Exploring the impact of contextual input on the evolution of word-meaning

Vogt, P.A.; Coumans, H.

Published in:
From animals to animats 7

Publication date:
2002

Exploring the impact of contextual input on the evolution of word-meaning

Paul Vogt and Hans Coumans
IKAT - Universiteit Maastricht
P. O. Box 616 - 6200 MD Maastricht - The Netherlands
p.vogt@cs.unimaas.nl

Abstract

This paper investigates how different types of non-verbal input influence the bootstrapping and evolution of lexicons. This is done by comparing three language game models that differ in the type of input they use. The simulations show that the language games that use either joint attention or corrective feedback as a source of contextual input are better capable of bootstrapping a lexicon than the game without such precise and directed input. The simulation of the latter game, however, does show that it is possible to develop a lexicon without using directed input when the lexicon is transmitted from generation to generation.

1 Introduction

One important question in the scientific field of language evolution and acquisition is: how do language learners acquire the meaning of novel words? A famous example that perfectly illustrates one of the hardest problems these learners face has been introduced by (Quine, 1960). Suppose a linguist is out in the field with a native speaker of a, to him unknown language when suddenly a rabbit passes by. Apparently in response the native speaker says 'gavagai'. A natural reaction of the linguist would be to assume that 'gavagai' means *rabbit*. The native, however, could have referred to a specific rabbit, a part of the rabbit, a mammal or even to the sunny weather. The linguist's assumption is therefore uncertain and he needs more input in order to understand what the native really meant.

When learning a language, a variety of input is exposed to the learner. Important questions are: what kind of input is there? And, is this input functional and/or necessary for language learning? The most obvious form of input are the spoken words. In addition there are often pragmatic of contextual cues to indicate the meaning of spoken words. For young language learners the input usually includes a context that is present in the 'here and now'. Often there is joint attention on the topic of communication and sometimes learners receive corrective feedback on their language use. Many psycholinguists assume that directed contextual input such as joint attention or corrective feedback is necessary to learn word-meanings, but there is some evidence that children do not need such directed cues to learn the meaning of their first words (Lieven, 1994). There is much debate in the literature on the nature of input that language users, and children in particular use to learn the meaning of words, see, e.g., (Bloom, 2000) for a discussion. Similar debates are also prominent in the field of language evolution (Dessalles, 2000, Tomasello, 1999, Kirby and Hurford, 2001).

Recent computational studies on the evolution of language have shown how agents can learn word-meanings successfully, for an overview see, e.g., (Cangelosi and Parisi, 2001). In most of these studies it has been assumed that either joint attention was established, cf. (Oliphant, 1999) or that agents receive corrective feedback, cf. (Steels and Kaplan, 1999). These two conditions have been modeled on real robots using variants of the language game model originally introduced in (Steels, 1996a). The joint attention condition has been modeled with *observational games* (Vogt, 2002a) and the corrective feedback condition with *guessing games* (Steels and Kaplan, 1999, Vogt, 2002b).

Few studies have investigated whether cues such as joint attention and corrective feedback are really necessary. Smith has shown in simulations how two agents can develop a shared lexicon without any of the mentioned cues, as long as the objects to which words refer are 'visible' in some context (Smith, 2001). Robotic experiments that use such a strategy (called *selfish games*), however, revealed results equal to chance (Vogt, 2002a). But these results were obtained with an experimental setup that is very minimal in terms of the environmental complexity and the robots' architecture. In Smith's simulations the environment is more complex, because for every interaction a context of randomly generated objects is constructed. It is therefore likely that Smith's results are more realistic.

In this paper three research questions are addressed. (1) How do the three language game models relate to each other? (2) Are the models scalable in terms of population size? And (3) can the selfish game deal with a population dynamics? The first question is investigated by comparing the three language game models with each other under the different conditions that are used to answer the other questions. The second question is addressed in simulations that have larger populations than in the previous experiments. The last question is investigated by modeling a flow of agents using the *iterative learning model* (Kirby and Hurford, 2001). As the last two questions have already been answered affirmative for the observational and guessing games (Oliphant, 1999, Steels and Kaplan, 1999), the focus for these questions will be on the selfish games for which these questions have not been addressed before.

In contrast to some previous work, e.g, (Vogt, 2001), all studies in this paper are done in simulations where the meanings are predefined and the interactions are modeled without noise. As a consequence the symbols these agents develop are Saussurean - i.e. they are word-meaning associations - and not grounded in reality. Furthermore, the research in this paper is not concerned with the question how joint attention or corrective feedback can be achieved.

The remainder of this paper is organized as follows: The next section presents some psycholinguist background and some relevant computational studies relating to the problem statement. The three language game models are presented in section 3. Section 4 presents the results of the simulations, which are discussed in section 5. Section 6 presents conclusions.

2 Learning the meaning of words

When children learn the meaning of words, they tend to learn some types of words earlier than other types. Children usually acquire the meaning of basic-level categories such as *dog* before they acquire superordinate and subordinate categories such as *animal* and *terrier*. This can be explained by the observation that basic-level names occur more frequently in the input (Brown, 1958) and because children "find it more natural to categorize a novel object as an instance of a basic-level kind than as an instance of a superordinate kind. As a result, children (and adults) find it easier to learn novel basic-level names than novel superordinates" (Bloom, 2000, p. 90).

It is widely assumed that most words are learned by associative learning, i.e. words are associated with the meaning of referents that are simultaneously presented. This requires joint attention on the referent that can be established, e.g., by checking, following or directing an adult's attention (Tomasello, 1999). Many parents, for instance, teach their children novel words by holding something in front of the child and giving it a name ('look, it's a "toma"'). In some Eastern cultures, however, adults in a child's environment do not speak directly to their children until they used at least a few words meaningfully (Lieven, 1994). These adults only speak with each other in front of children, who start speaking at some moment and therefore must grasp the meaning of words by observing adult speech without receiving clear cues concerning its meaning. These children learn somewhat slower than those raised in Western cultures (Lieven, 1994), which is consistent with the finding that children, who use joint attention learn faster than those who do not (Tomasselo and Todd, 1983).

An alternative for associative learning is reinforcement learning, in which children receive corrective feedback on their language use. This means that when a child uses a word for the first time (or perhaps somewhat later), it receives positive feedback when it uses the word properly; otherwise it receives negative feedback. Word-meanings that receive positive rewards tend to be used more often than those that receive negative feedback. Again such learning strategies are widely observed in Western cultures, but not in some Eastern cultures (Lieven, 1994).

Most computational studies on the origin of word-meanings use the Western strategies. Associative learning, for instance, is used by (Oliphant, 1999, Billard and Dautenhahn, 1999), while reinforcement learning is used by, e.g., (Steels and Kaplan, 1999, Yanco and Stein, 1993). Both learning types have also been implemented on real mobile robots as observational games and as guessing games (Vogt, 2001). The first type of game models associative learning and the second reinforcement learning. The experimental results show that agents using either strategy can develop a coherent lexicon rather well. Comparing the two games reveal small differences in the development of a grounded lexicon. The success rate of the observational game experiments increased and stabilized faster than for the guessing game, although for both experiments the level of success approached approximately the same level. The amount of polysemy and synonymy[1] emerged from the guessing games, however, was substantially lower than from the observational games. Thus the lexicons developed with the guessing games are more informative. For a detailed discussion of this comparison consult (Vogt, 2001).

Models that try to explain the origins of language - or language acquisition - must also be able to explain learning word-meanings using the non-Western strategy. To study this, the *selfish game* has been developed (Vogt, 2002a). In the selfish game, neither joint attention nor corrective feedback is used.[2] Agents in a selfish game observe a context of several objects (or meanings). The speaker of the game selects a topic and tries to name it. The hearer has to guess what the name refers to, but

[1] Polysemy is the association of one word with several meanings and synonymy is the association of one meaning with many words.

[2] As the agents in this game do not explicitly 'care' whether they communicate about the same meaning, they behave a kind of selfish.

has no means to verify whether it was successful. When agents are exposed to different contexts in which one object always re-occurs together with a particular name, they may infer that the name refers to this re-occurring object. Learning such knowledge has been modeled with a *Bayesian learning* technique (Mitchell, 1997). In a way this learning mechanism is similar to associative learning, but deals better with uncertain relations that are present in the selfish games.

Robotic experiments with the selfish game were disappointing as they revealed results equal to chance (Vogt, 2002a). The reason for this has to do with the use of robots which are minimal in terms of their physical architecture that reveals poor sensorimotor skills and that operate in a limited environment containing only three to four objects. In the selfish game, the robots calculate within the context of a game the probability of the occurrence of a meaning, given the occurrence of some word, based on previous games. The association that has the highest probability determines the meaning of the word. So if all contexts would have the same meaning, the distribution of this probability would be flat and a word's meaning is highly uncertain. When the context varies sufficiently, the distribution will bring forward word-meaning associations that co-occurred most often. In the experiments some variation was forced by removing and adding one of the four objects dynamically, but this appeared to be insufficient. That the principle of the selfish game works has been shown in simulations (Smith, 2001), so the negative results of the robotic experiments are perhaps only valid for the minimal setup. While a great number of experiments have been done with the guessing games and the observational games - or other similar models, the selfish game is rather unexplored. Both Smith and Vogt have only reported on one or two experiments with a population size of 2, so more experiments are required to test, for instance, the scalability and population dynamics of the selfish game.

One important question that is investigated in this paper is how the three different language games compare to each other. The simulations will be compared in terms of success and in terms of information, i.e. which game performs better in communication and which one reveals the most informative lexicon?

3 Language games

The simulations reported in this paper are all primarily based on the language game model introduced in (Steels, 1996a). In the language game experiments, a population of agents develop a shared lexicon of word-meaning associations from scratch by means of cultural interactions, individual adaptation and self-organization. Many successful experiments have been conducted both in simulations (Steels, 1996a, De Jong, 2000) and on physical robots such as the Talking Heads (Steels and Kaplan, 1999) and small LEGO vehicles (Vogt, 2002b, Vogt, 2002a).

Each agent in the population has a private lexicon. The lexicon is a set of word-meaning associations where each entry has a score σ indicating the effectiveness of the association. The words are constructed from arbitrary strings of consonants and vowels and meanings are given and represented by integers in the simulations. The association scores σ are real values between 0 and 1. At the start of each experiment, the agents have empty lexicons and non-empty ontologies that contain all meanings.

The remainder of this section explains each type of language game. The iterative learning model is explained in the final part of this section and applies to all types of language games.

3.1 The observational game

The observational game uses joint attention and to enable associative Hebbian learning. The game is organized as follows:

1. Two agents are randomly selected from the population. Arbitrarily, one agent is assigned the role of the speaker, the other is the hearer.

2. The speaker selects randomly one meaning as the *topic* from the ontology and informs the hearer what the topic is, thus establishing joint attention.

3. The speaker searches its lexicon for words that are associated with this topic and selects the association that has the highest score σ. If the speaker fails to find a matching association, it invents a new word and adds the word-meaning association to the lexicon with an initial association score of $\sigma = 0.01$. The word is communicated to the hearer.

4. The hearer searches its own lexicon for an association of which the word matches the received word *and* of which the meaning corresponds to the topic.

5. If the hearer succeeds in finding a proper association, the observational game is a success. Otherwise it fails. The outcome is known to both agents.

6. Depending on the outcome the lexicon is adapted as follows:

    (a) If the game is a failure, the hearer adopts the word and adds the word-meaning association to its lexicon. The speaker lowers the used association score by $\sigma := \eta \cdot \sigma$, where $\eta = 0.9$ is a constant learning rate.

    (b) If the game is a success, both robots increase the association score of the used association by $\sigma := \eta \cdot \sigma + 1 - \eta$ and they laterally inhibit all competing associations by $\sigma := \eta \cdot \sigma$. An association is competing when either its meaning is the same as the topic, but its word differs from the uttered word, or when the word is the same, but not its meaning.

3.2 The guessing game

In the guessing game, the non-verbal cue is given by evaluating the correctness of the game and the lexicon is learned by reinforcement learning. This game differs slightly from the observational game and goes as follows:

1. Two agents are randomly selected from the population. One is assigned the role of the speaker, the other is the hearer.
2. Both agents establish a context of a limited size that contains meanings from the ontology. Both agents share the same context.
3. The speaker selects one meaning from the context as the topic, but it does not inform the hearer about this.
4. The speaker searches its lexicon for words that are associated with this topic and selects the association with the highest score. If the speaker fails to find a matching association, it creates a new word and adds its association with the topic to the lexicon. The word is communicated to the hearer
5. The hearer searches its lexicon for an association of which the word matches the received word *and* of which the meaning corresponds to one of the meanings in its context. If there are one or more matching associations, the hearer selects the association that has the highest association score. The corresponding meaning becomes the hearer's topic.
6. If the hearer succeeds in finding an association, the guessing game is a success when its topic matches the speaker's topic. Otherwise it fails because either there is a mismatch in the topic or the hearer does not know the uttered word in relation to the context. The verification of the outcome implements the corrective feedback, which is known to both agents.
7. Depending on the outcome of the game, the lexicon is adapted as follows:
   (a) If there is a mismatch in the topic, the hearer adopts the word and associates it with an arbitrary meaning from the context that is not yet associated with this word. In addition, both agents lower the association score by $\sigma := \eta \cdot \sigma$.
   (b) If the hearer does not know the word, it adopts the word and associates it with the topic as in the observational game. The speaker lowers the used association score as above.
   (c) If the game is a success, both agents increase the association score of the used association and they laterally inhibit all competing associations as in the observational game.

3.3 The selfish game

In the selfish game there is no non-verbal input to indicate exactly the topic of a game. The only input is a context that contains a number of meanings and the speaker's utterance. As in the selfish game the agents have no means to verify whether their communication was successful, they cannot use the association score as an indication of the effectiveness of a lexical element. The only information the hearers have with respect to the meaning of an utterance are the meanings in the context. As the context may consist of more than one meaning[3], the meaning of an utterance is uncertain for the hearer. When the contexts vary sufficiently from game to game, the cross-section of these contexts in co-occurrence with a particular word indicates the meaning of this word. Learning this relation can be done using a *Bayesian learner*. For this the association score is now given by the following equation:

$$\sigma = P(m \mid w) = \frac{P(w \mid m)P(m)}{P(w)} = \frac{P(m \wedge w)}{P(w)}$$

In this equation $P(m \mid w)$ is the conditional probability that given a word w the meaning m can be expected. Using Bayes' law, this can be translated into the quotient between the probability that m co-occurs with w and the probability of w. Note that this way the association score can be calculated in the same way as the *confidence probability* used in (Smith, 2001) by $\sigma = \frac{U_{wm}}{U_w}$, where U_{wm} is the co-occurrence frequency of w with m and U_w is the occurrence frequency of w.

Applying this new association score to the selfish game leads to the following algorithm:

1. to 5. are identical to the guessing game.[4]

6. Instead of evaluating the game's success, the agents adapt their lexicon immediately as follows:
   (a) The hearer assures that the word is associated with all meanings in the context. In addition the hearer increments the co-occurrence frequency U_{wm} by 1 for all meanings in the context and increases the occurrence frequency U_w with the context size, i.e. it increments U_w by 1 for all meanings in the context.
   (b) The speaker increments both U_{wm} and U_w by 1 for the topic.

[3] In the current paper, the context is always larger than one.

[4] When the speaker invents a new association (point 4), U_{wm} and U_w are initialized with 1.

The selfish game implemented here is very similar to the implementation of (Smith, 2001), except that his agents use discrimination games (Steels, 1996b) to acquire meanings and that they use *obverter* learning (Oliphant and Batali, 1997). The principle of obverter learning, where the speaker selects a word-meaning by pretending he is a hearer, however, is adopted by the Bayesian learner that evaluates the probability of the topic, given the occurrence of a word. In obverter learning speakers may select words of which the meaning does not match the topic. In such cases no utterances are made and new words must be invented to cover the topic. In the current application, words are always selected in relation to the topic, even if the word is better understood for a different meaning.

3.4 The iterative learning model

The iterative learning model (ILM) implements a population dynamics and the transmission of linguistic knowledge over generations (Kirby and Hurford, 2001). It can be applied to all kinds of language games. In the ILM there are two types of agents in the population: adults and learners. Adults have passed the stage of learners and are assumed to have mastered the language. Learners enter the population as novices and learn the language from the adults. In the current simulations adults only take the role of speakers, while learners only take the role of hearers. As a result, the adults are the only agents that can invent new symbols, while the learners can only adopt words expressed by the adults.

Assuming that the sets of learners and adults are initially empty, the ILM is as follows:

1. Start with an initial population of N equal agents that are neither learners nor adults. This way each agent can take the role of speaker and hearer.
2. A series of X observational, guessing or selfish games are played, where the type of game depends on the experiment. This way the lexicon is bootstrapped.
3. When finished, set the initial population to the set of learners.
4. Iterate the following:
   (a) Remove all adults and replace them by the set of learners.
   (b) Add N new 'empty' agents to the set of learners.
   (c) Play a series of X observational, guessing or selfish games.

Note that the first time step 4 is reached, the set of adults is empty and is equalized by the initial population in 4 (a) after which the population is doubled in step 4 (b).

4 Experimental results

With the three models, a number of simulations have been done. Each simulation has been repeated 10 times with different random seeds. In the simulations all agents had identical ontologies of 100 meanings and the context size[5] in each game was fixed at 5. The three conditions that have been investigated are:

Exp. 1: Simulation of the games with a population size of 2.

Exp. 2: Simulation of the games with a population size of 5.

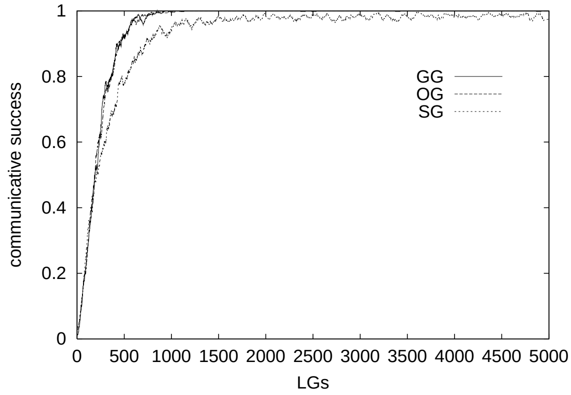
Exp. 3: Simulation of the games using the ILM and with a population size of **(a)** 4 and **(b)** 8.

Exp. 1 In figure 1 (a) the communicative success of the simulations with a population size of 2 for the three types of games is shown. The simulations have been run for 5,000 language games. The communicative success is the number of successful games averaged over the past 50 games, and over the 10 different runs. It can be seen that both the observational and the guessing game reach a communicative success of 1 rather fast. The selfish game, however, never reaches a success of exactly 1, it stabilizes around 0.95 from approximately game 1,500.

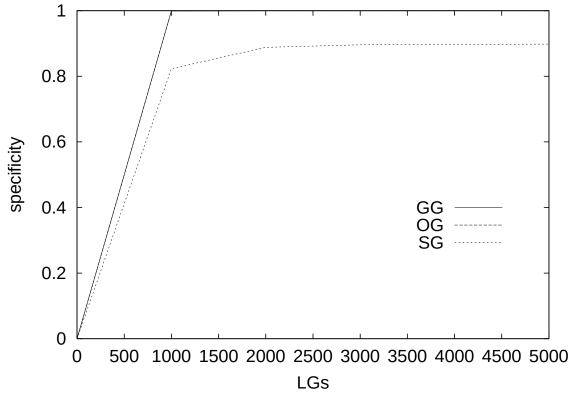
Figure 1 also shows (b) the *specificity* and (c) *consistency*, which are calculated from the mutual information of meanings and words relative to the uncertainty of meanings in case of specificity and relative to the uncertainty of words for the consistency. The uncertainty, like the mutual information is calculated using the entropy measures introduced by (Shannon, 1948). The two measures are developed by (De Jong, 2000) to indicate how specific or consistent the agents' lexicons are used. A lexicon is specific if each word use predicts its meaning perfectly; it is consistent when each meaning is named with one word. Thus the specificity indicates the amount of polysemy in the lexicon and the consistency the amount of synonymy. For details how to calculate the specificity and consistency consult, e.g., (De Jong, 2000, Vogt, 2002b). Although the measures are originally calculated from the agents' viewpoint, in this paper it is calculated for the global lexicon use. They are calculated over each 1,000 games.

Figure 1 (b) shows that the specificity increases to 1 for the observational and guessing game, but increases to approximately 0.9 for the selfish game. This means that in the selfish game there is some polysemy, which is not true for the other two games. The consistency too increases to 1 for the observational and guessing

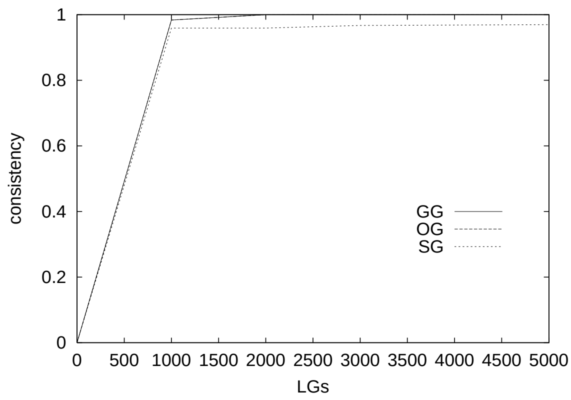
[5] The context size only applies to the guessing game and the selfish game.



(a) Communicative success



(b) Specificity



(c) Consistency

Figure 1: The results of **exp. 1**. Each figure shows the results of the guessing game (GG), observational game (OG) and the selfish game (SG). The figures show the various measures as a function of the language games (LGs) played.

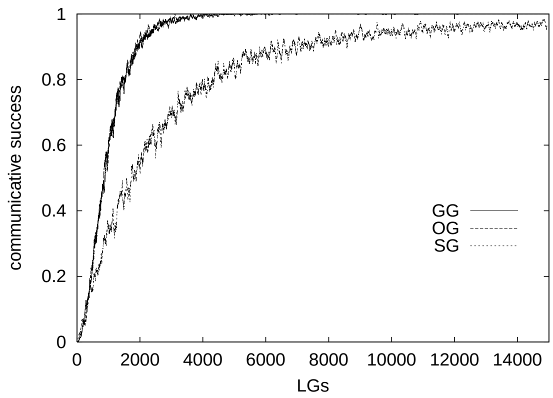
game, while for the selfish game it increases to approximately 0.95. This result indicates that for the selfish game there is also some synonymy, which is not present in the two other games. This is confirmed by the *coherence*, which measures more directly the level of synonymy as it calculates the average number of agents of that use the same words for all meanings (Steels, 1996a). The coherence (not shown here) reaches 1 for the observational and guessing games slightly after the communicative success reaches 1. It stabilizes near 0.8 from game 3,000 in the selfish games. The results of **exp. 1** are consistent with the results reported by (Oliphant, 1999, Steels and Kaplan, 1999, Smith, 2001).

Exp. 2 When the number of agents is increased, the results are rather different (figure 2). In these simulations 15,000 language games have been played. Although it takes somewhat longer, both the guessing game and the observational game converge to 1 for the communicative success, which increases to a value slightly below 1 for the selfish game. The specificity evolves more or less the same as in **exp. 1**, but the consistency is lower. For the observational and guessing game, the consistency reaches a value of 1 only near the end of the experiment. The consistency of the selfish game evolves to a value close to 0.8, which means that there emerges more synonymy when the population size increases. An observation that is also confirmed by the coherence of the lexicon, which stabilizes but does not exceed 0.2 for the selfish game, even when the simulation is run for 50,000 games. This value indicates that each agent uses its own words to name the various meanings, but that they can nevertheless understand each other - otherwise the communicative success would not approach 1. Similar results were obtained for all tested population sizes from 3 to 20, but there is no linear dependency between the population size and the coherence.

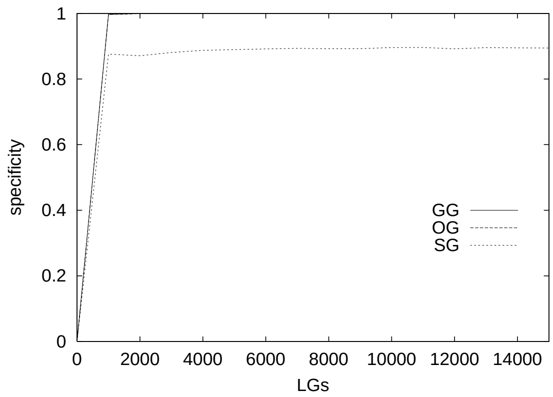
Exp. 3 (a) Up to now the population has been kept fixed in each simulation. Figure 3 shows what happens when the ILM is applied with an initial population size of 2, so that the population size becomes 4 after the first iteration. The ILM has been run for 8 iterations of 4,000 games each.

Figure 3 (a) shows that in each iteration of the communicative success rapidly returns to its old values, which are 1 for the observational and guessing games and approximately 0.95 for the selfish game. It is hard to see in the figure, but the performance improves slightly in each iteration, except for the second iteration where the population size becomes double the size of iteration 1.

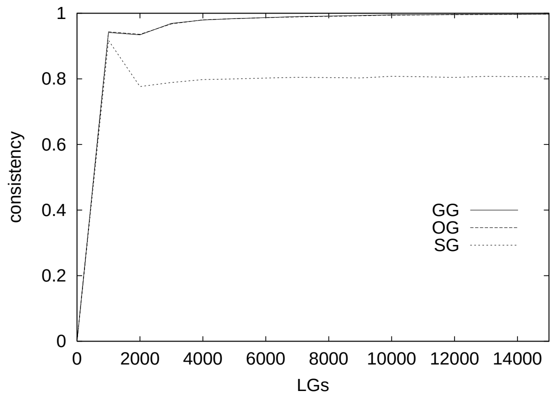
Figures 3 (b) and (c) show that the specificity and consistency have a similar evolution as the previous experiments. They do not drop after each iteration because these values do not measure success, but the quality of the lexicon use by speakers, which is not affected by the



(a) Communicative success

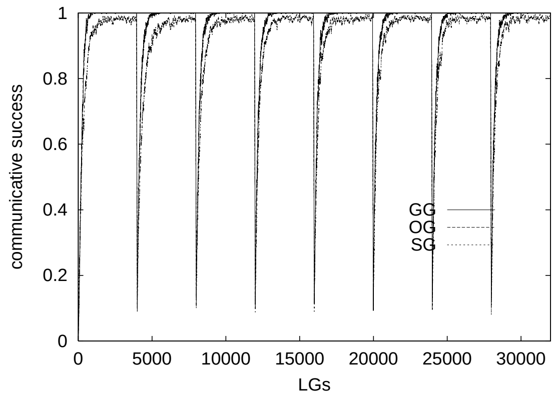


(b) Specificity

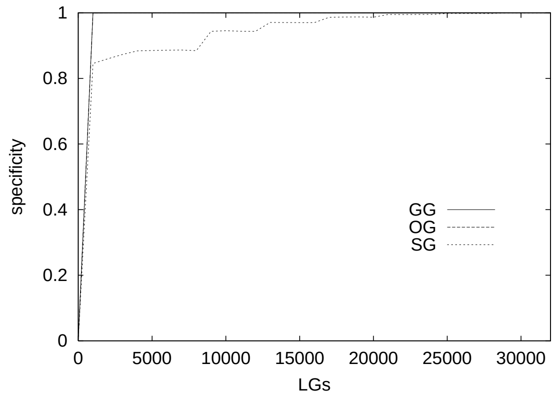


(c) Consistency

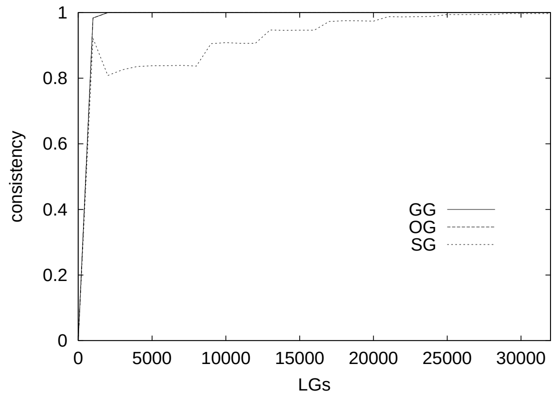
Figure 2: The results of **exp. 2**, which are simulations with a population size of 5.



(a) Communicative success

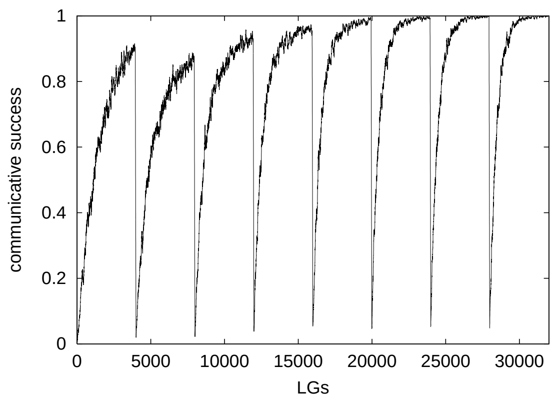


(b) Specificity

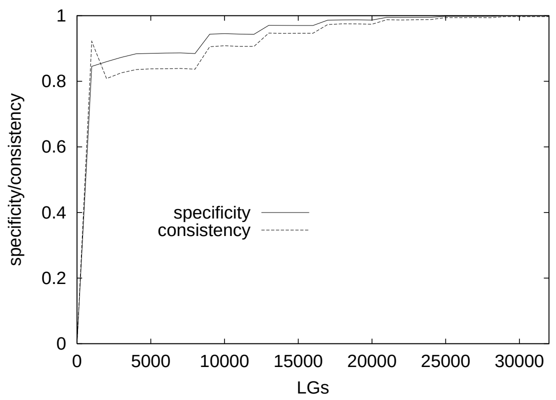


(c) Consistency

Figure 3: Results of **exp. 3 (a)**: applying the ILM to the simulations with an initial population size of 2.



(a) Communicative success



(b) Specificity / consistency

Figure 4: The results of **exp. 3 (b)**: applying the ILM to the selfish with a population size of 4. The figure shows (a) the communicative success and (b) the specificity and consistency.

introduction of novel learners. The specificity and consistency of the observational and guessing games reach 1 after 2,000 games in the first iteration. For the selfish game these measures approach 1 in the last iteration, which is also true for the coherence.

Exp. 3 (b) Figure 4 (a) shows the results of applying the ILM to the selfish game with a population size of 4, hence with a population size of 8 after the first iteration of 4,000 games. The two other games are not shown here as they evolve in a similar way as in **exp. 3 (a)**. The communicative success approaches 1 at the end of iteration 5. The specificity and consistency too show an evolution from relative low values (between 0.8 and 0.85) to values near 1 in the eighth iteration. The coherence does not reach 1 after 8 iterations, but is still increasing and it may be expected that a value of 1 is reached in few more iterations. These results indicate that - even from more difficult conditions, i.e. an initial population size of 4 - the agents agents benefit from the dynamical flow induced by the ILM.

5 Discussion

The first experiment confirm the findings reported earlier by (Steels and Kaplan, 1999, Oliphant, 1999, Smith, 2001). When comparing these results with the robotic experiments reported in (Vogt, 2001), the observational and the guessing game yield similar results, although the robotic experiments revealed small differences between the two models, which is not the case in the simulations. This has most likely has to do with the 'perfect' conditions of simulations that are not met in the robotic experiments. That the selfish game in the simulations outperforms the robotic experiments can be explained by the limited environment used in the robotic experiments (Vogt, 2002a). As a result the robots obtained little variation in the context, which is very important for the Bayesian learner.

The observational and guessing games seem unaffected by increasing the population size (**exp. 2**), but the selfish game performs worse as the global lexicon shows more uncertainty by the lower consistency. It seems as though each agent develops its own lexicon which the other agents understand, but they do not converge in their production. This convergence does occur when the iterative learning model is applied to the selfish game (**exp. 3**). Because in the ILM there are only adult speakers, the results suggest that the strategy of the selfish game only works for larger populations when the learners learn from experienced speakers. The learners are only exposed to the lexicon use of adult speakers and do not invent parts of the lexicon themselves. Thus they observe a relatively stable lexicon, which stabilizes even more while they are learning. So, when they become adults, they can transfer their improved lexicon to the next generation and the cycle continues.

It is interesting to see that the selfish games converge much slower than the other two games. This has much to do with the uncertain nature of the input. The hearer has to guess what the speaker is referring to, while it cannot verify whether it guessed right. It may take quite some games before the hearer has enough information to make the decisions. This is consistent with the slower learning in the Eastern cultures (Lieven, 1994) and with the observation that children learn word-meanings faster when joint attention is established than when not (Tomasselo and Todd, 1983). Moreover, many studies reveal a fast mapping phenomenon, which is the observation that the meaning of many novel words are learned within one or two exposures (Carey, 1978). The results of the simulations suggest that fast mapping can only occur when the observational or guessing game models are used, because learning in the selfish game model is too slow.

Another reason why the selfish game converges slower and emerges less informative lexicons is that the selfish game cannot use lateral inhibition within the scope of a context, it only inhibits associations that are outside in the context. Previous experiments have shown that lateral inhibition is a crucial factor for lexicon development, e.g., (Oliphant, 1999, De Jong, 2000). It also implicitly implements the notion of *lexical contrast* (Clark, 1993), which is the principle that language learners have a bias that words have only one meaning and vice versa. Lateral inhibition models the mechanism that one word-meaning association is preferred, while others are pushed to the background. It can, however, only work when an agent has more precise knowledge, which requires input modeled with joint attention or corrective feedback.

The selfish game nevertheless explains why children raised in some Eastern cultures do not need joint attention or corrective feedback to learn word-meanings (Lieven, 1994). But the child requires input from experienced speakers that expose a consistent language use as the experiments with the ILM reveal. The selfish game, however, is not a likely mechanism for all language learning or for bootstrapping symbolic communication. The results of **exp. 2** might indicate why non-human primates have not evolved symbolic communication spontaneously, i.e. in their natural habitats. As, for instance, chimpanzees do not engage in joint attention behaviors (Tomasello, 1999), they are bounded to use the selfish game strategy. This strategy is slow and symbolic communication that chimpanzees may try out will be ineffective at first and is therefore less likely to be picked up by conspecifics. When non-human primates do engage in joint attention behavior, which occurs in experimental settings controlled by humans, they may indeed acquire symbolic communication. Moreover, it has been observed that an infant bonobo chimpanzee learned symbolic communication from his mother that has been trained for communication (Savage-Rumbaugh et al., 1986). So this occurred once infant bonobos could observe the communication behaviors of an adult 'speaker', which is quite well possible with the selfish game.

As for humans all three studied strategies are potential candidates to explain what input is used by infants to learn word-meanings, it might be the case that humans use a mixed strategy to learn words. Reconsider the linguist faced with the 'gavagai' problem. The (joint) attention of the linguist is focused on the passing rabbit, because the speaker might have looked at the rabbit and the linguist followed the speaker's eye gaze. As there is a tendency that novel words are first associated with whole-objects and most naturally with basic-level categories (Bloom, 2000), the linguist associates 'gavagai' with *rabbit*. But now suppose that 'gavagai' means *large ears*. If the linguist does not interrogate the speaker about the meaning of 'gavagai', he can only find out when he hears the word 'gavagai' in a different context where large ears are present. In that case negative feedback may lower his certainty about 'gavagai' meaning *rabbit* and form new hypotheses about its meaning. If his attention is not completely focused on the large ears, i.e. the context may still be uncertain, the linguist can only make uncertain assumptions. When he is exposed more often with the word in a context where large ears are present, he may finally learn the meaning of 'gavagai'.

6 Conclusion

The three types of language games that are investigated in this paper all work rather well, although the observational and guessing games that use precise contextual input converges faster than the selfish game which does not use such precise cues.

The selfish game that uses a Bayesian learning technique performs worse than the other two games when no population dynamics is used. When there are two agents, the lexicon that emerges has some ambiguities, but the results are similar to those reported by (Smith, 2001). When the population size increases the results are much worse in terms of mutual information between words and meanings, i.e. the lexicons become more ambiguous. Only when a population dynamics is incorporated, modeled with the iterative learning model (Kirby and Hurford, 2001), agents can develop a highly informative lexicon after a few generations, even with relative large populations.

The results suggest that it is not unlikely that all types of input can be used to learn language, which is conform empirical data from the psycholinguistic literature. While the use of joint attention or corrective feedback appears to be necessary to explain the phenomenon of fast mapping (Carey, 1978), the absence of such input in some Eastern cultures (Lieven, 1994) can be explained by the selfish game. The results additionally suggest that joint attention or corrective feedback are necessary to bootstrap symbolic communication, conform (Dessalles, 2000, Tomasello, 1999).

Future research on language evolution and acquisition - either in simulations or robotic experiments - should take a mixture of the investigated, and possibly more strategies into account. Furthermore research is required to investigate how joint attention and corrective feedback can emerge in grounded experiments and to further scale up the selfish games.

Acknowledgements

The authors wish to thank Floris Wiesman for providing critical comments on an earlier version of this paper.

References

Billard, A. and Dautenhahn, K. (1999). Experiments in social robotics: grounding and use of communication in autonomous agents. *Adaptive Behavior*, 7(3-4).

Bloom, P. (2000). *How Children Learn the Meanings of Words*. The MIT Press, Cambridge, MA. and London, UK.

Brown, R. (1958). How shall a thing be called? *Psychological Review*, 65:14–21.

Cangelosi, A. and Parisi, D., (Eds.) (2001). *Simulating the Evolution of Language*. Springer, London.

Carey, S. (1978). The child as word-learner. In Halle, M., Bresnan, J., and Miller, G. A., (Eds.), *Linguistic theory and psychological reality*, Cambridge, MA. MIT Press.

Clark, E. V. (1993). *The lexicon in acquisition*. Cambridge University Press.

De Jong, E. D. (2000). *The Development of Communication*. PhD thesis, Vrije Universiteit Brussel.

Dessalles, J.-L. (2000). Two stages in the evolution of language use. In Dessalles, J.-L. and Ghadakpour, L., (Eds.), *The evolution of language: 3rd International Conference, Paris 2000*, Paris. ENST.

Kirby, S. and Hurford, J. R. (2001). The emergence of linguistic structure: An overview of the iterated learning model. In A., C. and D., P., (Eds.), *Simulating the Evolution of Language*, London. Springer.

Lieven, E. V. M. (1994). Crosslinguistic and crosscultural aspects of language addressed to children. In Gallaway, C. and Richards, B. J., (Eds.), *Input and interaction in language acquisition*, cambridge. Cambridge University Press.

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.

Oliphant, M. (1999). The learning barrier: Moving from innate to learned systems of communication. *Adaptive Behavior*, 7 (3-4):371–384.

Oliphant, M. and Batali, J. (1997). Learning and the emergence of coordinated communication. *Center for Research on Language Newsletter*, 11(1).

Quine, W. V. O. (1960). *Word and object*. Cambridge University Press.

Savage-Rumbaugh, E., McDonald, K., Sevcik, R. amd Hopkins, W., and Rubert, E. (1986). Spontaneous symbol acquisition and communicative use by pygmy chimpanzees (*pan paniscus*). *Journal of Experimental Psychology: General*, 115:211–235.

Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 623–656.

Smith, A. D. M. (2001). Establishing communication systems without explicit meaning transmission. In Kelemen, J. and Sosík, P., (Eds.), *Proceeding of the 6th European Conference on Artificial Life, ECAL 2001*, LNAI 2159, pages 381–390, Berlin Heidelberg. Springer-Verlag.

Steels, L. (1996a). Emergent adaptive lexicons. In Maes, P., (Ed.), *From Animals to Animats 4: Proceedings of the Fourth International Conference On Simulating Adaptive Behavior*, Cambridge Ma. The MIT Press.

Steels, L. (1996b). Perceptually grounded meaning creation. In Tokoro, M., (Ed.), *Proceedings of the International Conference on Multi-Agent Systems*, Menlo Park Ca. AAAI Press.

Steels, L. and Kaplan, F. (1999). Situated grounded word semantics. In *Proceedings of IJCAI 99*. Morgan Kaufmann.

Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press.

Tomasselo, M. and Todd, J. (1983). Joint attention and lexical acquisition style. *First Language*, 4:197–212.

Vogt, P. (2001). The impact of non-verbal communication on lexicon formation. In *Proceedings of the Belgian/Netherlands Artificial Intelligence Conference, BNAIC'01*.

Vogt, P. (2002a). Evolution of grounded symbols in communicating robots. *Evolution of Communication*. to appear.

Vogt, P. (2002b). The physical symbol grounding problem. *Cognitive Systems Research Journal, Special Issue on Embodied Cognition*. to appear.

Yanco, H. and Stein, L. (1993). An adaptive communication protocol for cooperating mobile robots. In Meyer, J.-A., Roitblat, H. L., and Wilson, S., (Eds.), *From Animals to Animats 2. Proceedings of the Second International Conference on Simulation of Adaptive Behavior*, pages 478–485, Cambridge Ma. The MIT Press.