Evolution of Linguistic Diversity in a Simple Communication System

Takaya Arita
Yuhji Koyama
Graduate School of
Human Informatics
Nagoya University
Furo-cho, Chikusa-ku
Nagoya 464-8601
JAPAN
ari@info.human.nagoya-u.ac.jp
koyama@shiro.gs.human.nagoya-u.ac.jp

Abstract This article reports on the current state of our efforts to shed light on the origin and evolution of linguistic diversity using synthetic modeling and artificial life techniques. We construct a simple abstract model of a communication system that has been designed with regard to referential signaling in nonhuman animals. We analyze the evolutionary dynamics of vocabulary sharing based on these experiments. The results show that mutation rates, population size, and resource restrictions define the classes of vocabulary sharing. We also see a dynamic equilibrium, where two states, a state with one dominant shared word and a state with several dominant shared words, take turns appearing. We incorporate the idea of the abstract model into a more concrete situation and present an agent-based model to verify the results of the abstract model and to examine the possibility of using linguistic diversity in the field of distributed AI and robotics. It has been shown that the evolution of linguistic diversity in vocabulary sharing will support cooperative behavior in a population of agents.

Keywords
evolution, linguistic diversity, communication, genetic algorithms

1 Introduction

Chomsky's famous claim that from a Martian's-eye-view all humans speak a single language is surely plausible. However, in our view it is true that we have thousands of mutually unintelligible languages. Terrestrial scientists have no conclusive answer as to why this linguistic diversity exists [12]. While the quest for the origin of diversity in languages is a challenging theme, diversity in species is also one of the most important themes in biology. Charles Darwin stressed the importance of language difference and linked the evolution of languages to biology [5].

The study of communication/language from an alife perspective has received a great deal of attention lately [14]. Some of the first experiments were conducted by MacLennan [10] and Werner and Dyer [16]. MacLennan considered a population of simple organisms, represented genetically by truth tables, and created a shared environment through which the organisms could pass initially arbitrary signals. It was observed that effective communication evolved in the population based on their scoring function. The simulation experiment by Werner and Dyer successfully demonstrated the evolution of a system for signaling between members of opposite sexes to coordinate mating behavior. In their model, explicit scoring functions were not used; instead, effective communication allowed males to find females more rapidly, thus increasing the reproductive rate of the individuals that communicated effectively.

Concerning the evolution of grammar, Batali [4] constructed a model for the evolution of grammar and performed the simulations of evolution on populations of simple recurrent networks in which the selection criterion was the ability of the networks to recognize strings generated by grammars. The results suggest a new explanation for the "critical period" effects observed in language acquisition. Hashimoto and Ikegami [6] studied the evolution of grammar systems in networks using an agent model. Here, the individual grammar was expressed by a symbolic generative grammar, and each agent was ranked explicitly by three scores in each round: speaking, recognizing, and being recognized. It was observed that two processes, a module-type evolution and a loop-forming evolution, were significant. The number of recognized words rapidly increased when a module emerged in a grammar system, and many words could be derived recursively by a grammar processing a loop structure.

There have not been many studies concentrating on the issue of the linguistic diversity from an evolutionary perspective. Werner and Dyer [16] showed that "dialects" that are bilingual (i.e., correctly interpreting several signaling protocols) have an increased chance of dominating over time. Also, Hashimoto and Ikegami [6] studied the diversity of spoken words produced by symbolic grammar systems in terms of the computational ability of automata, where their computational ability was the ratio of recognizable words to the total number of possible words.

The most straightforward explanation for the origin of linguistic diversity is based on spatial distribution of individuals [3]. The following two studies have supported this view. Arita and Taylor [2] constructed a simple communication model in which a population of artificial organisms with neural networks inhabited a lattice plane and each organism communicated information with neighbors by uttering words. The results of the experiments showed that the accumulation of mutation, propagation delay, and the effects of inheritance produce very complex dynamics, while learning by neural networks and selection of parents have large effects on language unification. Through their experiments on naming games, Steels and McIntyre [15] showed that agent interaction, which depends on spatial distribution, determines the degree of diversity in vocabulary. Their research takes the view that linguistic information evolves and is transmitted culturally, not genetically.

There have been other explanations of the origin of linguistic diversity. Hutchins and Hazlehurst [9] presented simulations employing communities of simple agents to model how a lexicon could emerge from interactions between agents in a simple artificial world. Their models were not based on the evolutionary perspective, but on the connectionist approach. They occasionally observed that the random initial starting points of the networks in a community were incompatible with each other, and this led to divergence in the verbal representations of these individuals.

Recently, Werner and Todd [17] have extended a previous model [16] to focus on exploring the idea that the origin of diversity in communication signals is due to sexual selection. In their new model, communication signals were used to attract females as mates, and sexual selection drove the evolution of male songs and female song preferences. Each male had genes that directly encoded the notes of his songs, and females' genes encoded a transition matrix used to rate transitions from one note to another in male songs. Each entry in the transition matrix represented the female's expectation that one pitch would follow another in a song. Werner and Todd investigated three methods for scoring the male songs, one of which is based on the idea in ethology that females exposed to the same song repeatedly will become bored and respond to that song less. They have shown that sexual selection could lead to maintenance of signal diversity, which was at its maximum in an initial population with many different male songs.

The first goal of our article is to investigate the origin and evolution of linguistic diversity from an evolutionary perspective. To do this we construct minimal models that are designed with regard to referential signaling in nonhuman animals and analyze their evolutionary dynamics based on the synthetic experiments. The second goal is to examine the possibility of utilizing linguistic diversity in the fields of distributed AI and robotics, based on the results of the above experiments. We believe that a very simple communication system can continue to generate linguistic diversity in an environment without spatial distribution. This supports the hypothesis that in an environment with limited amounts of resources that contains individuals with poor linguistic facilities, linguistic unification is not necessarily adaptive.

In Section 2 we discuss the design of the abstract model based on the communication systems found among nonhuman animals and show the results of the experiments. We then construct an agent-based model (Section 3) by introducing the evolutionary mechanism of the abstract model into a concrete situation to verify the results obtained in Section 2. In Section 3 we also examine the possibility of utilizing the mechanism in engineering fields. In Section 4 we discuss several issues concerning the origin and evolution of linguistic diversity and its application, based on the results described in the previous sections. Section 5 is a summary of the article.

2 Abstract Model

2.1 Background

Seyfarth, Cheney, and Marler's pioneering work [13] on the vervet monkey's alarm call system revealed that they produce acoustically distinct and discrete alarm call types, and in response to hearing such calls, individuals react with appropriate escape behaviors. It is a remarkable point that vervet monkeys are born with the ability to respond appropriately to general predator categories (e.g., things up in the air, slithering things on the ground), where learning plays virtually no role in modifying signal structure, either during early development or later in life [8]. A referential system is functionally significant because when an individual hears an alarm call, an appropriate antipredator response can be initiated without having to see what is going on. In fact, the vervet monkey's alarm call system is a beautiful illustration of how selection pressures might have favored signal diversification [8]. An all-purpose alarm call would not work for vervet monkeys, because it would not provide sufficient information about the type of predator or escape response that would be most appropriate.

Since the work on the vervet monkey's alarm call system, several other studies have focused on the problem of referential signaling in nonhuman animals, including other simian primates (e.g., rhesus macaques), prosimians (e.g., ringtailed lemurs), and a few other species (e.g., domestic chickens). It has become clear that these signals are used in various contexts such as predator encounters, discovering food, and social relationships. For example, when a food call is given, listeners obtain information about the availability of alternative food sources, which can serve to guide their foraging decisions. Characteristics of these communication systems, especially in primates, are as follows:

- The communication systems are composed of speakers and listeners. Those who encounter the predators (or food) produce acoustically distinct and discrete alarm (or food) calls, and in response to hearing such calls, listeners behave appropriately.
- The signals are referential in the sense that they are reliably associated with objects and events in the environment.
- They do not react instinctively as a direct expression of their internal states. They send the signals with some primitive type of intention on the assumption of the existence of listeners.

- They are born with the ability to respond appropriately to general categories. Learning plays a relatively small role in modifying signal structure.
- These types of communication systems illustrate how natural selection might have driven signal diversification.

The first steps toward human languages are still shrouded in mystery despite the studies and controversies in many fields, but the above-described communication systems might be strong candidates for the immediate steps, in other words, the "protolanguages." This article aims at exploring the origin and evolution of linguistic diversity using two different types of models (in Section 2 and Section 3) with a communication system that is based on that observed in nonhuman animals.

2.2 Definition

The communication system in our models is composed of N_{pop} individuals. Each has a simple vocabulary system that is represented by a table that relates words and meanings as shown in Figure 1. Identical words can appear more than one time, which corresponds to homonyms (word 12 in this figure), while each meaning appears one time in this table. These tables describe innate information and are transmitted to offspring by genetic operators.

First an initial population of N_{pop} individuals with randomly generated vocabulary tables is generated. A signaler and N_{rec} listeners are randomly selected at the beginning of each "conversation." In a conversation, a word is uttered by the signaler, and each listener is one of the following three types, based on the interpretation of the word:

- a listener that has the word in its vocabulary table, and its meaning is equal to the meaning in the signaler's vocabulary table ("right listener");
- a listener that has the word in its vocabulary table, but its meaning is not equal to the meaning in the signaler's vocabulary table ("misunderstanding listener");
- a listener that does not have the word in its vocabulary table ("ignorant listener").

In the case that the received word is a homonym in the listener's vocabulary table, one meaning is randomly selected as its interpretation. Figure 2 shows an example where a signaler sends the word 5 which expresses the meaning 2.

Here, we divide the "right listeners" into "successful listeners" and "unsuccessful listeners" because it would be necessary to take these constraints into consideration in many situations investigated. For example, in the case of the food call, some of the listeners that wish to obtain the food might nonetheless fail to do so, because of feeding competition. In the case of the alarm call, some of the listeners that intend to respond with behaviorally appropriate escape responses might nonetheless fail in their effort to escape from the predator.

In every conversation each individual belongs to one of the following categories: signaler, successful listeners, unsuccessful listeners, misunderstanding listeners, ignorant listeners, or nonparticipants, as shown in Figure 2, and they are rewarded with R_{send},

Meaning	0	1	2	3	4
Word	87	12	34	60	12

Figure 1. An example of a vocabulary table.

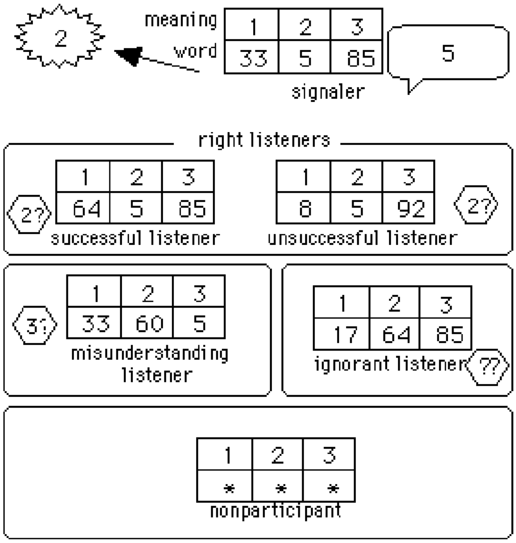


Figure 2. An example of conversation.

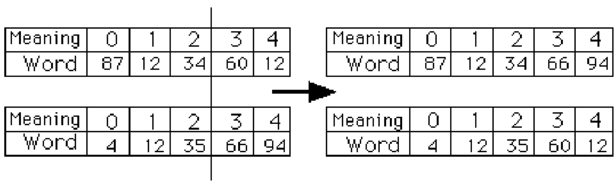


Figure 3. An example of crossover on a vocabulary table.

R_{share}, R_{unshare}, R_{wrong}, R_{ignorant}, or R_{out}, respectively. There can be positive, negative, and zero values. These rewards are genetic fitness scores for signaling.

After this process of conversation is repeated N_{conv} times, the information in the vocabulary tables is passed on to offspring by genetic operations. The next generation, which is also composed of N_{pop} individuals, is created by roulette selection based on the scores, where mated vocabulary tables cross over at a randomly selected point between columns (Figure 3). Then, mutation is performed on each word in the vocabulary tables with some probability P_{mut}, where the word is changed to a randomly selected word.

2.3 Experiments

We have conducted simulations following the procedure described above. The abstract model is general in the sense that it can represent many situations depending on the values of the rewards. In this article, we examine the communication system in the context that an individual finds a food source and utters the word for its meaning (the type of the food). We leave until later the case of alarm calls, though we see no reason why it should be different.

The number of the population (N_{pop}) was 64. If the number of the right listeners was not more than 4 in a conversation, all of the right listeners were considered to be successful and to obtain R_{share}. Otherwise, 4 successful listeners were randomly selected from the right listeners, and the remaining right listeners were considered unsuccessful because of competition. The individual that found the food source and successful listeners shared the food source equally, that is, $R_{send} = R_{share} = R_{food}/(n + 1)$, where the amount of the food source was R_{food} and the number of the successful listeners was n. R_{food} was set to be an arbitrary constant, 20. The reward for the individuals that interpreted the uttered word correctly but could not obtain the food source ($R_{unshare}$), was -3. The reward for the individuals that misunderstood the uttered word (R_{wrong}) was -2. The reward for the individuals that did not have the uttered word in their vocabulary tables ($R_{ignorant}$), and the reward for the individuals that did not join the conversation (R_{out}), were -1 and 0, respectively. The number of individuals that joined the conversation was always 20 ($N_{rec} + 1$). Each generation had 500 conversations (N_{conv}). Each word was expressed by an integer I ($0 \leq I \leq 99$). In this article, we investigate the case that there is only one type of food source (the size of the vocabulary table is 1) for convenience of analysis.

Plate 1a–d shows the evolutionary dynamics in vocabulary sharing where mutation rates (P_{mut}) are 0.01, 0.015, 0.04, and 0.1, respectively. The horizontal axes represent the generations. The vertical axes represent the distribution of words corresponding to meaning, and each same gray level means that an identical word is attached to the meaning.

Overall, as these figures show, the lower the mutation rate becomes, the more individuals there will be that have the same word for the meaning. The states of how the meaning was typically shared among the population were classified into the following four classes (the threshold values are approximate numbers).

Class A (P_{mut} is less than 0.015, Plate 1a):

A dominant word emerges, and the state becomes stable.

Class B (P_{mut} is nearly 0.015, Plate 1b):

The state that three to six words coexist and the state that one word spreads appear in turn.

Class C (P_{mut} is more than 0.015 and less than 0.07, Plate 1c):

Several words coexist. New words appear and then disappear repeatedly.

Class D (P_{mut} is more than 0.07, Plate 1d)):

The state changes in a chaotic manner.

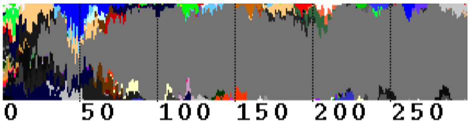
In Class B, the latter state was broken by an individual that had a new word generated by mutation. The reason this occurred is considered to be that the benefit to the mutant of monopolizing the food sources it found was larger than the benefit of sharing the sources found by the others by receiving the information of their existence at that moment. It is shown here that the unification of vocabulary tables in the population is not necessarily adaptive, which is a remarkable point.

Plate 2 shows the relation between the state of vocabulary sharing and the scores of agents when $P_{mut} = 0.015$. The upper part of this figure shows the state of vocabulary sharing, the middle part shows the average score of individuals and the Shannon's entropy of the words for the meaning, and the lower part shows the number of the words shared by more than three agents. The entropy was obtained by calculating

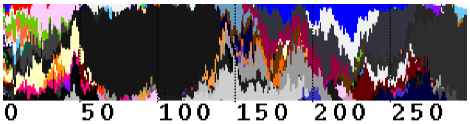
$$H = \Sigma p(I) \ln p(I),$$

where a word for the meaning is word I with probability $p(I)$.

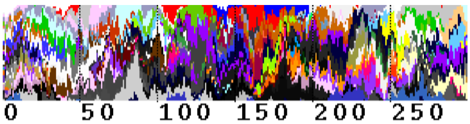
It is easy to make a distinction between the occasion where several words coexist and the occasion where there is only one dominant word in the middle graph. It is



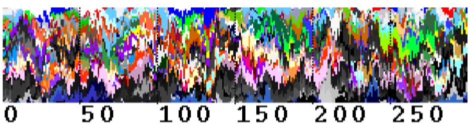
(a)



(b)



(c)



(d)

Plate 1. Evolution of vocabulary sharing: P_{mut} = (a) 0.01; (b) 0.015; (c) 0.04; (d) 0.1.

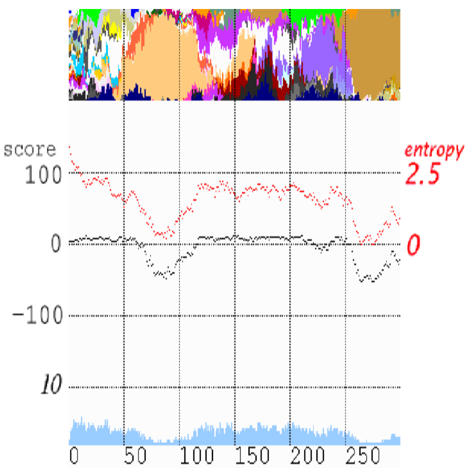
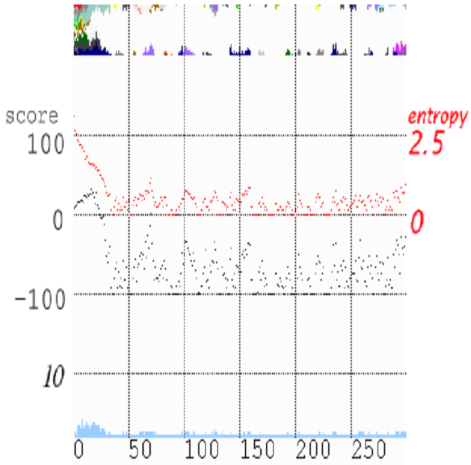
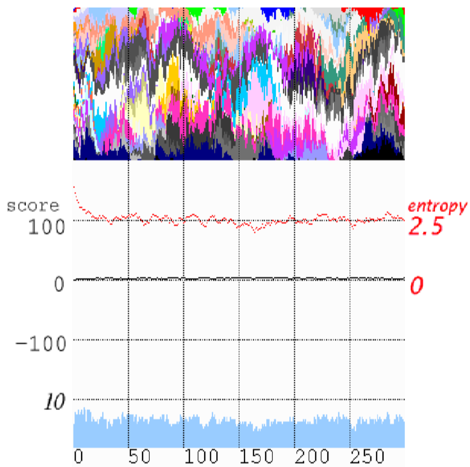


Plate 2. Average score, entropy, and the number of shared words.



(a)



(b)

Plate 3. Effect of varying population size: N_{pop} = (a) 32; (b) 128.

regarded as the cause of reduced scores in the state with a dominant word that a large number of individuals with the identical vocabulary obtained the reward (cost) R_{unshare} frequently in this state. The entropy dropped initially and moved according to the state of vocabulary sharing (a state of one dominant word or a state with several dominant words) after some kind of order emerged.

We have conducted another series of experiments concerning the effects of population size and the amount of the source (R_{food}). Some of the results are shown in Plate 3 and Plate 4. It can be found from these figures that an increase (decrease) in population size, or in the amount of food source, has similar effects to an increase (decrease) in mutation rate. One of the things that we notice is that there is a difference between those occasions where many words coexist because of increased mutation rate, and those occasions where many words coexist because of increased population size. In the former, the individuals with a new word appear repeatedly and the states change. In contrast, in the latter, the state has a tendency to be stable without allowing the individuals with a new word to appear. The experiments on the effects of varying the amount of the food source have shown that the greater the amount of the food source, the more individuals there will be that have the same word for the meaning.

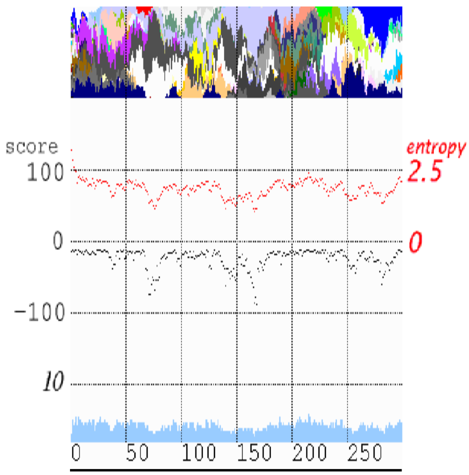
It has been assumed in all of the experiments to this point that any individual that has found the food source always signals. Here, we make a minor modification in the settings to investigate the motivation of signalers. We interpret that a specified word (the word 0 in this series of experiments) means being silent. If an individual that has found a food source has the word 0 corresponding to the food source, then it will not signal at all. Therefore, it could monopolize the food source, which will be a benefit, but at the same time it cannot obtain the information about the existence of the other food sources when the other individuals find them, which will be a disadvantage. The experiments have been conducted under the same conditions ($P_{\mathrm{mut}} = 0.015$) but with this modification. The results are shown in Plate 5a. A silent individual, that is, a mutant with this newly defined word 0, was generated by mutation at about the 180th generation, and then the silent group spread through the population rapidly. Communication died out in all experiments when silent individuals were allowed. The reason for this is thought to be that the silent individuals pay no penalty when they cannot obtain food sources, and at the same time, they have a slimmer chance of being sent signals from the individuals with a nonzero word, as the number of the silent individuals increases.

In the above-described experiment, when a silent individual found a food source, it monopolized all of that food source if it could. Next we modified this setting so that it could obtain half of the food source at most. The results are shown in Plate 5b. In this case, the silent group does not become dominant. The reason is believed to be that the silent individuals made less-efficient resource distribution than nonsilent individuals in the sense that occasionally the silent individuals left food sources without transferring information of the source. The issues concerning the silent individuals are worth examining, and some of them will be discussed in Section 4.

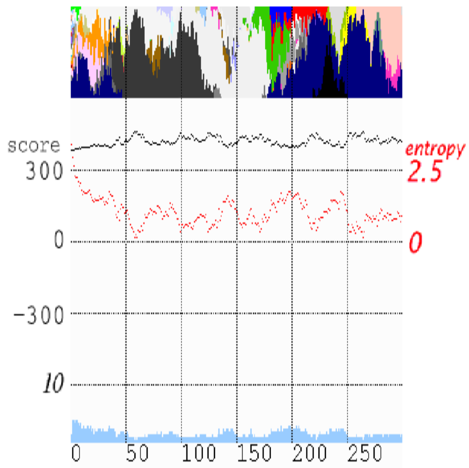
3 Agent-Based Model

3.1 Definition

We have introduced the evolutionary mechanism of the abstract model generating the linguistic diversity into a concrete situation and have constructed an agent-based model. The first objective of its design and experiments is to verify the results of the experiments concerning the abstract model, which depend on the explicit reward setting, by defining a concrete task done by agents. The second objective is to explore the possibility of

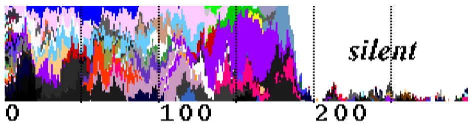


(a)

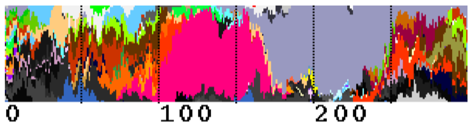


(b)

Plate 4. Effect of varying the amount of food: R_{food} = (a) 10; (b) 100.



(a)



(b)

Plate 5. Effects of allowing silent individuals: (a) full monopolization; (b) half monopolization.

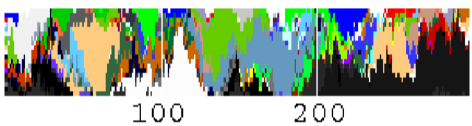


Plate 6. Evolution of vocabulary sharing.

applying the evolutionary dynamics of the linguistic diversity to issues in various fields, such as robotics and distributed AI.

Foraging behavior in a population of simple mobile agents (robots) has been taken up as the theme of the agent-based model. The task described in this section could be interpreted in many ways, as energy supply in robotics, or garbage collection in distributed AI, for example, because we have assumed a situation in which mobile agents move and gather food sources using a simple communication system.

The field has N_{pop} mobile agents and N_{food} food sources. Each agent has a vocabulary table and has an energy value as an internal state, which corresponds to a genetic fitness, though it could be negative. If the energy value of an agent is less than E_{hungry}, then the agent is "hungry." When the energy value is E_{full}, the agent is "full," and it can no longer eat the food source. Each agent consumes one unit of energy every time step.

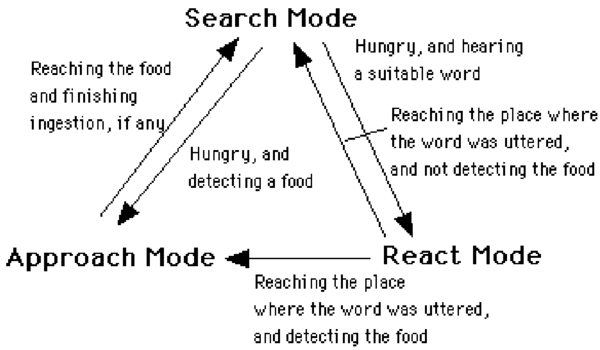


Figure 4. Transitions among behavioral modes.

The behavioral state of each agent is one of the three modes: search mode, react mode, or approach mode (Figure 4). At the beginning of every generation, agents and the food sources are located at randomly selected positions in the field. All agents are in search mode, and the energy values are E_{full}. Each agent in search mode selects randomly and engages in one of the following five behaviors: halting, moving forward, moving backward, turning right, or turning left. The distance of moving forward/backward, D, is randomly determined every time ($0 < D < L_{\mathrm{move}}$, in pixels). The speed of moving in any mode is constant (V_{agent} pixel/step). The angle of turning right/left, X, is randomly determined every time ($0 < X < A_{\mathrm{turn}}$, in degrees). It takes 1 time step to turn right/left. Agents in search mode detect food sources within a distance of L_{detect}. When a hungry agent in search mode finds a food source, it utters the word for it, and its state changes into approach mode. This signaling process takes 1 time step.

An agent in approach mode approaches the food source. Each food source also has an energy level, which is E_{food} initially. When an agent reaches a food source, it ingests the food until it becomes full or the energy of the food source becomes zero. If the energy of a food source becomes zero, it is removed from the field. Food sources are generated only when a new generation of agents is created. Other agents cannot get the information about the exhaustion of the food source. Therefore, when the food sources are removed, the agents that are in react mode, in other words, devoting themselves to going for the location where the word was uttered, would generate loss of time and energy for themselves. This cost, which is represented implicitly and naturally in this agent-based model, is equivalent to the value expressed by R_{unshare} in the abstract model.

Agents that are in search mode and are within a distance of L_{hear} can hear an uttered word. If an agent is hungry and is a "right listener," its state changes into react mode. Each agent in react mode approaches the location where the word was uttered. When an agent in react mode reaches the location, if it detects a food source, its state changes into approach mode; otherwise its state changes into search mode.

In this manner, the agents repeat searching for food, approaching food or the places the words were uttered, uttering words, and hearing the words, until N_{step} time steps pass from the beginning, or all food sources are consumed. Next, the information on vocabulary tables is passed on to offspring by genetic operations in a manner similar to that in the abstract model. The next generation, also composed of N_{pop} agents, is created by roulette selection based on the agents' energy values after scaling, where mated vocabulary tables cross over at a randomly selected point between columns. Then, mutation is performed on each word in the vocabulary tables with some probability

P_{mut}, where the word is changed to a randomly selected word. In this manner, these processes are repeated again and again for populations in subsequent generations.

3.2 Experiments

We have conducted some preliminary experiments with the following parameters: $N_{pop} = 20$; $N_{food} = 20$; $N_{step} = 10{,}000$; $L_{move} = 100$; $L_{detect} = 100$; $L_{hear} = 200$; $V_{agent} = 1$; $A_{turn} = 100$; $E_{hungry} = 3{,}000$; $E_{full} = 5{,}000$; $E_{food} = 4{,}500$; Field size was 1,000 by 1,000. Also, only one meaning was set up in this series of experiments. In other words, there was one type of food in the field. Evolution was observed for 300 generations.

Plate 6 shows the evolutionary dynamics in vocabulary sharing for 300 generations, where P_{mut} was 0.01. We have observed similar evolutionary dynamics to those in the abstract model, except that the effect of the mutation rate is slightly different. The threshold value is approximately 0.01 in this agent-based model, which divides Class A and Class C, while P_{mut} around 0.015 is the threshold in the abstract model.

The following two additional methods were investigated for comparative evaluation:

Method 1: All agents have the identical word-meaning relation (vocabulary table) a priori. Therefore, when an agent utters a word, each listener is either successful or unsuccessful and cannot be a misunderstanding listener or an ignorant listener. No genetic operators are used, and there is no evolution.

Method 2: There is no communication at all. All agents are silent all the time. There is also no evolution.

We have conducted 10 trials of the comparative experiments. The parameters have the same values as in the experiment shown in Plate 6. Results are shown in Table 1. Table 1 shows the average energy value and the maximum energy value among all agents, and the number of occurrences of all food sources being exhausted. We refer to the method based on the original agent-based model as Method 0 in this table. It is shown that the maximum energy value and the average energy value are higher in Method 0 than in Method 1 and Method 2. This means that the evolution of the vocabulary table contributed to the efficient task execution in these experiments. However, the number of occurrences of all food sources being exhausted is slightly higher in Method 1 than in Method 0. The cause of this seems to be that the communication with the identical word increased the cases of all food sources being exhausted, although it made the agents that heard the word waste time and energy. It is also shown that Method 2 (no communication) shows poor performance as compared with the other two methods. These results mean that the role of the evolving communication system with linguistic diversity is significant for the foraging behavior in the population of agents.

4 Discussion

4.1 Observed Linguistic Diversity

The results of the experiments imply that linguistic diversity grows when population size, mutation rate, or restriction on resources becomes greater. Figure 5 shows this implication roughly. From another point of view, it can be said that the communication system adapts to the growth of population size, mutation rate, or restriction on resources by increasing its linguistic diversity. One extreme case is that there is no diversity. This corresponds to the case where all agents shared an identical vocabulary table in the experiments with small mutation rates, or the case where all agents were silent in the experiments allowing silent individuals. The other extreme case is where they share

Table 1. Results of the comparative experiments.

Trial No.	Method 0 Avg E. (max E.) Exhaustion	Method 1 Avg E. (max E.) Exhaustion	Method 2 Avg E. (max E.) Exhaustion
1	−1254 (4004) 175	−3608 (4354) 184	−4508 (4409) 131
2	−2667 (4529) 166	−2715 (4392) 179	−5016 (4634) 143
3	−3560 (4476) 166	−4096 (4200) 170	−5595 (4305) 140
4	−3654 (4242) 168	−3698 (3824) 171	−2219 (4193) 137
5	−3362 (4143) 162	−1908 (4319) 170	−3716 (3905) 145
6	−1039 (4680) 170	−4300 (3952) 175	−2052 (3953) 127
7	−2933 (4186) 165	−396 (3933) 177	−3764 (4195) 139
8	−2011 (4371) 168	−2601 (4676) 183	−5492 (3926) 147
9	−2239 (4427) 171	−3509 (4498) 169	−3531 (3675) 139
10	−2035 (4060) 161	−1759 (4174) 173	−4203 (3619) 122
Average	−2475 (4312) 167	−2859 (4232) 175	−4010 (3081) 137

no stable and identical vocabularies at all, and they thus cannot efficiently transfer information by communication systems. This corresponds to the case with a quite large mutation rate in the experiments. The results of the experiments on the agent-based model have shown that the evolutionary dynamics could maintain a proper level of linguistic diversity and attain effective task execution.

There was a tradeoff between the monopoly of the resources discovered by an agent itself and the sharing of the resources discovered by other agents (to be exact, sharing with risks of additional competition). When the former exceeded the latter, the linguistic diversity observed in the experiments was generated by the selection pressure. This selection pressure allowed the individuals with new words to increase in the population.

In other words, individuals with new words can increase by making others respond with inappropriate reactions through misinterpreted words, which can be called functional deception, but not intentional deception (deception based upon manipulation of belief states). All agents became silent when we allowed individuals to be silent. In this case, they withheld information about food sources and thereby increased their fitness relative to others. We can call it another primitive form of deception. This type of deceptive behavior in nonhuman animals has been reported. In chimpanzees, food calls are given by individuals at relatively large food sources (implying that the costs of increased feeding competition may be negligible) [18]. Also, in some species, the probability of calling in the context of food is less than 100%, suggesting the possibility that individuals sometimes suppress their calls [8].

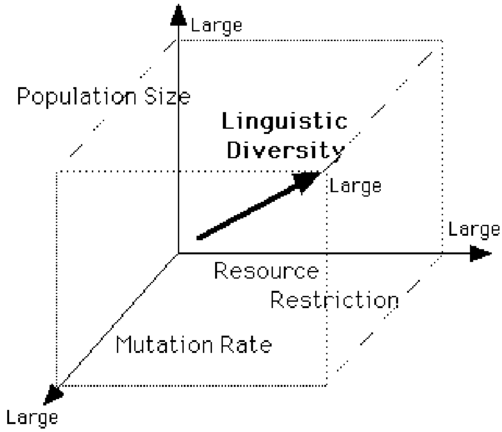


Figure 5. Growth of linguistic diversity.

The invasion of a silent population and the generation of linguistic diversity discussed in this article are closely linked to the issue of the origin of altruism [1]. Food calls would appear to be altruistic in general, because those who announce their discoveries are essentially inviting increased food competition and, consequently, potentially decreasing their own access to food. Kin selection and reciprocal relationships are strong candidates for its explanation. It has also been reported that there is social pressure making individuals call. Individual rhesus who found food but failed to call and were detected by other group members received more aggression than individuals who called upon discovery [7]. The result that the silent population disappeared when we reduced the maximum amount of food sources that individuals could obtain to half might be a candidate for its explanation at the lowest level.

The effects of incorporating a learning mechanism into these models would be worth investigating, though we have focused on the evolutionary dynamics of the linguistic diversity in this article. It is clear that the effects depend on the adopted learning algorithm. If we adopt a learning algorithm that uses the rewards in conversations as teacher signals in learning and modifies the word-meaning relations gradually, we would expect the learning mechanism simply to accelerate the evolutionary dynamics observed in these experiments. However, the contributions of population size, resource restriction, and mutation rate to linguistic diversity could be rather complex.

4.2 Possibility of Utilizing the Diversity

Application of the results of alife studies has been investigated, and it has begun to bear fruit in various fields. One of the promising fields is robotics. We have conducted the experiments on the agent-based model partly based on the idea that the communication system that evolves and maintains linguistic diversity would be beautifully fit to be used as the flexible mechanism for communication among a population of autonomous robots that attain cooperative behavior. The results in the experiments concerning the agent-based model are encouraging in the sense that the communication system supported the cooperative task execution.

The result that the energy value in the evolutionary model is higher than that in the model where the agents share a vocabulary table supports the hypothesis that in an environment with limited amounts of resources, containing individuals with poor linguistic facilities, linguistic unification is not necessarily adaptive. Here, we can grasp

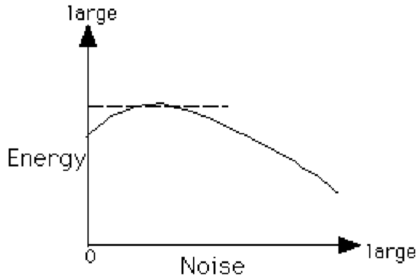


Figure 6. Another role of noise.

the observed evolutionary dynamics in the agent-based model from another point of view. Unification of word-meaning relations generated the loss of time and energy found in Method 1. In that case, if noise disturbed the communication to a certain extent, the loss would be reduced (Figure 6). We believe that in Method 0, the evolutionary mechanism realized the optimal linguistic diversity instead of the noise in this context. The results of the preliminary experiments support this.

The complexity in the mechanism of the communication system is extremely reduced, because we have aimed to implement a minimal communication system that generates linguistic diversity that could be utilized in engineering fields. Communication systems with far richer facilities, for example, those with which agents can negotiate on sharing the resources, would surely rank higher. On the other hand, slightly extended versions of the current communication system can be investigated, for example, as follows:

Version 1: The agents that found the sources signal only when they finish feeding and there are food sources left. This modification of setting can reduce the cost of listener agents.

Version 2: The volume of the food calls is set to be proportional to the amount of the food sources. This modification makes the number of the listening agents vary correspondingly to the amount of the food sources, which can reduce the cost of listener agents.

We expect both versions will rank higher than the results of our experiments, though we do not have enough evidence that in nonhuman animals there are such communication systems. Some species, however, have a food call that refers to the quality of food sources.

One of the most difficult hurdles to overcome to achieve physical realization based on the evolutionary dynamics, in general, is the relationship between simulations and actual robot execution. Even the experiments on this simple communication system were slow to evolve in the agent-based model. It is very difficult and may take as much time to run detailed simulations as it would take to build the actual robot systems. At the same time, it is also impractical to build and observe many actual robots during many generations. Therefore, we plan to adopt a hybrid simulated/embodied selection regime [11]. Large numbers of simulated robots are examined in simulation, but only the promising subset of these are actually built and examined, thereby reducing the scope of the problem. Simulated evolution of communication systems will also be necessary for speeding up the adaptation in the physically realized robotic systems in the near future, because the communication systems will be able to adapt to rapid changes in dynamic environments.

5 Summary

This article reports on the current state of our efforts to shed light on the origin and evolution of linguistic diversity, using synthetic modeling and artificial life techniques. We have constructed a simple abstract model for a communication system that is designed with regard to referential signaling in nonhuman animals. The evolutionary dynamics of vocabulary sharing were analyzed based on these experiments.

The results have shown that only a subset of initial conditions leads to the unification of vocabulary, and linguistic diversity evolves corresponding to the changes in population size, mutation rate, and restriction of resources. These facts support the hypothesis that in an environment with limited amounts of resources, containing individuals with poor linguistic facilities, linguistic unification is not necessarily adaptive. We have also observed that unification of vocabulary causes a decrease in genetic fitness of the individuals.

We have incorporated the idea of the abstract model into a more concrete situation and have presented an agent-based model to verify the results of the abstract model and to examine the possibility of utilizing the linguistic diversity in the field of distributed AI and robotics. It has been shown that selection pressure could explain the linguistic diversity in the cooperative behavior of multiple agents.

The proposed models can be extended in several directions. One obvious direction would be to investigate a model that focuses on not only evolution but also learning. Another direction would be to analyze the evolutionary dynamics using embodied multi-agent systems. One of the issues that we currently focus on is the relation between linguistic diversity and noise, which is partly described in Section 4.

Acknowledgments

This work is partially supported by the Artificial Intelligence Research Promotion Foundation and by the Hori Information Science Promotion Foundation. The authors would like to thank Charles Taylor for fruitful discussions. The helpful suggestions of Masaki Sasai and two anonymous reviewers are much appreciated.

References

1. Ackley, D. H., & Littman, M. L. (1994). Altruism in the evolution of communication. In R. A. Brooks & P. Maes (Eds.), *Artificial life IV* (pp. 40–48). Cambridge, MA: MIT Press.

2. Arita, T., & Taylor, C. E. (1996). A simple model for the evolution of communication. In L. J. Fogel, P. J. Angeline, & T. Bäck (Eds.), *Evolutionary programming V (Proceedings of the Fifth Annual Conference on Evolutionary Programming* (pp. 405–409). Cambridge, MA: MIT Press.

3. Arita, T., Unno, K., & Kawaguchi, K. (1995). A primitive model for language generation by evolution and learning. In *Proceedings of the International Workshop on Biologically Inspired Evolutionary Systems* (pp. 181–186).

4. Batali, J. (1994). Innate biases and critical periods: Combining evolution and learning in the acquisition of syntax. In R. A. Brooks & P. Maes (Eds.), *Artificial life IV* (pp. 160–171). Cambridge, MA: MIT Press.

5. Darwin, C. (1871). *The descent of man and selection in relation to sex.* London: John Murray.

6. Hashimoto, T., & Ikegami, T. (1995). Evolution of symbolic grammar systems. In F. Morán, A. Moreno, J. J. Merelo, & P. Chacón (Eds.), *Proceedings of the Third European Conference on Artificial Life* (pp. 812–823). Berlin: Springer-Verlag.

7. Hauser, M. D. (1992). Costs of deception: Cheaters are punished in rhesus monkey. In *Proceeding of the National Academy of Sciences, 89*, 12137–12139.

8. Hauser, M. D. (1996). *The evolution of communication.* Cambridge, MA: MIT Press.
9. Hutchins, E., & Hazelhurst, B. (1995). How to invent a lexicon: The development of shared symbols in interaction. In N. Gilbert & R. Conte (Eds.), *Artificial societies: The computer simulation of social life* (pp. 157–189). London: UCL Press.
10. MacLennan, B. (1991). Synthetic ethology: An approach to the study of communication. In C. G. Langton, C. Taylor, J. D. Farmer, & S. Rasmussen (Eds.), *Artificial life II* (pp. 631–658). Reading, MA: Addison-Wesley.
11. Miglino, O., Nafasi, K., & Taylor, C. E. (1995). Selection for wandering behavior in a small robot. *Artificial Life, 2*, 101–116.
12. Pinker, S. (1994). *The language instinct.* New York: William Morrow.
13. Seyfarth, R., Cheney, D. L., & Marler, P. (1980). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science, 210,* 801–803.
14. Steels, L. (1997). The synthetic modeling of language origins. *Evolution of Communication, 1*(1), 1–34.
15. Steels, L., & McIntyre, A. (1997). Spatially distributed naming games. http://arti.vub.ac.be/www/steels /spatial.ps.
16. Werner, G. M., & Dyer, M. G. (1991). Evolution of communication in artificial organisms. In C. G. Langton, C. Taylor, J. D. Farmer, & S. Rasmussen (Eds.), *Artificial life II* (pp. 659–687). Reading, MA: Addison-Wesley.
17. Werner, G. M., & Todd, P. M. (1997). Too many love songs: Sexual selection and the evolution of communication. In P. Husbands & I. Harvey (Eds.), *Proceedings of the Fourth European Conference on Artificial Life* (pp. 434–443). Cambridge, MA: MIT Press.
18. Wrangham, R. W. (1977). Feeding behaviour of chimpanzees in Gombe National Park, Tanzania. In T. H. Clutton-Brock (Ed.), *Primate ecology: Studies of feeding and ranging behaviour in lemurs, monkeys and apes* (pp. 504–538). London: Academic Press.