The Sixteenth International Symposium on Artificial Life and Robotics 2011 (AROB 16th '11),
B-Con Plaza, Beppu,Oita, Japan, January 27-29, 2011

Visualizing language evolution as an emergent phenomenon based on biological evolution and learning

Tsubasa Azumagakito, Reiji Suzuki and Takaya Arita
Graduate School of Information Science, Nagoya University
Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan
Tel : +81-52-789-4266; Fax : +81-52-789-4258
E-mail : azumagakito@alife.cs.is.nagoya-u.ac.jp

Abstract: From the artificial life perspective, language can be viewed as a complex adaptive system emerging from linguistic interactions between individuals. Language and the human brain have evolved in parallel and interacting with each other. In this study, we propose a model of language evolution based on biological evolution and learning. In our model, the linguistic space is expressed in the polar coordinate system in which each possible language is expressed as a point. We conduct evolutionary experiments based on the model and visualize the results in the linguistic space. The trajectory of distribution of innate linguistic abilities shows the diversification and complexity growth of language. In the extended experiment, in which the angular coordinate represents the additional effect on cost for the plasticity, we observe a general tendency that the cost of plasticity evolves to become smaller. However, it never evolves to be zero, which might suggest that some cost of plasticity producing the Baldwin effect is adaptive in language evolution.

Keywords: language evolution, visualization, learning, artificial life.

1 Introduction

Language distinguishes humans from other animals. Language allows us to accumulate knowledge and transmit it across both space and time. This has led to a greater understanding of the world and accelerated cultural achievement. The evolution of language has been the subject of numerous debates and speculations. Nevertheless, it is difficult to study in a scientific manner and remains an open research question.

Recently, a constructive approach has been adopted to investigate language evolution [1]. It is characterized by the use of computational models from the viewpoint that language is a complex adaptive system emerging from linguistic interactions between individuals. Another viewpoint states that language and the human brain have evolved in parallel and interacting each other. In other words, they have coevolved. If we focus on the evolution of the human brain, there are two typical adaptive processes at different time scales: biological evolution and learning (phenotypic plasticity) [2].

Based on these viewpoints, in previous work [3], we investigated the coevolution between communication ability and phenotypic plasticity to clarify whether and how leaning can facilitate evolution in dynamic environments arising from communicative interactions among individuals. To do this, a simple computational model was devised to do this. The levels of adaptive communication of signaling and receiving processes are determined by different sets of traits. Each level represents the expected value of fitness contribution for a successful communication. A communication is successful only when the levels of the signaler and the receiver are the same. The agents try to improve their communication levels through learning in which the values of plastic traits can be modified from their genetically determined values. The evolutionary experiments showed that the population with a learning ability successfully increased its shared level of communication while the population with no leaning was not able to increase the level. It was also shown that the Baldwin effect (typically interpreted as a two-step evolution of the genetic acquisition of a learned trait without the Lamarckian mechanism [4]) repeatedly occurred and facilitated the evolution.

The purpose of this research is to study the general roles of biological evolution and learning in the evolution of language. For this purpose, we construct a generalized model for the coevolution between the communication ability and phenotypic plasticity. It is a generalization over the model devised in the previous work [3] in the following two aspects. 1) The linguistic space is expressed in the polar coordinate system in which each possible language is expressed as a point, and the success in the conversation is determined geometrically (instead of using a specific task as in the previous work). Therefore, we can observe the coevolution as trajectories of innate linguistic abilities of agents in the linguistic space. 2) The fitness can be defined by adjusting the benefit from the communicating agents, the benefit from the complexity of the used language, and the cost of learning, independently.

2 Model

2.1 Agent and communication

N agents in the population communicate with each other using their language capacity. The linguistic space is expressed in the polar coordinate system in which each possible language is expressed as a point (Figure 1). The distance from the origin (r) represents the complexity of the language, and the angle from the positive x-axis (θ) represents a language type. Each agent is represented as a point and a field surrounding the point in the linguistic space. The former corresponds to the agent's innate language, and the latter corresponds to linguistic plasticity as an innate attribute of the agent. The plasticity is expressed as a fan-shaped field with area determined by r_p

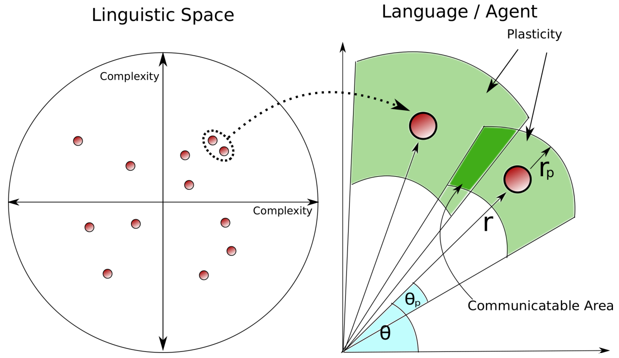


Figure 1: Linguistic space and agents.

and θ_p as shown in Figure 1.

Overlap of two plasticity fields means that corresponding two agents can communicate with each other by learning (using their plasticity). Agents with large plasticity can communicate with many agents although they incur a large cost proportional to the area of their plasticity field. The polar coordinate system reflects the situation where the difficulty and cost of communication is proportional to the complexity of the language.

2.2 Fitness Evaluation

The fitness of each agent depends on the number of communicating agents, the complexity of the innate language, and the area of the linguistic plasticity. The fitness function is defined as:

$$Fitness = (L \quad \frac{L^2}{2G})^{w^1} \quad r^{w^2} \quad (4\theta_p r_p r)^{w^3}, \qquad (1)$$

where w^i (i=1, 2 and 3) are weights for three components of the fitness function. The first term represents the benefit from the number of communicating agents. L is number of communicating agents and G is the population size.   is a parameter that determines the change in the benefit of communications with the increase in the number of communicating agents. There are three possible situations: a) a linear increase (=0), b) an exponential increase ($<$ 0), c) the existence of an optimal number of communicating agents for the best benefit ($>$ 0), as shown in Figure 2. Case b) corresponds to the situation in which there is a synergetic effect in information sharing, and c) corresponds to the situation in which the benefit of the information decreases if it is shared by too many agents due to some restrictions (e.g., the limitation of resources). The second term represents the benefit of communications depending of the complexity of the language. The more complexity will bring about the greater benefits. The third term represents the cost of learning. We assume that it is proportional to the area of plasticity ($4\theta_p r_p r$), as it is probable that more complex and different languages are difficult be learned.

2.3 Evolution

The agents are selected using roulette wheel selection to reproduce according to their fitness to form the next generation. Mutation is performed with probability P_m. The genotypes of offspring are mutated by adding a small random value: R(0, 12) for r_p, r, and $5/r$ R(0, 1) for θ, θ_p, where R(, ^{2}) is a normal random number with mean

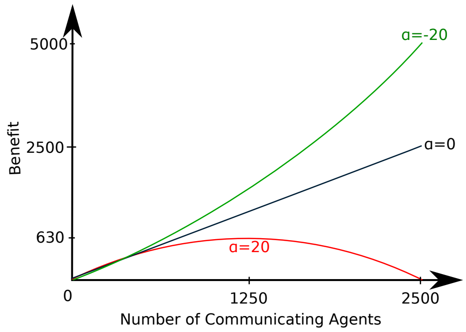
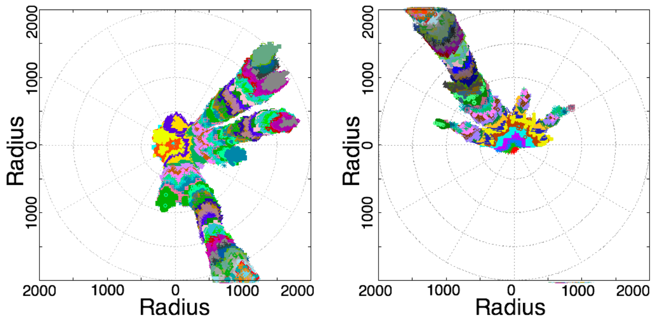


Figure 2: The effect of   on the benefit.



(a)   = 20   (b)   =   20

Figure 3: Results of language evolution.

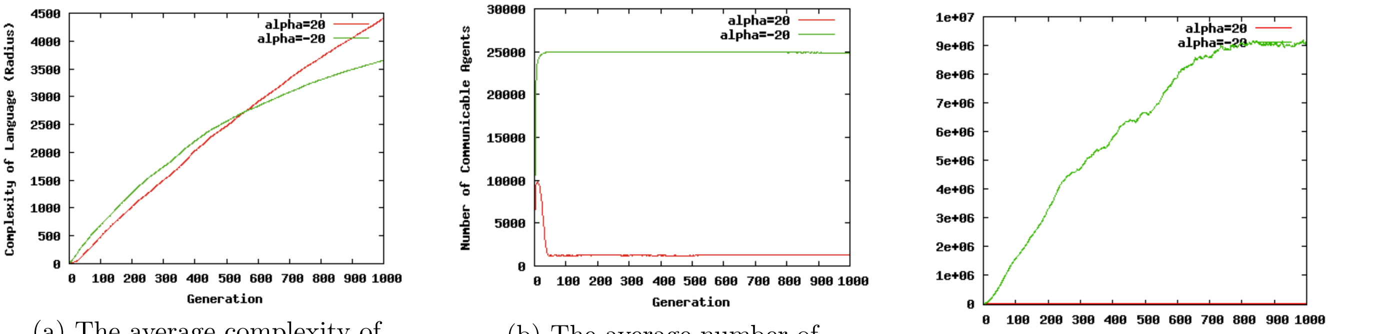
and variation ^{2}. Note that the range of a random value for θ, θ_p is inversely proportional to r of the parent. This property keeps the amount of displacement of the innate language or change in its plasticity due to mutation constant independent of the location of the agent in the linguistic space.

3 Results

We conducted evolutionary experiments for 1000 generations and visualized the results in the linguistic space. The following parameters were used: N=25000, w^1=1, w^2=1, w^3=1.1, P_m=0.8. The initial values of the genotypes of the agents $(r, r_p, \theta, \theta_p)$ were all zero.

3.1 Evolution of linguistic diversity and complexity

We conducted two experiments in which   was 20 or 20. The results using the two-dimensional polar coordinate system are shown in Figure 3. The innate linguistic abilities of all agents were plotted with a unique color for every 10 generations. Figure 3(a) and (b) show the trajectory in the case of =20 and 20 respectively. Both figures indicate that the agents formed some linguistic clusters from the initial population at the origin, then they increased their complexity of language gradually. However, the clusters converged to one large cluster with high complexity by the end of experiments. It is notable that distant clusters coexisted for a longer generation when   was positive (Fig 3(a) vs. (b)). This means that the negative effect of information sharing on the fitness caused by the excess number of communicating agents contributed to the maintenance of high linguistic diversity.



(a) The average complexity of language.

(b) The average number of communicating agents.

(c) The average plasticity of agents.

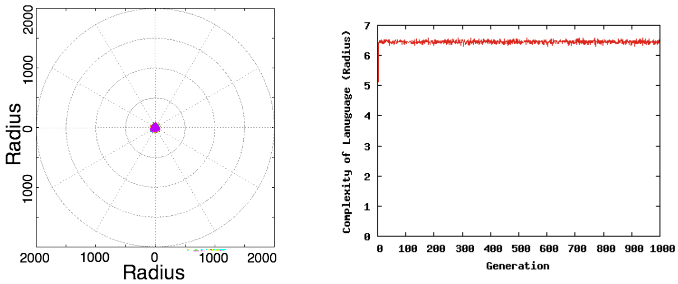
Figure 4: The comparison of agent abilities between experiments with =20 and =-20.

Figure 4 compares some properties of agents in both cases. Figure 4 (a) shows the evolution of the average innate complexity of language. The horizontal axis represents the generation. We see that the average complexity in the case with =20 was higher than the one in the case with =20 after around the 550th generation. This implies that the negative effect of information sharing could also accelerate the later evolution of the complexity of language.

Figure 4(b) and (c) shows the evolution of (b) the average number of communicating agents and (c) the average plasticity for each agent respectively. We define the plasticity of each agent as its area $(4\ \theta_p\ r_p\ r)$. Figure 4(b) showed a sharp increase in the number of communicating agents for about 20 generations and its subsequent decrease to the small value (about 1250) in the case of =20. In contrast, the number of communication in the case of =-20 sharply increased to its maximum 25,000 (equal to the population size). Figure 4(c) shows that the sufficient amount of plasticity increased drastically from the initial generation in the case of = 20. In the case of =20, the plasticity evolved to the relatively smaller value 12694 at 1000th generation, although it is too small to see its value in Figure 4(c). This indicates that agents in the case of =20 formed linguistic clusters that keep the most beneficial size (about 1250) by controlling their plasticity. Note that, although Figure 3(a) shows that the linguistic clusters seem to have converged to one cluster in the last generation, that cluster is composed of small sub-clusters that kept the most beneficial size. Furthermore, the linguistic clusters that were close to each other were more robust against a mutation of θ. This is an explanation of why the linguistic clusters converged to one cluster in spite of =20.

As a whole, the results indicate that the evolutionary scenario of linguistic diversity as follows. 1) In the early stages of evolution, there was an increase in the number of agents with more plasticity to communicate each other. 2) Then, in the case of =20 (means there is negative effect of information sharing), the plasticity is adjusted to small value to keep population size in optimal value.

In addition, we conducted experiments without learning in which agents have no plasticity. Specifically, a communication between two agents results in success when they share the same r and θ perfectly. The results with =20 are shown in Figure 5. We see that the complexity of language rapidly converged to about 6.5 in early generations.



(a)Trajectory of evolution.

(b)Complexity of language.

Figure 5: Results of experiments without learning.

It shows that the evolutionary process of the complexity of language tended to constant without learning.

3.2 Effects of learning cost

We conducted further experiments to understand the effects of the leaning cost on the evolution of the population. We used various weights for learning cost (w^3) ranging from 0 to 2 at intervals of 0.05. Here we focused on the case of =20. Figure 6 shows four typical trajectories of the population when w^3= 0, 1.2, 1.4 and 1.6 respectively. It shows that as the weight of learning cost increased, the increasing rate of the linguistic complexity decreased, and the linguistic clusters tended to converge around the origin.

The average plasticity and complexity of language in the last generation in these cases are shown in Figure 7.

When the weight was relatively small ($w^3 < 1.0$), the complexity of language reached a high value around 4800. The plasticity value has a wide distribution between low value and significant high value comparatively. In these cases, the increase in the complexity of language brings about the higher benefit compared with its effect on the cost. Thus, the higher complexity of language was essential for survival of agents. Besides, due to the smaller cost of learning, the plasticity often became high value. As a result, the language tended to become more complex and the plasticity often reached high values.

When $1.0 \leq w^3 < 1.4$, both indices tended to become smaller as w^3 increased due to the increased cost of learning. There is a possibility that the smaller plasticity retarded the evolution of the complexity because agents might not be able to communicate with their own mutants or not be able to keep the optimal number of communicating agents.

When $w^3 \geq 1.4$, the complexity converged to the small

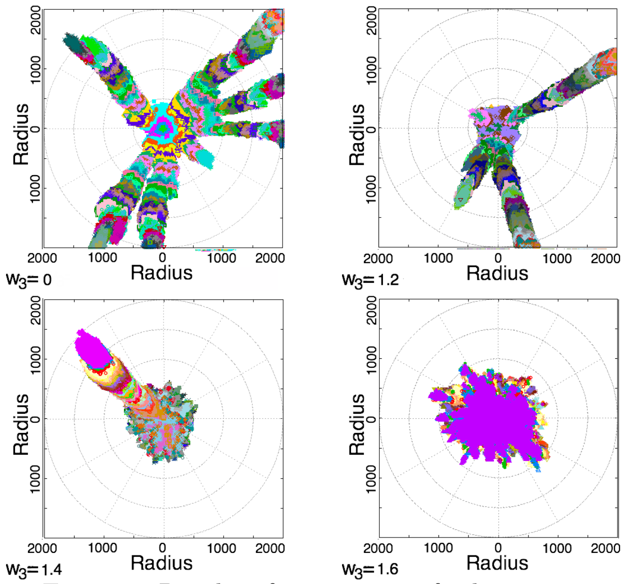
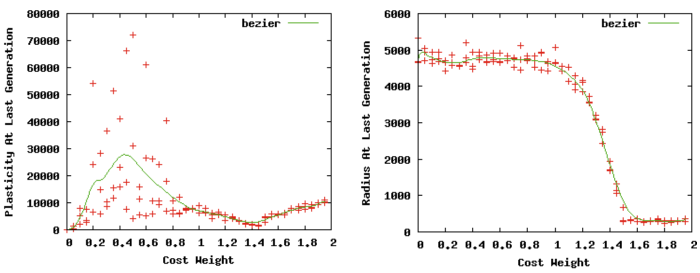


Figure 6: Results of experiments for learning cost.



(a) The average plasticity of agents.

(b) The average complexity of agents.

Figure 7: The results comparison of learning cost.

value around 200 as w^3 increased. In these conditions, the increase in the complexity of language yields much larger cost. Thus, the population was supposed to evolve to use the less complex languages. The gradual increase in the average plasticity when $w^3 \geq 1.4$ was due to the slower convergence to the small values caused by the higher cost.

As a whole, it turned out that the degree of learning cost strongly affected the diversification of languages and their properties.

3.3 Additional cost experiments

Finally, we conducted experiments to consider the evolution of learning cost. In these experiments, the angular coordinate of an agent θ was associated with an additional cost of learning *Cost*, which was added to the fitness value calculated as follows:

$$Cost = \begin{cases} (|\frac{\pi}{2} \quad \theta| \quad \frac{2(4\theta_p r_p r)^{w^3}}{\pi}) & if\ 0 < \theta \leq \pi/2, \\ (|\frac{\pi}{2} \quad \theta| \quad \frac{2(4\theta_p r_p r)^{w^3}}{\pi}) & if\ \pi/2 < \theta \leq \pi, \\ (|\frac{\pi}{2} \quad \theta| \quad \frac{2(4\theta_p r_p r)^{w^3}}{\pi}) & if\ \pi < \theta \leq 3\pi/2, \\ (|\frac{\pi}{2} \quad \theta| \quad \frac{2(4\theta_p r_p r)^{w^3}}{\pi}) & otherwise. \end{cases}$$

When θ is 0 or 2π, there is no effect of the additional cost. As θ gets closer to π, the additional cost increases. We used 20 as .

The typical results are shown in Figure 8. Each figure shows the result of a trial with a different random seed. They illustrate that linguistic diversification occurred in the first or the fourth quadrant due to the higher cost

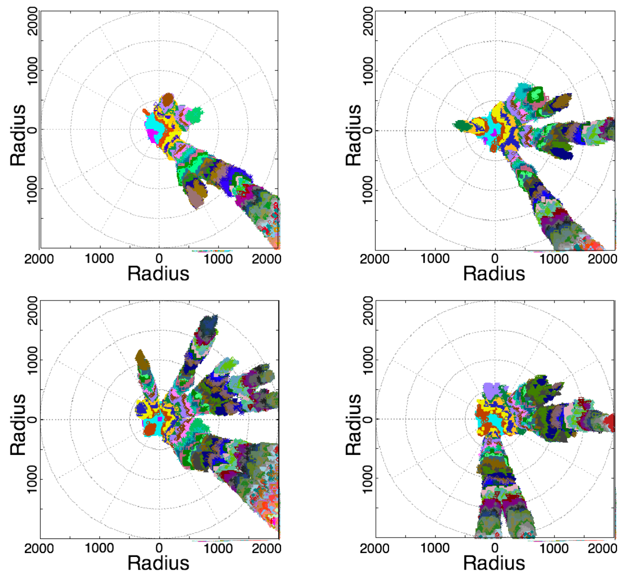


Figure 8: Results of extended experiments.

in the second and third quadrant. Thus, the language that had low cost to learn was selected in the evolution of language. The results show that excessive cost of learning prevents evolution of language.

4 Conclusion

In this paper, we conducted experiments with a comprehensive model of language evolution. The results of our experiments showed some implications for language evolution as follows. First, the linguistic complexity and diversity can emerge through interactions between evolution and learning. Second, the negative effect of information sharing on the fitness caused by the excess number of communicating agents contributed to the maintenance of high linguistic diversity. Third, agents may not be able to advance their linguistic complexity without the plasticity of linguistic abilities. Finally, the excessive cost of learning can prevent evolution of language because the plasticity is not enough to cover the mutation range. It is not necessarily the case that the low cost of leaning accelerate evolution of language due to the plasticity covering the whole mutation range.

Future work includes considering geographical factors in linguistic diversification and the asymmetric aspects of benefit of communications by extending our model.

References

[1] Luc Steels. The synthetic modeling of language origins. *Evolution of Communication*, Vol. 1, No. 1, pp. 1–34, 1997.

[2] Yusuke Watanabe, Reiji Suzuki and Takaya Arita. Language evolution and the Baldwin effect. *Artificial Life and Robotics*, Vol. 12, pp. 65–69, 2008.

[3] Reiji Suzuki and Takaya Arita. How learning can guide evolution of communication. *Proceedings of Artificial Life XI*, pp. 608–615, 2008.

[4] Peter Turney, Darrell Whitley and Russell Anderson. Evolution, learning, and instinct: 100 years of the Baldwin effect. *Evolutionary Computation*, Vol. 4, No. 3, pp. 4–8, 1996.