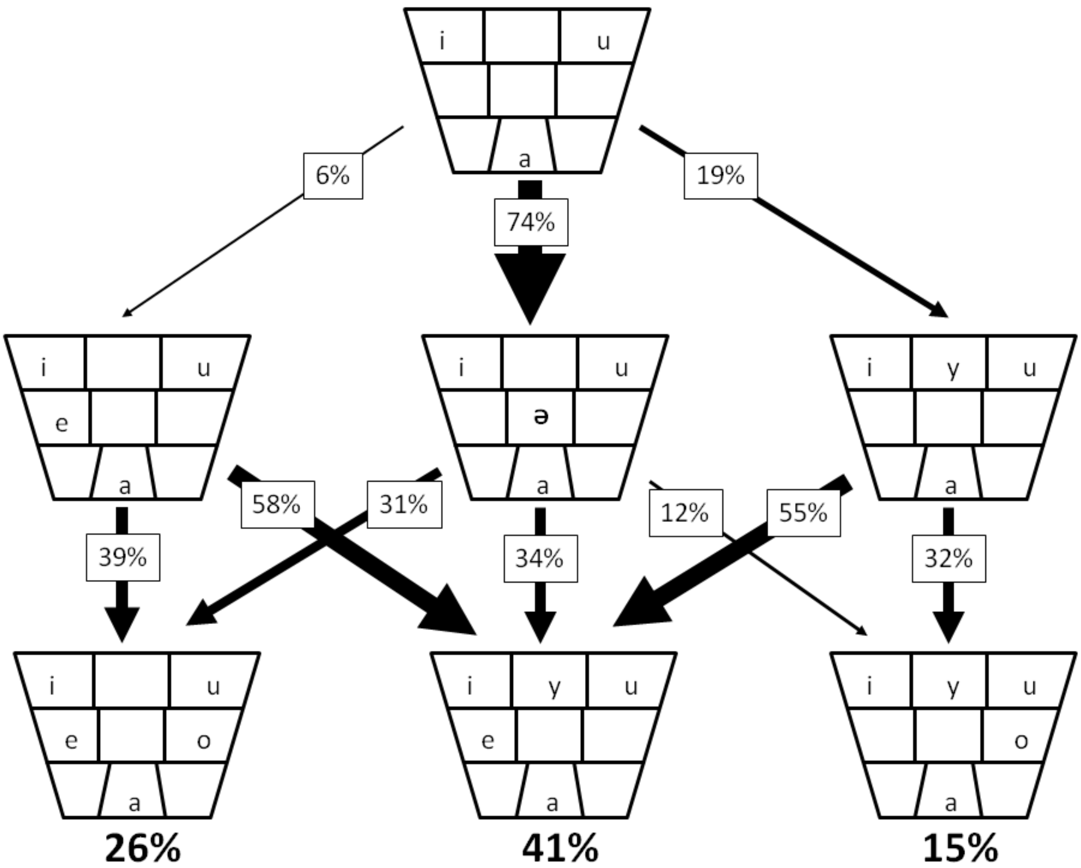


Figure 14: The result of the simulation experiments: a directed acyclic graph of the probabilistic dependencies of the emerged vowel space categorization systems. Further rarely emerged systems are discounted in this graph. The thickness of the arrows corresponds to the percentage.

extended by one. Furthermore the edges are labeled with the percentage of simulation runs, for which B was the result of A. As you can see in Figure 14, the thickness of the edges also represents this percentage value.

Let's take a closer look on this graph. Interestingly by starting with [*i a u*], all simulation runs resulted in one of three systems. For each system it holds that the three old categories are still represented by *i*, *a* and *u*, whereas the new category is represented by something in-between. In the majority of all cases, namely 74%, the resulting system was [*i* ə *a u*], i.e. the new category is represented by the ə. In further 19% of all cases the resulting system was [*i y a u*], so *y* represented the new category. And in the remaining 6% the new category had in *e* its representative and therefore [*i e a u*] was the resulting system. Thus this three systems roughly corresponds to three of the four different 4-vowel systems of the survey table in Figure 11. Only [*i a o u*] never emerged, but all in all these are still promising results.

Furthermore the extension from 4 to 5 categories revealed interesting facts. First of all the two 5-vowel systems [*i e a o u*] and [*i e y a u*] of the survey table of Figure 11 emerged. Also some other 5-vowel systems emerged in our experiments, of which the infrequent ones are not considered in this graph and only a third one, namely [*i y a u o*], was noteworthy and depicted. Interestingly according to the survey [*i e a o u*] is the predominant 5-vowel system, at least by the number of languages. For comparison: the leaves of the directed acyclic graph of Figure 14 are labeled with the percentage of all simulation runs that stopped with the appropriate vowel system, when a 5-vowel system stabilized. Ergo here the system [*i e a o u*] emerged only in a quarter of all runs and is only the second most frequent one, superseded by the system [*i e y a u*] with 41%. Nevertheless it is a nice result that the two most frequently emerged vowel systems correspond to the only two systems in the survey table of Figure 11.

Another interesting result involves the dependencies between the different vowel systems. First of all in each transition step from a 3-vowel to a 4-vowel or from a 4-vowel to a 5-vowel system, the already established regions are never shifted that strong, that the average coordinate switches to a new subspace; and so the new category positions itself somewhere in intermediate space between the already established regions of categories; with one exception: the ə-region emerged in [*i* ə *a u*]. This region and therefore its average coordinate is shifted in any case to a new subspace when it comes to a 5-vowel-system, no matter which one succeeds. This transition effect could possibly be a hint for finding an explanation why there isn't a 5-vowel system with a ə, at least according to the survey.

6 Conclusion and Outlook

Let's begin with a short recap: we started with comparing different reinforcement learning accounts applied on 2-players repeated signaling games: Roth-Erev and Bush-Mosteller reinforcement. We were able to see that Bush-Mosteller especially in combination with lateral inhibition yields better results for repeated signaling games, but far from perfect - extending the domain to more than two states resulted in a fraction of 9,6% for a 3×3-game up to 59,4% for a 8×8-game of all runs failing to develop perfect communication. In more detail we observed that for $n \times n$-games a higher value for n resulted in a lower probability of perfect communication. And results were even worse for multi-agent accounts with 3 agents: with increasing n less agents develop a signaling language in the first place, especially pooling strategies turned out to be a common outcome. Furthermore the results do not change very much by switching from three to five agents, assuming this will hold for bigger populations as well.

In a next step we extended the reinforcement learning account by adding next to negative reinforcement and lateral inhibition a possibility of innovation and extinction of messages. We found that these tweaks result in perfect communication between 3 agents in $n \times n^*$-games for $n < 8$ and between 5 agents for $n < 6$, since higher values for n or the number of agents require much higher runtime that exceeded the simulations' runtime limit. But we expect perfect signaling to evolve in those cases as well. Especially the force of innovation seems to be responsible for this achievement, since it makes sure that new messages are introduced when communication is not successful, while the combination of negative reinforcement and lateral inhibition takes care of all unused or useless messages to become extinct. Consequently, the result is an agreement on one single population-wide perfect signaling language.

In a following step we changed our hitherto account by replacing the discrete message space by a metric message space and let communicative success be dependent on the distance of the messages used by the sender and available in the receiver's repertoire, respectively. With this account we simulated the emergence of human vowel spaces and revealed that agents don't reach an agreement for $n \times n^*$-games with $n > 3$, at least not in a manageable runtime. But we could show that by extending an $n \times n^*$-game to an $(n+1) \times (n+1)^*$-game during runtime and right after agents have already found an agreement for the $n \times n^*$-game, they will also find an agreement for the $(n+1) \times (n+1)^*$-game. Therefore it is possible to find vowel space systems for any number of states $n \in \mathbb{N}$ by increasing n incrementally. These results revealed also interesting succession dependencies between 3-vowel, 4-vowel and 5-vowel systems. Furthermore it may deliver an explanation for the fact that there is no 5-vowel system with a ə, at least according to all languages reconsidered in Figure 11.

An open issue is the examination of n-vowel systems with $n > 5$, which will possibly topic of a succeeding article. For the analysis of the impact of innovation it remains to be shown that our results in fact hold for higher numbers of agents and states. It would further be interesting to see what influence different especially more realistic network-types like e.g. small-world or scale-free networks would have on the agreement performance and what happens if e.g. two or more languages interact.

References

Barrat, Alain et al. (2007). "Agreement dynamics on interaction networks with diverse topologies". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 17.2 (cit. on pp. 1, 12).

Barrett, Jeffrey A. (2009). "The Evolution of Coding in Signaling Games". In: *Theory and Decision* 67, pp. 223–237 (cit. on pp. 1, 5).

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 (cit. on p. 1).

Bush, Robert and Frederick Mosteller (1955). *Stochastic Models of Learning.* New York: John Wiley & Sons (cit. on p. 4).

Hofbauer, Josef and Simon M. Huttegger (2008). "Feasibility of communication in binary signaling games". In: *Journal of Theoretical Biology* 254.4, pp. 843–849 (cit. on p. 1).

Huttegger Simon M.; Skyrms, Brian; Smead Rory; Zollman Kevin J. (2010). "Evolutionary dynamics of Lewis signaling games: signaling systems vs. partial pooling". In: *Synthese* 172.1, pp. 177–191 (cit. on p. 1).

Jäger, Gerhard (2008). "Applications of Game Theory in Linguistics". In: *Language and Linguistics Compass* 2/3, pp. 406–421 (cit. on pp. 1, 12).

Lewis, David (1969). *Convention.* Cambridge: Harvard University Press (cit. on p. 1).

Liljencrants, Johan and Bjorn Lindblom (1972). "Numerical Simulation of Vowel Quality Systems: The Role of Perceptual Contrast". In: *Language* 48.4, pp. 839–862 (cit. on p. 12).

Mühlenbernd, Roland (2011). "Learning with Neighbours". In: *Synthese* 183.S1, pp. 87–109 (cit. on pp. 1, 5).

Mühlenbernd, Roland and Michael Franke (2011). "Signaling Conventions: Who Learns What Where and When in a Social Network". In: *Proceedings of EvoLang IX*, pp. 242–249 (cit. on pp. 1, 5).

Nowak, M. A. et al. (1999). “An error limit for the evolution of language”. In: *Proceedings of The Royal Society of London. Series B, Biological Sciences* 266.1433, pp. 2131–2136 (cit. on p. 12).

Pierrehumbert, Janet B. (2001). “Exemplar dynamics: Word frequency, lenition and contrast”. In: *Frequency effects and the emergence of linguistic structure*. John Benjamins, pp. 137–157 (cit. on p. 13).

Roth, Alvin and Ido Erev (1995). “Learning in extensive-form games: experimental data and simple dynamic models in the intermediate term”. In: *Games and Economic Behaviour* 8, pp. 164–212 (cit. on p. 3).

Schwartz, Jean-Luc et al. (1997). “The Dispersion-Focalization Theory of vowel systems”. In: *Journal of Phonetics* 25.3, pp. 255–286 (cit. on pp. 13, 14).

Skyrms, Brian (2010). *Signals: Evolution, Learning & Information*. Oxford: Oxford University Press (cit. on pp. 1, 8).

Steels, Luc (1998). “The Origins of Ontologies and Communication Conventions in Multi-Agent Systems”. In: *Autonomous Agents and Multi-Agent Systems* 1.2, pp. 169–194 (cit. on pp. 1, 12).

Wagner, Elliott (2009). “Communication and Structured Correlation”. In: *Erkenntnis* 71.3, pp. 377–393 (cit. on pp. 1, 5).

Zollman, Kevin J. S. (2005). “Talking to neighbors: The evolution of regional meaning”. In: *Philosophy of Science* 72.1, pp. 69–85 (cit. on pp. 1, 5).