UNIVERSITÉ PARIS 1 PANTHÉON SORBONNE

Documents de Travail du Centre d'Economie de la Sorbonne

CES Working Papers

Cooperation and evolution of meaning in senders-receivers games

Claude MEIDINGER

2018.36

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
https://centredeconomiesorbonne.univ-paris1.fr/
ISSN : 1955-611X

Cooperation and evolution of meaning in senders-receivers games

Claude Meidinger[1]

Abstract.
Whether there is a pre-existing common "language" that ties down the literal meanings of cheap talk messages or not is a distinction plainly important in practice. But it is assumed irrelevant in traditional game theory because it affects neither the payoff structure nor the theoretical possibilities for signaling. And when in experiments the "common-language" assumption is implicitly implemented, such situations ignore the meta-coordination problem created by communication. Players must coordinate their beliefs on what various messages mean before they can use messages to coordinate on what to do. Using simulations with populations of artificial agents, the paper investigates the way according to which a common meaning can be constituted through a collective process of learning and compares the results thus obtained with those available in some experiments.
Key words: Experimental Economics, Computational Economics, Signaling games

Résumé.
Le fait de savoir s'il existe ou non un « langage » commun préexistant qui détermine les significations littérales des messages cheap talk est manifestement important en pratique. Cependant ce fait est considéré comme non pertinent dans la théorie traditionnelle des jeux car il n'affecte ni la structure des gains ni les possibilités théoriques de signaler. Et quand dans les expériences l'hypothèse d'un « langage commun » est implicitement implémentée, de telles situations ignorent le problème de méta-coordination créé par la communication. Les joueurs doivent coordonner leurs croyances sur ce que signifient les différents messages avant d'utiliser les messages pour coordonner leurs actions. A l'aide de simulations au sein de populations d'agents artificiels, ce papier étudie la manière selon laquelle une signification commune de messages peut se constituer dans le cadre d'un processus collectif de learning et compare les résultats obtenus avec ceux résultant d'expériences.
Mots clés : Economie expérimentale, Economie Computationnelle, Jeux avec communication

Classification JEL: C73, C91, D03.

[1] *Professeur émérite, CES Université Paris 1.*
E-Mail : claude.meidinger@gmail.com

Introduction

According to Tomasello (2008), if we are to understand the ultimate origins of human communication, we must look outside of communication itself and into human cooperation more generally. Human communication is cooperatively structured in ways that the communication of other primates is not. However, every animal communication system cannot be evolutionary stable if it is not beneficial to the signalers and receivers of the signal. « If there is on average no information of benefit to the receiver in a signal, then receivers should evolve to ignore that signal. If receivers ignore the signal, then signaling no longer has any benefit to the signaler », Searcy and Nowicki (2005, 8). Certainly, remarks Sterelny (2003, 183), « deception and trickery are part of human linguistic interaction, but perhaps they are background noise ». Even lying, notes Tomasello (2008, 190), « requires collaboration to get the deceptive message across and a sense of trust on the part of the recipient (or else the lying could not ever work) ».

But before to use communication in an honest or manipulative way, a common meaning must be acquired. Because an audience has an adaptative interest in recognizing a communicative intention, it is first in the interest of all parties to have this intention properly identified. But, in identifying communicative intentions, « there is no arm race between deceptive signaling and vigilant unmasking. Where there is no temptation to deceive, coevolutionary interactions will tend to make the environment more transparent and the detection task less informationally demanding », Sterlny (2003, 180). There is therefore in the language decoding function a first stage that involves no temptation to defect, a stage that not only recognizes communicative intentions as such but also recognizes the social intention of the communicator, whether or not it will be used as an information source about the world and whether or not it will make the receiver's behavior to conform to the wishes of the speaker. Outside of such a context of interactions, a common meaning of signals cannot evolve. We want here to explore that question in the framework of sender-receiver's games by using simulations with artificial agents

I - Games, communication and focal point.

As noted by Crawford (1998), in spite of the fact that games with communication are just extensive-form games with special restrictions on payoff, they are worth considering separately because of the special issues raised by the possibility that a pre-existing common language influences the meaning of messages. At the formal level, because there is nothing in the model to favor one language for messages over another, there is an inessential multiplicity of equilibria associated with many different ways to use messages to support a given equilibrium outcome, besides a possibility of an essential multiplicity of equilibria whenever players' preferences are similar enough. Let us consider two examples in order to make clear those special issues.

Let us begin by considering the very simple matching game used by Blume et al. (1998) in one of their experiments. In this experiment, there is a cohort of $n = 12$ players, 6 senders and 6 receivers. These players are randomly designated as either a sender or receiver at the start of the experiment and keep their designation throughout. In each period of the game, senders and receivers are paired using a random matching procedure. Sender types (t_1 or t_2) are equally likely and independently and identically drawn in every period for each sender that is privately informed about his type. Each sender then chooses a costless message (m_1 or m_2) sent to the receiver that next has to choose an action (a_1 or a_2). Both players' payoffs depend on the sender's type and on the action taken by the receiver, but not on the message sent. Talk is here cheap. The payoffs for senders and receivers are given in the following table 1.

Table 1

	a_1	a_2
t_1	0, 0	700, 700
t_2	700, 700	0, 0

In such a configuration, can a commonly understood communication emerge through repeated local learning interactions between senders and receivers that afford senders to correctly identify their types? Of course, both players prefer that all information be revealed, but if there is a separating equilibrium (for example senders send message m_2 when they are t_1 and message m_1 when they are t_2), there is also another separating equilibrium (in which senders send message m_1 when they are t_1 and message m_2 when they are t_2).

Let us next consider in table 2 the canonical game BOS (Battle of the Sexes) similar to the one used by Cooper et al. (1990, 1994) in their experimental investigation of coordination.

Table 2

	a_1	a_2
x_1	3, 1	0, 0
x_2	0, 0	1, 3

This is a well-known game with asymmetric payoffs so that players disagree about which equilibrium is best. Both would like to match but row player prefers the equilibrium (x_1, a_1) and column player the equilibrium (x_2, a_2). Note that the two pure equilibria are symmetric, up to the identity of the players. Hence, only selection principles that distinguish one player from the other will help them to coordinate. Now, we can add to that game a process of preplay communication in which one of the two players, the sender can, before choosing an action, send to the other player, the receiver, a message, either m_1 or m_2. After that, each player simultaneously chooses an action. With row player being the sender, we can model the situation as an extensive-form game and in the reduced normal form of this game:
- the sender strategies are (m_1, x_1), (m_1, x_2), (m_2, x_1) and (m_2, x_2)
- the receiver strategies are (a_1, a_1), (a_1, a_2), (a_2, a_1) and (a_2, a_2), where we first list the decision that the receiver would make if the sender sends m_1 and second the decision that he would make if the sender sends m_2.
The reduced normal representation of this game is shown in table 3, with the different equilibria in bold style.

Table 3

	$(a_1,\ a_1)$	$(a_1\ a_2)$	$(a_2,\ a_1)$	$(a_2,\ a_2)$
$(m_1,\ x_1)$	**3, 1**	**3, 1**	0, 0	0, 0
$(m_1,\ x_2)$	0, 0	0, 0	1, 3	**1, 3**
$(m_2,\ x_1)$	**3, 1**	0, 0	**3, 1**	0, 0
$(m_2,\ x_2)$	0, 0	1, 3	0, 0	**1, 3**

It is interesting to note here that the adjunction of a preplay process of communication does not allow to select one of the two equilibria of the game without communication. These two equilibria are still possible with different meaning of messages. For instance, in the equilibrium [$(m_2,\ x_1),\ (a_2,\ a_1)$], the message m_2 is for column player the signal that row player will choose x_1 and row player choose m_2 because m_1 means for both players that x_2 will be chosen. However, to the strategy $(m_2,\ x_1)$ is associated another equilibrium (a babbling one) in which column player, playing always the same action, does not give to the messages any signification. In that case too, $(m_1,\ \ x_1)$ is also a best response to $(a_1, a_1,)$. In other terms, without a common preexisting understanding of messages, it is the different possible equilibria with communication that determine the meaning of messages. At the formal level, because there is nothing in the model that favors a particular meaning, there is always an inessential multiplicity of equilibria associated with the different ways of using messages to support a given equilibrium outcome.

But let us suppose with Myerson (1991) that a common language is used by the players and that decisions x_1 and a_1 are "I go to the football match" and x_2 and a_2 are "I go to shopping". Now, if m_1 is the expression "let's meet at the football match" and m_2 is the expression "let's meet at the shopping center", then the obvious focal equilibrium is $\{(m_1,$ $x_1)$, $(a_1,\ a_2)\}$. In this equilibrium, row player says "let's meet at the football match" and goes to the football match, and column player would try to meet row player wherever he suggests. As was emphasized by Myerson (1991, 111), this equilibrium is the unique equilibrium in which row player' statement is interpreted according to its literal meaning. So, the players' shared understanding of language which is part of their cultural heritage, may make focal this equilibrium.

II – Learning to communicate

II-1 – Where does meaning come from?

In experiments, the common-language assumption is usually implemented by requiring subjects to communicate in a restricted language that consists of the labels with which their decisions in the underlying game are presented. In that way, some equilibria might be focal because of precedent or physical or semantic distinctions in the way strategies are labeled. But in such situations, the meta-coordination problem created by communication is ignored. Players must coordinate their beliefs on what various messages mean before they can use messages to coordinate on what to do. And "the point is that creating meaning from scratch is deceivingly difficult and may take a long time. Where does meaning come from? Common sense is not a satisfactory answer for theorizing since we would like to know where common sense comes from and how it is sustained", Camerer (2003, 357).

In sender-receiver games, trying to answer that question is closely related to many problems encountered in the traditional conception of equilibrium, Fudenberg and Levine (1998). In particular, in case of multiple equilibria, without an explanation of the way according to which players coordinate on one of them, no guarantee exists concerning the realization of any of those equilibria. Therefore, to answer the question, two principles are important, Peyton Young (1998). The first one tells us that an equilibrium cannot be understood without a dynamical construction that explains its realization. The second one tells us that we must have a model defining the way according to which agents coordinate: without being hyper-rational, agents are not devoid of rationality in their learning process of coordination.

Messages can have two different roles - signaling intentions (BOS game) or signaling private information (Blume et al. (1998)). But in both cases, if theory is unable to specify conditions under which specific meanings arise, one can put people in a situation where there is no meaning and see how they created it. In an interesting paper, Blume et al. (1998) did such an experiment in which players in sender-receiver games with private information could learn to create a homemade "language" whose meaning could be understood by receivers, using abstract message spaces and a label-scrambling technique to eliminate any possible effects of a common preexisting language. In this process, meaning slowly emerges through an evolutionary process of adaptation and selection of messages, but without explicit formalization of the way according to which players collectively agree to coordinate on an equilibrium. We want here to use that approach to try to answer the following question[2]: How can people succeed to come to an agreement about a common meaning through a collective process of learning in which messages may have different roles - either signaling intentions or signaling private information? In investigating the answer with artificial agents, we have to specify the process of learning used by the agents and to keep an eye on the possibility previously noted that where there is no temptation to deceive, coevolutionary interactions may favor a common meaning.

II-2 – Learning to communicate: EWA models.

There is a wide range of learning models, from naïve behavioral models to sophisticated ones. In experimental economics, one can distinguish two broad classes of learning models. In the first class are simple ad hoc stimulus response models in which the learning process is only pure reinforcement-based, see for instance Roth and Erev (1995). In the second class, one can find models in which individuals try to exercise more cognitive ability by using other available information in order to try to guess their opponents' decisions and thus play a best

[2] Such an approach is very similar to that used in many experiments with robots. In the Computer Science Lab of SONY for instance, experiences called "Talking Heads" with robot-dogs AIBO are used to show how repeated interactions between automatons allow them to agree on a common vocabulary. From a limited register of randomly and personal terms, communicating by radio, these robot-dogs succeed to come to an agreement about terms used to designate things in their environment.

response to those decisions. These belief-based learning models embody a higher level of rationality, see for instance Cheung and Friedman (1997)[3]. One can also try to estimate a more general learning model such as EWA (Experience-Weighted Attraction) learning models that hybridize the two popular approaches to learning in games, reinforcement and belief formation, Camerer and Ho (1999). Adaptation of that model to signaling games shows that for some empirical investigations, this type of model performs significantly better than its choice reinforcement and belief-based special cases, Anderson and Camerer (2000). We therefore choose such a EWA learning model for our artificial agents.

EWA learning models belong to a family of stochastic dynamic models of individual choices. These models determine the probability that a player will choose a given strategy at time t and how these probabilities evolve over time in response to the player's experience. Individuals start with initial propensities or attractions in terms of which probabilities of choice are expressed. If A_{ts} is the attraction at round t of pure strategy s for a player X, the probabilities of choice pA_{ts} of the different (pure) strategies are determined according to

$$pA_{ts} = \frac{exp(\lambda A_{ts})}{\sum_i exp(\lambda A_{ti})} \qquad (1)$$

with i index of the different strategies of player X. The parameter $\lambda \geq 0$ measures the sensitivity of players to attractions. The probability to play a strategy increases with its attraction, except for $\lambda = 0$, in which case a purely random choice is made by the player.

These propensities or attractions are next updated in response to the player's experience. Let us therefore suppose that at round t, player X playing with another player Y, plays a strategy s and player Y a strategy r. With g_X the payoff function of player X, at round t+1, his attractions are then updated according to the rules[4]

- for the strategy s played at round t

$$A_{(t+1)s} = h[A_{ts} + g_X(s,r)] \qquad (2)$$

- for the strategies i not played at round t

$$A_{(t+1)i} = h[A_{ti} + \delta g_X(i,r)] \qquad (3)$$

In EWA learning models, because choices can both reflect past success of chosen strategies (reinforcement-based) and the history of how others played (belief-based), the crucial feature is how strategies are reinforced. In pure reinforcement-based models, only chosen strategies are reinforced by the payoffs they earn, so that $\delta = 0$. But if one wants to mix elements of reinforcement and belief learning, one has to reinforce both chosen and unchosen strategies. Chosen strategies are reinforced by the full amount of the payoffs they earn, unchosen ones only by a fraction δ of the payoffs they would have earned if played.

---

[3] Using the data of their 1998 experiment and focusing only on senders' behavior, Blume et al. (2002) compare the two types of models and find that both types fit the data reasonably well in games where the preferences of senders and receivers are perfectly aligned and where the population history of the senders is known by all players. But they also conclude that they are quite agnostic about the winning learning model and for that matter which one of the alternative models wins. Some econometric results suggest that both models may fail to capture the strategic aspects of the games and the respective learning involved.

[4] Les fonctions h s'écrivent $\frac{\varphi N_t A_{ts} + g(s,r)}{N_{t+1}}$ pour la stratégie jouée et $\frac{\varphi N_t A_{ti} + \delta g(i,r)}{N_{t+1}}$ pour les stratégies non jouées, avec $N_{t+1} = \rho N_t + 1$. Les paramètres du modèle sont donc N, λ, φ, δ and ρ.

However, sender-receiver games are extensive-form games and this property demands some modifications if we want to extend EWA models to these cases. "The key problem", note Anderson and Camerer (2000, 691), "is that players do not always know the forgone payoff to a signal which they do not choose (because its payoff depends on other players' reactions to the unsent signal, which is usually not known)". For instance, in a simple signaling game with private information in which senders can be of different types (t_1 or t_2) and can choose between different messages (m_1 or m_2) in order to inform receivers of their type and induce them to choose the right action (a_1 or a_2), special modifications have to be introduced[5].

For sender i, let AS_{itm} be her attraction for message m when she is of type t. If that sender is of type t_1, sends message m_2 and gets response a_1, she knows her payoff and updates AS_{it1m2}, reinforcing it by the full amount of the payoff it earns. But by having sent that message, the sender also knows that if she had been of type t_2 and sent m_2, she would have earned the payoff of the issue (t_2, a_1). She therefore also updates AS_{it2m2}, but only by a fraction δ of the payoff it would have earned if played. For messages no sent, because the possible response of the receiver is unknown, the corresponding attractions are not updated.

For receivers, the problem is easier because these players can condition only on the senders' message and therefore they know all their forgone payoffs. Let AR_{jma} be the attraction of receiver j for action a when she receives message m. If, for example, receiver j take action a_1 in response of message m_1 when the sender was a t_2 she updates AR_{jm1a1} by the full amount of the payoff it earn, and AR_{jm1a2} only by a fraction δ of the payoff it would have earned if played. For the messages not received, the attractions are unchanged.

III - Coordination by evolution of common meaning in a matching game with private information.

We investigate that question in populations of 12 artificial agents, 6 senders and 6 receivers that revised their decisions according to a EWA process, in a configuration similar to that implemented by Blume et al. (1998). In each round of the simulation, senders and receivers are paired using a random matching procedure. Sender types (t_1 or t_2) are equally likely and independently and identically drawn in every round for each sender privately informed about his type. Each sender then chooses a costless message (m_1 or m_2) sent to the receiver that next choose an action (a_1 or a_2). Table 4 below shows what can happen in a particular round.

[5] Following Anderson and Camerer (2000), one must note that we do not define for senders contingency complete strategies which specify a message for each type, for instance ($m_1|t_2$ and $m_2|t_1$). As noted by the authors, this approach begs the question of how to update attractions for several complete strategies which have the same used portion but different unused portions. One rather takes another approach in which a sender chooses a strategy (a message) only after knowing his type. This is similar to the "agent form" game in which each node is played by a "different" agent for a single player.

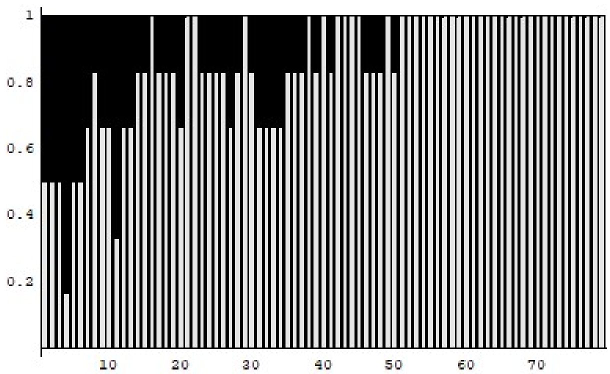
In this round for example, the sender 1 of type t_2 is associated with the receiver 5. She sends message m_2 and receiver 5 responds by choosing a_2. The payoff of both agents is zero. These payoffs are then used by the agents to update their attractions and to make choices at the next round.

Table 4

sender	type	message	receiver	action
1	t_2	m_2	5	a_2
2	t_2	m_2	2	a_1
3	t_1	m_1	3	a_2
4	t_1	m_2	4	a_1
5	t_2	m_1	1	a_2
6	t_2	m_1	6	a_2

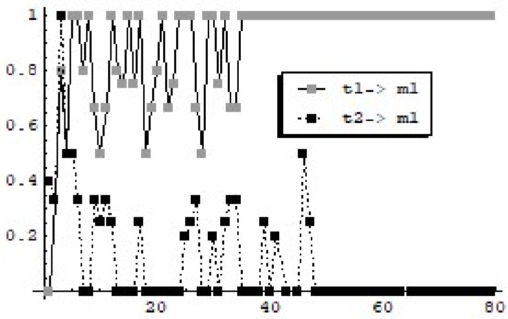
Considering the 48 attractions concerned by the updating process at each round (pairs of type-message for senders and pairs of message-action for receivers), we run simulations starting with initial values of zero for attractions that give identical initial conditional probabilities respectively to messages and to actions. For both senders and receivers, the parameters of the learning model are $N = 1,\ \lambda = 1,\ \varphi = 0.5,\ \delta = 0.5$ and $\rho = 0.6$. The following graphics display the results obtained in a simulation over 60 rounds.

Graphic 1 below displays a frequency histogram obtained over 60 rounds (in white frequencies of issues (t_1, a_2) or (t_2, a_1) that are consistent with a population playing a separating equilibrium, in black frequencies of other issues). There is here clearly an evolution toward a perfect coordination between types and actions through messages.

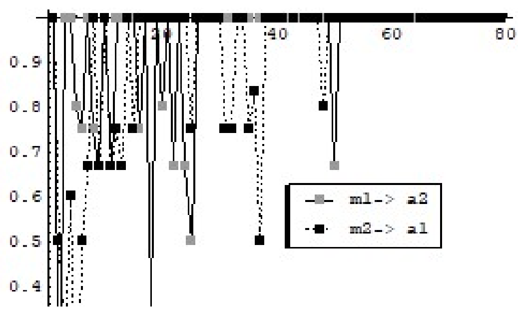


Graphic 1

Note that for computers, there is no initial common language, no special connection respectively between m_1 and t_1 on the one hand and between m_2 and t_2 on the other hand that could suggest, by the similarity of the indexes, a particular focal point of meaning. The way according to which a common language progressively evolves in the population is displayed by graphics 2 and 3 below. Graphic 2 shows the evolution of the proportion of type t_1 that send message m_1 and the evolution of the proportion of type t_2 that send the same message. Here, types t_1 specialize in messages m_1 and types t_2 in messages m_2 . Graphic 3 shows the evolution of the proportion of action a_1 that follows m_1 and the evolution of the proportion of action a_2 that follows m_2 . For receivers, m_1 means t_1 because they always play a_2 when receiving that message and m_2 means t_2 because they always play a_1 when receiving that message. We can have an inessential multiplicity of equilibria with other simulations leading to different meanings in which t_1 is identified by m_2 and t_2 identified by m_1.



Graphic 2



Graphic 3

Note also that with EWA learning processes, contrasting to some theoretical and experimental works that study how particular dynamics can guarantee communication in cheap-talk games[6], it is not necessary that players have access to information about the history of the population play. With only the private information they could gather through their personal experiences, they participate to a collective learning of a common meaning that may allow them to realize a coordination through messages. Table 5 below displays for each group of 500 simulations ran for different numbers of rounds the percentage of simulations that succeed in realizing such a coordination (characterized by the fact that the population plays the issues of a separating equilibrium at least over the 20 last rounds) in populations of size 12 and 16.

Table 5: Matching Game

Number of periods	40	60	80	100	200
% perfect coordination, n=12	1.4	10.2	28.8	52.8	98
% perfect coordination, n=16	0	4.2	19.4	45.2	97

[6] In Blume et al. (1998), it is important that players have access to information about the history of population play and not only about the private information they could gather through their personal experiences. Except for one special treatment *NH* (No History), they use protocols that allow all players to receive at the end of each round information about all sender types and all messages sent by the respective sender types. For the special treatment *NH* however, in contrast with the other treatments that clearly show over 20 periods the emergence of a "common language", without information about the history of population's play no consistent evolution toward a meaningful communication emerges.

IV - Coordination by evolution of common meaning in games with asymmetric payoffs. Signaling intentions.

Let us now consider the game BOS (Battle of the Sexes) below used by Cooper et al. (1990, 1994) in their experimental investigation of coordination.

Table 6

	a_1	a_2
x_1	0, 0	200, 600
x_2	600, 200	0, 0

This is a game with asymmetric payoffs so that players disagree about which equilibrium is best. Both would like to match but row player prefers the equilibrium (x_2, a_1) and column player the equilibrium (x_1, a_2). There is also a third mixed-strategy equilibrium with prob(x_1) = prob (a_1) = 0.25, in which the expected payoffs 150 are worse for both players than either of the pure- equilibrium payoffs strategy.

Table 7 displays some results obtained from different experimental treatments made by Cooper et al. (1994) and Straub (1995): percentage of pairs choosing each of the two equilibria and either of the disequilibria, from the last half of a twenty-two rounds experiment. In *Coop* and in *Straub*, the game is played without communication and in *Coop-Com*, it is played as a sender-receiver game in which one player, the row player, makes a non-binding preplay announcement.

Table 7

Treatment	(x_1, a_2)	(x_2, a_1)	Other issues
Coop	22%	19%	59%
Straub	26%	14%	60%
Coop-Com	1%	95%	4%

Obviously, without communication, coordination failure is common. And a one-way cheap talk communication breaks the symmetry of the roles and works for the profit of the sender. But in this experiment, players have a common language that restricts their beliefs about the meaning of the announcements and can participate to a focal point because psychological or semantic associated distinctions.

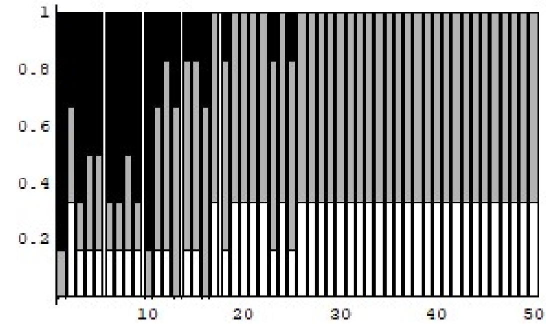
We use here simulations with artificial agents to investigate if EWA learning allows the evolution of common meanings in a population of players playing the BOS sender - receiver game of table 7. We do that in the same way as before, by considering a cohort of 12 players, 6 senders and 6 receivers. In each period of the game, senders and receivers are paired using a random matching procedure. Each sender first chooses a costless message (m_1 or m_2) sent to the receiver. Next, sender and receiver simultaneously choose an action, (x_1 or x_2) for the sender, (a_1 or a_2) for the receiver. Both players' payoffs depend on the actions chosen, but not on the message sent: talk is here also cheap. Table 8 shows what can happen in a particular round.

Table 8

Sender	Sender's choice	Receiver	Receiver's choice
1	(m_2, x_1)	4	a_2
2	(m_1, x_2)	1	a_1
3	(m_1, x_1)	3	a_2
4	(m_2, x_2)	5	a_2
5	(m_1, x_1)	2	a_1
6	(m_1, x_2)	6	a_1

In each period, according to the EWA learning process, attractions are updated. For each sender, we have four attractions to update, AS_{imx} being for sender *i* the attraction for (*m, x*). We also have for each receiver four attractions to update, AR_{jma} being for receiver j the attraction for action a when she receives message m. Here too, chosen strategies are reinforced by the full amount of the payoffs they earn, unchosen ones only by a fraction δ of the payoffs they would have earned if played. If, for example, sender i sends message m_1, take action x_2 and gets response a_1, she updates AS_{im1x2} , reinforcing it by the full amount of the payoff it earn. But that player also knows that if she had played x_1, she would have earned the payoff of the issue (x_1, a_1). She therefore also updates AS_{im1x1} , but by a fraction δ of the payoff it would have earned if played. For the message no sent, because the possible response of the receiver is unknown, the attractions are not updated. In the same way, for a receiver j choosing a_1 in response of message m_1, we update AR_{jm1a1} by the full amount of the payoff it earns, and AR_{jm1a2} only by a fraction δ of the payoff it would have earned if played. Considering the 48 attractions possibly concerned by the updating process at each period, we run simulations, starting with initial values of attractions (equal to zero) that give identical initial conditional probabilities of choice and with the same values of parameters as before.

Graphic 4 displays the frequency histogram obtained for the evolution of the different types of issues (in white *S*-issues (x_2 , a_1) that are favored by senders, in grey *R*-issues (x_1, a_2) that are favored by receivers and in black the other issues). The graphic shows that after a few numbers of rounds, the evolution converges to a stationary state in which *S*-issues are played with a frequency of 2/6 and *R*-issues with a frequency of 4/6.



Graphic 4

The manner according to which players in the population succeed in avoiding issues of disequilibrium is depicted in the following Table 9 that shows for each sender and each receiver the equilibrium probabilities, $prob(m,x)$ for senders and $prob(m,a)$ for receivers. In each line corresponding to a sender, we have the following elements

$$\begin{pmatrix} prob(m_1,x_1) & prob(m_2,x_1) \\ prob(m_1,x_2) & prob(m_2,x_2) \end{pmatrix}$$

and in each line corresponding to a receiver, we have

$$\begin{pmatrix} prob(m_1,a_1) & prob(m_2,a_1) \\ prob(m_1,a_2) & prob(m_2,a_2) \end{pmatrix}$$

Table 9

Senders		Receivers	
$\begin{pmatrix} 3.25749\times10^{-70} \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.25768\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 3.4566\times10^{-209} \\ 1. \end{pmatrix}$
$\begin{pmatrix} 3.4566\times10^{-209} \\ 1. \end{pmatrix}$	$\begin{pmatrix} 3.4566\times10^{-209} \\ 3.4566\times10^{-209} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.28249\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 3.4566\times10^{-209} \\ 1. \end{pmatrix}$
$\begin{pmatrix} 3.25749\times10^{-70} \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.26112\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 3.4566\times10^{-209} \\ 1. \end{pmatrix}$
$\begin{pmatrix} 3.4566\times10^{-209} \\ 1. \end{pmatrix}$	$\begin{pmatrix} 3.4566\times10^{-209} \\ 3.4566\times10^{-209} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.25788\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 3.4566\times10^{-209} \\ 1. \end{pmatrix}$
$\begin{pmatrix} 3.25749\times10^{-70} \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.28246\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 3.4566\times10^{-209} \\ 1. \end{pmatrix}$
$\begin{pmatrix} 3.25749\times10^{-70} \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.46171\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 3.4566\times10^{-209} \\ 1. \end{pmatrix}$

Here two senders (sender 2 and 4) always choose m_1 followed by x_2 and four senders always choose m_2 followed by x_1. Because in each round senders and receivers are paired using a random procedure, such choices lead to a stationary state only because receivers have acquired a common understanding of messages. For all of them, m_1 means "I play x_2" because they always respond to m_1 by a_1 and m_2 means "I play x_1" because they always respond to m_2 by a_2. But though communication succeeds in creating a homemade meaning that is well understood by the receivers, such a common understanding is not known by the senders. It is clear that if such a common understanding of receivers would have been known by senders, all of them would have sent m_1 and plaid x_2.

Contrary to the evolution of meaning in the matching game with private information in which players have no conflict of interest, in the BOS game, some stationary states obtained can be considered as babbling equilibria as exemplified by the following two simulations.

A babbling equilibrium that favors senders.

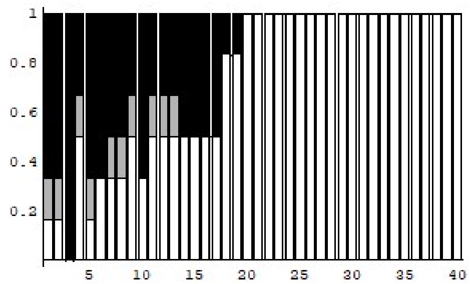


Table 10

Senders		Receivers	
$\begin{pmatrix} 3.4566\times10^{-209} \\ 3.4566\times10^{-209} \end{pmatrix}$	$\begin{pmatrix} 3.4566\times10^{-209} \\ 1. \end{pmatrix}$	$\begin{pmatrix} 1. \\ 4.41753\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.25749\times10^{-70} \end{pmatrix}$
$\begin{pmatrix} 3.4566\times10^{-209} \\ 3.4566\times10^{-209} \end{pmatrix}$	$\begin{pmatrix} 3.4566\times10^{-209} \\ 1. \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.31914\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.25749\times10^{-70} \end{pmatrix}$
$\begin{pmatrix} 3.4566\times10^{-209} \\ 3.4566\times10^{-209} \end{pmatrix}$	$\begin{pmatrix} 3.4566\times10^{-209} \\ 1. \end{pmatrix}$	$\begin{pmatrix} 1. \\ 1.5496\times10^{-68} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.25749\times10^{-70} \end{pmatrix}$
$\begin{pmatrix} 3.4566\times10^{-209} \\ 1. \end{pmatrix}$	$\begin{pmatrix} 3.4566\times10^{-209} \\ 3.4566\times10^{-209} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.27337\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.25749\times10^{-70} \end{pmatrix}$
$\begin{pmatrix} 3.4566\times10^{-209} \\ 3.4566\times10^{-209} \end{pmatrix}$	$\begin{pmatrix} 3.4566\times10^{-209} \\ 1. \end{pmatrix}$	$\begin{pmatrix} 1. \\ 2.11059\times10^{-69} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.25749\times10^{-70} \end{pmatrix}$
$\begin{pmatrix} 3.4566\times10^{-209} \\ 3.4566\times10^{-209} \end{pmatrix}$	$\begin{pmatrix} 3.4566\times10^{-209} \\ 1. \end{pmatrix}$	$\begin{pmatrix} 1. \\ 1.02324\times10^{-69} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.25749\times10^{-70} \end{pmatrix}$

In this stationary state, senders send different messages while choosing all the same action x_2, (5 senders send m_2 and one m_1), and all the receivers play a_1 whatever the message.

A babbling equilibrium that favors receivers

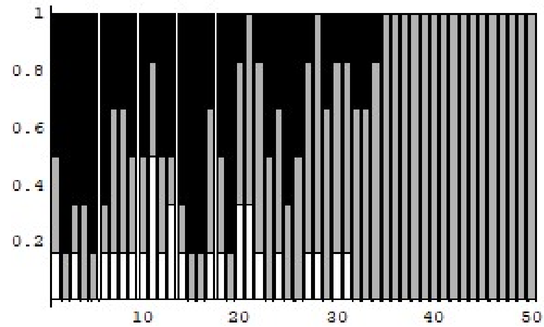


Table 11

Senders		Receivers	
$\begin{pmatrix} 3.25749\times10^{-70} \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 4.25765\times10^{-209} \\ 1. \end{pmatrix}$	$\begin{pmatrix} 3.45664\times10^{-209} \\ 1. \end{pmatrix}$
$\begin{pmatrix} 1. \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 3.25749\times10^{-70} \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 3.4985\times10^{-209} \\ 1. \end{pmatrix}$	$\begin{pmatrix} 3.49816\times10^{-209} \\ 1. \end{pmatrix}$
$\begin{pmatrix} 1. \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 3.25749\times10^{-70} \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 3.46157\times10^{-209} \\ 1. \end{pmatrix}$	$\begin{pmatrix} 3.61863\times10^{-209} \\ 1. \end{pmatrix}$
$\begin{pmatrix} 3.25749\times10^{-70} \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 3.46669\times10^{-209} \\ 1. \end{pmatrix}$	$\begin{pmatrix} 3.46469\times10^{-209} \\ 1. \end{pmatrix}$
$\begin{pmatrix} 3.25749\times10^{-70} \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.2575\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 5.2337\times10^{-209} \\ 1. \end{pmatrix}$	$\begin{pmatrix} 3.45677\times10^{-209} \\ 1. \end{pmatrix}$
$\begin{pmatrix} 3.25749\times10^{-70} \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 1. \\ 3.25749\times10^{-70} \end{pmatrix}$	$\begin{pmatrix} 3.48122\times10^{-209} \\ 1. \end{pmatrix}$	$\begin{pmatrix} 3.45737\times10^{-209} \\ 1. \end{pmatrix}$

In this stationary state, senders send different messages while choosing all the same action x_1, (4 senders send m_2 and two m_1), and all the receivers play a_2 whatever the message.

Finally, table 12 below displays for each group of 500 simulations ran for different numbers of periods the percentage of simulations that succeed in realizing a stationary state (here characterized by the fact that players in the population play only issues (x_2, a_1) or (x_1, a_2) at least over the 20 last periods) in populations of size 12 and 16.

Table 12: BOS Game

Number of periods	40	60	80	100	200
% perfect coordination, n=12	68	97.6	99.6	99.8	100
% perfect coordination, n=16	54.6	95.8	99.4	99.8	100

V - Discussion.

V-1- Considering the agent communicative intention.

Simulations with artificial agents are good candidates to study the emergence of a common meaning through repeated interactions because they eliminate at the start problems connected with the existence of psychologically or culturally prominent equilibria and the possibility that a pre-existing common language influences the interpretation of messages. Tomasello (2008, 59) noted that "if we want to understand human communication, we cannot begin by language. Rather we must begin with unconventionalized uncoded communication." Excellent candidates for such unconventionalized communications are natural gestures as pointing and pantomiming. But to understand how they are used to communicate, one must first consider the fact that the information carried by such gestures can only be extracted via an intermediate stage of recognizing the agent communicative intention.

Habit machines, whose motivations are based on internal signals of physiological conditions that just perform whatever act has been reinforced in the presence of current stimulus inputs, can have behavioral plasticity, Sterelny (2003). They can learn by association that one response to an environmental signal is rewarded whereas a different response to that signal is punished. In that sense, in sender-receiver games, artificial agents whose learning process is only pure reinforcement-based do not differ from such habit machines. Senders monitored by an internal signal of type *t* react by a gesture *m* that is rewarded by a payoff $g_s(\mathrm{t}, a)$. Receivers, confronted to such a gesture stimulus, take an action *a* that is rewarded by a payoff $g_r(\mathrm{t}, a)$. In reacting like that, both senders and receivers do not have to possess an instrumental knowledge of a communicative context. It is only when they consider forgone payoffs that artificial agents in a sense change from motivations based on internal signals of physiological condition into motivations based on representations of the external world. To ask what would have been my payoff if, when confronted to gesture *m*, I had chosen another action, is a necessary condition for having access to such a world. To ask what would have been my payoff if, by sending the same message, I had been of another type, is to have knowledge of a causal connection between acts and their consequences. It is the sign of an intentional agency that "involves the formation of world representations functionally decoupled from any specific action, while being potentially relevant to many", Sterelny (2003, 92).

Recognizing a communicative intention translates itself into an important difference when we investigate the efficiency of learning by communication. For the matching game with private information, with a population of size 12, table 13 below compares the percentage of 500 simulations ran over different number of rounds that succeed in realizing an identification of types, either by considering forgone payoffs (EWA, see table 5) or by pure reinforcement-based learning.

Table 13: Matching Game with private information

Number of rounds	40	60	80	100	200
%, EWA	1.4	10.2	28.8	52.8	98
%, pure reinforcement-based	0	0.2	1.8	4	18

For signaling intentions with asymmetric payoffs, using the BOS sender-receiver game previously investigated, the same comparison leads to the results displayed by table 14.

Table 14: BOS Game with communication

Number of rounds	40	60	80	100	200
% ,stationary states, EWA	68	97.6	99.6	99.8	100
% stationary states, pure reinforcement-based	12.2	23	42.4	49.4	78.4

Clearly, considering forgone payoffs leads to better coordination in both cases. But note that, with private information, messages that players send to convey unobservable information about their types is essential for coordination whereas in the BOS game players could learn to coordinate through the observation of past actions that can also signal intentions. It is therefore not surprising to find in Table 15 below that when we run simulations for this game without communication there is a better coordination efficiency without messages. Actions alone speak better than actions plus messages. But in that case, coordination is only realized by the selection of one of the two possible equilibrium issues, all players playing the same issue, either S-issues (x_2, a_1) favorable to row player or R-issues (x_1, a_2) favorable to column player.

Table 15: BOS Game without communication

Number of rounds	40	60	80	100	200
% stationary states, EWA	91.6	99.4	99.8	99.8	100
% stationary states, pure reinforcement-based	25.8	41.6	58.4	71.6	91.6

V-2- Considering the cooperative nature of human communication.

In both the matching game and the BOS game, players have an interest to recognize the communicative and social intentions of senders if they want to avoid disequilibrium issues. In both games, senders have an interest to transmit a truthful information, revealing their type in the matching game or their intention in the BOS game. But when analyzing communication, one must not conflate understanding with acceptance of social intentions. In experiments conducted with a pre-existing common language, understanding the meaning of messages is clear at the start and the problem concerns acceptance. With simulations using artificial agents, the emergence of understanding and acceptance are intimately mixed up from beginning to end of the evolutionary process. It is only in the stationary state of the process that one can try to look at them separately.

In the matching game, interactions are operating in a prior context deprived of any conflict of interest. In this situation of mutual interest, a common meaning emerges because there is a common acceptance to use the senders' social intention to reveal their type as an information source about the world. But such a common meaning is far from having the kind of recursivity that characterizes communicative conventions, Lewis (1969). It is not shared among users in that they all would know that they all would know how to comprehend and use the communicative device for coordinative ends. As a result of a process of communication that does not place individual communicative acts in the public space, it nevertheless succeeds in realizing a perfect coordination.

In the BOS game, the prior context is different. In a context of conflict of interest between senders and receivers, intentional communication can be viewed as requests aimed at individualistic goals in which others are used as social tools. In this case, to understand the social intention of senders is not sufficient to make the receiver's behavior to conform to the wishes of the speaker. Understanding and acceptance may interact to impede the emergence of a common meaning. This is particularly clear when one looks at the different stationary states obtained in the simulations. Some stationary states are of a babbling type, in which both messages associated with a same action are used by senders so that receivers ignore the messages and best response to the senders' action. In such babbling states only one of the two possible equilibria is plaid, either S-issues favoring senders or R-issues favoring receivers, a situation not different from those obtained without communication. However, introducing communication makes a difference because stationary mixed states are also possible.

In such states in which both equilibria are plaid, communication has a particular function. Let us recall the characteristics of the mixed state obtained in IV, graphic 4, in which two senders (sender 2 and 4) always choose m_1 followed by x_2 and four senders always choose m_2 followed by x_1, so that for receivers, m_1 means “I play x_2” and m_2 means “I play x_1”. As we have seen before, such a common receivers' understanding is not known by senders who otherwise would have all sent m_2 and plaid x_2. Because here individual communicative acts are not placed in the public space, it allows senders to use implicitly communication to build personal reputations through the learning process. Some of them have acquired a tough reputation and are identified by receivers by the message m_2, others have acquired a soft reputation and are identified by the message m_1. A situation finally not very different from the situation of the matching game, except that the reputation types are not given at the start but progressively built and not publicly known among senders. Sometimes, all senders may succeed in acquiring a tough reputation (in which case only S-issues $(x_2\ , a_1)$ are plaid) or failed to acquire one (in which case only R-issues $(x_1,\ a_2)$ are plaid). In these two cases, just one message is used, the one through which the common reputation has been built. To measure the importance of reputation building, table 8 below displays for the BOS game with communication and for the group of 500 simulations previously ran the distribution of stationary states.

Table 8: BOS Game with communication

Number of rounds	40	60	80	100	200
% of stationary states, EWA	68	97.6	99.6	99.8	100
% of mixed stationary states	49.7	46.5	45	48.5	47.6
% with freq. of tough reputation >1/2	*(11.1)*	*(9.4)*	*(10.8)*	*(14.2)*	*(12.6)*
% with freq. of tough reputation = 1	1.8	2.2	1.6	2.2	1.2
% with freq. of tough reputation = 0	2	1.5	1.2	1.6	1.8
% of babbling states favorable to senders	34.7	36	39.2	37	35.6
% of babbling states favorable to receivers	11.8	13.8	13	10.7	13.8

In about 50% of stationary states, communication allow senders to build personal reputations. Note that contrary to what has been obtained in the BOS Cooper's experiment with a one-way communication that clearly revealed a first mover advantage for senders (96% of issues are S-issues), no such advantage appears in the simulations. For instance, with 40 rounds and by taking an extensive interpretation of this advantage, only 11.1 + 1.8 + 34.7 = 47.6% are stationary states favorable to senders.

Why are people in the BOS Cooper's experiment so easily accepting a first mover advantage when they communicate with a pre-existing common language? As we have already noted, in experiments conducted with a pre-existing common language, understanding the meaning of messages is clear at the start and the problem concerns acceptance. And in the BOS game, two reasons may be put forward to explain the experimental success of one-way communication. The first one has to do with rationality. As argued by Farrell (1987), it is reasonable for cheap talk announcements to be taken at face value if it will be indeed optimal for the sender to keep his promise if he expects the receiver to belief the message. In this fashion, one-way cheap talk permits sender to select the Nash equilibrium of his choice. The second one has to do with the large amount of common conceptual collaborative ground without which communicative conventions could not have evolved. In such conventions, requests were mainly expressed in contexts of mutual assumptions of helpfulness and in contexts where the role reversal imitation pointed out by Tomasello (2008) - each initiate to the convention understands that she can use the convention toward others as they have used it toward her - contributes to an easy acceptance of a first mover advantage.

References

Anderson C.M. and Camerer C.F. (2000): "Experience-weighted Attraction Learning in Sender-Receiver signaling Games", Economic Theory, 16, 689-718.

Blume A., Kim Y.G., and Sobel J. (1993): "Evolutionary, Stability in Games of Communication", Games and Economic Behavior, 5(4), 547-75.

Blume A., Dejong D., Kim Y.G. and Sprinkle G. (1998): "Experimental Evidence on the Evolution of Meaning of Messages in Sender-Receiver Games, American Economic Review, 88, n°5, 1323-40.

Blume A., Dejong D., Neumann G.R. and Savin S.E. (2002): "Learning and Communication in Sender-receiver Games: An Econometric Investigation", Journal of Applied Econometrics, 17, 225-47.

Blume A. and Arnold T. (2004): "Learning to Communicate in Cheap Talk Games", Games and Economic Behavior", 46, 240-59.

Camerer C.F. and Ho T.H. (1999): "Experience-weighted Attraction Learning in Normal Form Games", Econometrica, 67, 827-74.

Camerer C.F. (2003): *Behavioral Game Theory*, Princeton University Press.

Cheung Y.W. and Friedman D. (1997): "Individual Learning in Normal Form games: Some Laboratory Results", Games and Economic Behavior, 19, 46-76.

Cooper R., Dejong D., Forsythe B. and Ross T. (1990): "Selection Criteria in Coordination Games: Some experimental Results", American Economic Review, 80, 218-33.

Cooper R., Dejong D., Forsythe B. and Ross T. (1994): "Alternative Institutions for Resolving Coordination Problems: Experimental Evidence on Forward Induction and Preplay Communication", in J. Friedman (Ed*.), Problems of Coordination in Economic Activity*, Dordrecht: Kluwer.

Crawford V. (1998): "A Survey of Experiments on Communication via Cheap Talk", Journal of Economic Theory, 78, 286-98.

Farrell J. (1987): "Cheap talk, Coordination and Entry", Rand Journal of Economics, 18, 34-39.

Farrell J. (1988): "Communication, Coordination and Nash Equilibrium", Economic Letters, 27, 209-14.

Farrell J. (1993): "Meaning and Credibility in Cheap-Talk Games", Games and Economic Behavior, 5, 514-531.

Lewis D. (1967): *Convention*, Harvard University Press, Cambridge, Mass.

Myerson R.B. (1991): *Game Theory*, Harvard University Press, Cambridge Mass.

Rabin M. (1990): "Communication between rational Agents", Journal of Economic Theory, 51, 144-70.

Roth A. and Erev I. (1995): "Learning in extensive-form Games: Experimental Data and simple dynamic Models in the intermediate Term", Games and Economic Behavior, 8, 164-212.

Searcy W.A. and Nowicki S. (2005): The Evolution of Animal Communication, Princeton University Press, Princeton.

Sterelny K. (2013): *Thought in a Hostile World*, Blackwell Publishing Ltd.

Straub, P. (1995): "Risk Dominance and Coordination Failures in Static Games", Quarterly Review of Economics and Finance, 35,339-65.

Tomasello M. (2008): *Origins of Human Communication*, The MIT Press, Cambridge, Mass.