

## Cyclic Behavior in Gene-Culture Coevolution Mediated by Phenotypic Plasticity in Language

Tsubasa Azumagakito<sup>1,2</sup>, Reiji Suzuki<sup>1</sup> and Takaya Arita<sup>1</sup>

<sup>1</sup>Graduate School of Information Science, Nagoya University

<sup>2</sup>azumagakito@alife.cs.is.nagoya-u.ac.jp

### Abstract

The evolution of language has been the subject of much debate and speculation. It is difficult to study in a scientific manner and remains an open research question. This paper proposes an integrated computational framework for investigating possible scenarios of genetic and cultural evolution of language. Specifically, our framework aims to capture cultural evolution to allow for investigation phylogenetic dynamics of language, and at the same time to capture genetic evolution of phenotypic plasticity to allow for investigation of the role of the Baldwin effect in language evolution, while keeping the framework as simple as possible. In our evolutionary experiments and analysis, we discovered a coevolutionary scenario involving biological evolution of phenotypic plasticity, and a cyclic coevolutionary dynamic between genetic and cultural evolution, mediated by phenotypic plasticity.

### Introduction

Language distinguishes humans from other animals. Of course, other animals also engage in vocal communication. For example, Velvet monkeys can convey some simple information using alert calls (Sayfarth et al., 1980). However, such animals' vocal communication lacks the complex grammar and high expressiveness that characterizes human languages. Why do only humans have sophisticated language? This is one of the core questions on the path to understanding the human identity. This paper focuses on the evolution of the fundamental traits underlying communicative interaction in the context of biological evolution (genetic evolution of the language faculty), such as the rules or conventions for the effective communication necessary for collective behaviors. We assume that such traits can evolve under directional selection because the traits can be modified incrementally to increase the benefit from communicative interactions. We also assume such traits must be shared between individuals for communication to succeed. Accordingly, at least some of the selection will be positively frequency-dependent. This might obstruct evolution based on directional selection. We believe that this captures a fundamental and general problem in the evolution of communicative traits. For example, in the context of language

evolution it has been pointed out that mutations in grammar cannot be beneficial because the peers of an individual with a grammar mutation may not understand the mutant form (Pinker and Bloom, 1990; Glackin, 2010).

We believe that nature's solution to this challenging problem is found in the evolution of phenotypic plasticity. Phenotypic plasticity refers to the variability in the phenotype obtained from a given genotype resulting from development in different environments (West-Eberhard, 2003). In recent evolutionary biology, ontogenetic adaptation<sup>1</sup> based on phenotypic plasticity is recognized as one of the key factors that brings about adaptive evolution of novel traits (West-Eberhard, 2005; Gilbert and Epel, 2009). Wund provided a summary of eight hypotheses on how plasticity might influence evolution (including several pieces of empirical support), focusing mainly on adaptation to new environments (Wund, 2012). For example, the hypothesis that phenotypic plasticity promotes persistence in a new environment, and the hypothesis that a change in the environment can release cryptic genetic variation via phenotypic plasticity, in turn impacting the rate of evolutionary responses. Zollman and Smead (2010) analyzed simple models of language evolution based on Lewis's signaling game and the prisoner's dilemma game. They observed that the presence of plastic individuals alters the trajectory of evolution by directing the population away from a non-adaptive signaling and toward the optimal signaling. They termed this the "Baldwin optimizing effect". Suzuki and Arita also showed that such an adaptive shift can occur repeatedly, using a computational model of the coevolution of signal sending behavior and signal receiving behavior, that incorporated behavioral plasticity (Suzuki and Arita, 2008, 2012, 2013). These studies indicate that learning could be an important driving force for adaptive evolution in the context of communicative interactions.

This paper also looks at the relationships between two aspects of language evolution: biological evolution and cultural evolution. The relationship between genes and lan-

<sup>1</sup>Adaptive changes that occur during the lifetime of an organism (e.g., learning).

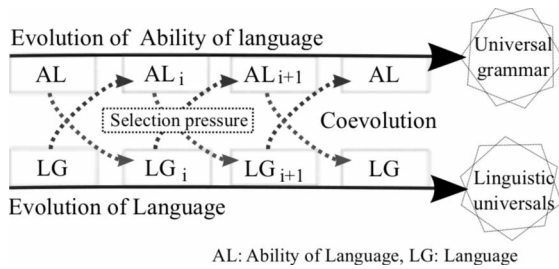


Figure 1: Coevolution between language and language ability. The upper arrow represents the evolution of language ability in general. The lower arrow represents the evolution of language itself. Language ability and language are represented by  $AL$  and  $LG$ , respectively.  $AL_i$  controls the selection pressures shaping  $LG_{i+1}$  and conversely  $LG_i$  controls the selection pressures shaping  $AL_{i+1}$ . Universal grammar (the theory proposing that the ability to learn linguistic grammar is hard-encoded into the brain) and Linguistic universals (general pattern that potentially exists in almost all of natural languages) may have emerged as an inevitable result of this coevolution.

guage is extremely complex and shrouded in controversy. Furthermore, rather than viewing language as a monolithic and independent entity, modern researchers typically break it down into its component mechanisms and analyze these independently (Fitch, 2011). Steels (2011) discussed various computational models of cultural evolution. He concluded that cultural evolution is a more powerful process than usually assumed, and that human language evolution’s dependence on genetic evolution is relatively limited. Several researchers argue that cultural evolution has fundamentally limited influence on the genetic evolution of the language faculties. This contrasts with various results that indicate that genetic biases are essential to language evolution. For example, Chater et al. (2009) have shown using a computational model that there are strong restrictions on the conditions under which the Baldwin effect can embed arbitrary linguistic constraints, and that the effect only emerges when language provides a stable target for natural selection. These approaches should be seen as complementary. There is a need to integrate these efforts and explore the relevant gene-culture coevolutionary interactions (Mesoudi et al., 2011).

We suggest the insights from these studies of language evolution together can be brought together using the concept of coevolution between language and brain, as it lets us integrate biological and cultural evolution. The idea of coevolution was originally suggested by (Darwin, 1871), and others have taken up his lead (Deacon, 1997), As illustrated in Fig. 1, the main idea is as follows: On one hand, a language is continuously changed its users, which brings about the linguistic variation much like mutation brings about ge-

netic variation. Language variants that can easily be learned by their users can survive and thus spread in the population of languages. On the other hand, having innate linguistic abilities (e.g., universal grammar) that equip an individual to handle the existing language variants and language changes provide a fitness advantage. Thus, genes for an innate language faculty will spread in the biological population. So we have two intertwined adaptation processes: language adapts to the brain, and the brain adapts to language. We believe that this insight is crucial to a comprehensive understanding of language evolution.

In this paper, we employ the coevolutionary framework described above, and propose a bottom-up computational model for investigating possible scenarios of the genetic and cultural linguistic evolution. We aim to capture cultural evolution to investigate the phylogenetic dynamics of language evolution, while at the same time capturing genetic evolution of phenotypic plasticity. The latter allows us to investigate the role of the Baldwin effect (typically interpreted as a two-step evolution of the genetic acquisition of a learned trait without the Lamarckian mechanism (Peter Turney and Anderson, 1996)) in language evolution, all the while keeping the framework as simple as possible. In order to do this, we extend our previous works on language evolution (Suzuki and Arita, 2008, 2012; Azumagakito et al., 2011, 2012). The main idea is to express the linguistic space as a polar coordinate system, in which individuals and languages gradually move about by genetic evolution and cultural evolution, respectively, as follows: 1) the more an individual can communicate, and the higher the expressivity of its languages, the more offspring it can produce. This nudges the population toward the locations of fit individuals. 2) Languages move towards the center of their user base, and divide, merge or go extinct in accordance with the distribution of their users.

Recently, there is much research on the diversification processes of real human language (Pagel, 2009; Lupyán and Dale, 2010; Atkinson et al., 2008; Atkinson, 2011). Levinson and Gray (2012) reviewed various methods to investigate the diversification of languages, and observed that phylogenetic analysis techniques from evolutionary biology are useful tools for elucidating the diversification of language. Phylogenetic patterns of language evolution were uncovered using such analyses, but the causal mechanisms of diversification are still unclear. These previous investigations proceed within the observational and deductive realm. The framework proposed here provides the means for experimentation, and a method to generate phylogenetic trees that can help elucidate the causal mechanisms of language diversification (see Fig. 2).

2) Analyze the feature of estimated tree, for example word-order relation between language families.

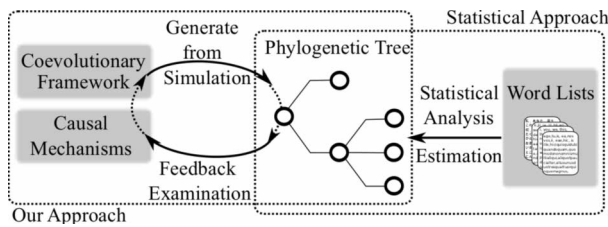


Figure 2: Phylogenetic comparative analysis using the proposed framework. Simulation results based allow us to generate phylogenetic tree of language families. By comparing the experimentally generated trees with the trees inferred from real linguistic data sets (such as standard word-lists), we can elucidate how and what causal mechanisms may have brought about the features of the inferred trees, and verify the validity of our simulation model. In this paper, we generate a phylogenetic tree from an evolution experiment. A comparison of it with the inferred trees from the real linguistic data is our future work.

### Model

We propose an integrated computational framework for investigating possible scenarios of genetic and cultural evolution of language. This framework allows us not only to capture coevolutionary interactions between languages and agents, but also to track the phylogenetic evolutionary process of languages.

### The Linguistic Space

There are  $N$  agents in a population, and they can communicate with each other using their shared languages. Agents and languages exist in a two-dimensional linguistic space, represented as a polar coordinate system, as shown in Fig. 3. Each language ( $L$ ) is defined as a point in the space. The distance  $r_L$  from a language to the coordinate system's origin represents the language's expressiveness, which contributes to the expected fitness benefit of a successful communication in that language. The angle of the language  $\theta_L$  to the origin represents its structural character. Each agent ( $A$ ) is represented as a point and a field surrounding the point. This point represents the agent's innate language ability, determined by its genotypes  $r_A$  and  $\theta_A$ . The agent can use the corresponding language in the linguistic space without learning. The field represents its linguistic plasticity (i.e. the range of its linguistic learning ability), as a  $p_A \times p_A$  fan-shaped field determined by its genotype  $p_A$ . The agent can use any language that falls within its plasticity field for communication. This polar coordinate system captures the fact that, as expressivity increases, the space of possible linguistic structures grows, such that agents with more expressive languages will be harder to communicate successfully with, due to the limited size of the plasticity field.

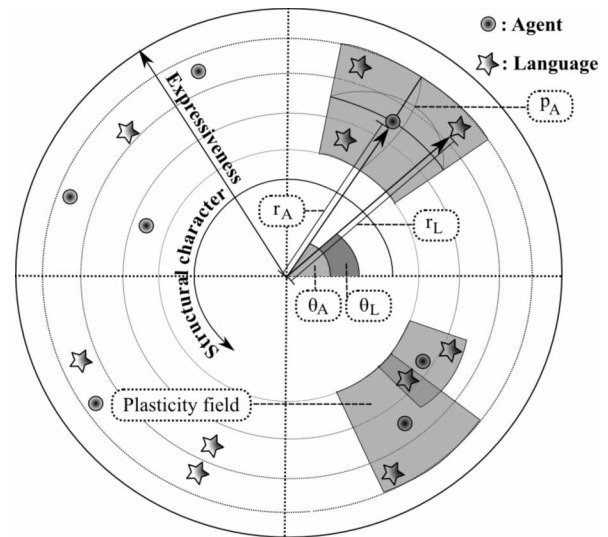


Figure 3: The linguistic space.

### Linguistic Interactions

In each generation, all possible pairs of agents make an attempt to communicate. If the two agents of a pair share one or more languages, they can communicate successfully. The fitness of each agent depends on its number of successful communicative interactions, the expressiveness of the languages used in those interactions, and the cost of its linguistic plasticity. The fitness function is defined as:

$$Fitness = N_{ca}^{w_1} \cdot \left( \frac{\sum_{i=0}^{N_{ul}-1} r_{L_i}}{N_{ul}} \right)^{w_2} - (p_A^2)^{w_3}, \quad (1)$$

where  $w_i$  ( $i=1, 2$  and  $3$ ) are weights for the three components of the fitness function. The first component represents the benefit from successful communicative interactions, which is proportional to the number of agents with which the focal agent successfully communicated  $N_{ca}$ . The second component represents the benefit from the expressivity of the languages available to the agent.  $N_{ul}$  is the number of languages within the plasticity field of the focal agent, and  $r_{L_i}$  is the expressiveness of the  $i$ -th language among them. The second component is an approximation of the average expressiveness of the languages used for successful communications. We adopted an approximation in order to reduce the computational cost. The third component represents the cost of linguistic plasticity. It is proportional to the area of plasticity ( $p_A^2$ ), because the further a language is removed from an agent's innate language ability, the bigger the demands on the agent's plasticity, and thus plasticity for further and further removed languages comes with higher and higher costs for maintaining the requisite plasticity. Overall, this fitness definition reflects that agents who can communicate with many other using expressive languages acquired

with limited linguistic plasticity will have high fitness.

### Biological Evolution of Language Ability

After the communicative interactions of agents, biological evolution of language ability occurs as follows: 1) Parent agents for the next generation are selected using roulette wheel selection (i.e. the probability that an agent is picked as a parent is proportional to its fitness). 2) Each genotype of each offspring is mutated with probability  $P_m$ . Mutation adds a small random value of  $R(0.01)$  to the values of  $r_A$  and  $p_A$ , and of  $R(0.01/r_A)$  to  $\theta_A$ , where  $R(x)$  produces random numbers between  $-x$  and  $x$  with a triangular distribution. Note that the range of a random value for  $\theta_A$  is inversely proportional to  $r_A$  of the parent. This keeps displacement of the innate language ability of an agent and the change in plasticity independent from the location of the agent in linguistic space.

### Cultural Evolution of Language

Subsequently, the language population evolves according to four cultural processes: extinction, cultural change, division and fusion.

**Extinction** Any language that was not used by any agents in this generation do not appear in the next generation, and thus are removed from the linguistic space. This models the extinction of unused languages.

**Cultural change** The users of a language change its characteristics. We model a cultural change of language as a change in the location of the language within the linguistic space. Each agent creates, for every language it used, an attraction force (vector)  $F$  that drags the languages toward the agent's location, as shown in Fig. 4(a). The length  $F$  is  $f_{max}/N_{ul}$ , where  $f_{max}$  is the parameter that determines the maximum length of  $F$  and  $N_{ul}$  is the number of languages within the plasticity field of the agent. Each language moves to the location determined by the resultant vector of all the forces exerted on it. Fig. 4(a) shows an example of cultural change of language.

**Division** A language is divided into two languages if the forces of cultural change strongly pull it in opposing directions. Fig. 4(b) shows an example process of a language division. To determine the direction of a division of a language, we adopt "Principal Component Analysis" (PCA) (Pearson, 1901). Because the first component axis of PCA on the forces working on a language corresponds to the direction of the forces' maximum variance, we use the second component axis ( $L$ ) to divide the forces into two groups. We calculate the resultant force for each group. A division process occurs when the length of either or both resultant forces is larger than the threshold parameter  $Div_f$ . These resultant forces are then used to determine the locations of the two languages resulting from the division.

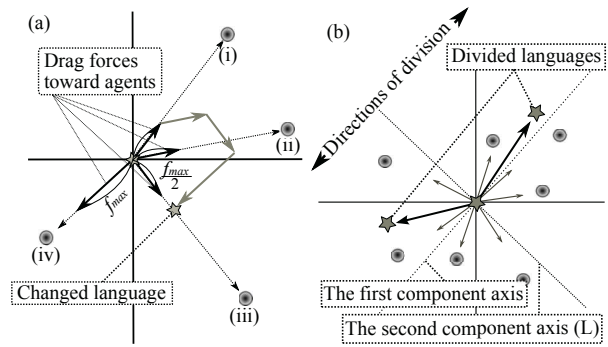


Figure 4: Panel (a) shows an example of a cultural evolution step of a language. Note that agents (i), (ii) and (iii) can each use another language (not shown) besides the focal one. Therefore, their pulling forces are only half that of agent (iv). Panel (b) shows an example of a language division process.

**Fusion** When the distance between two languages is close enough, these languages are united into one language. This process occurs when two languages' difference in distance to origin and angle to origin are smaller than the thresholds  $\beta$  and  $\gamma$  respectively.

Through the above processes, the agent population and the language population coevolve.

### Simulation Results

We conducted evolutionary experiments for 30000 generations and visualized the results in the linguistic space. The following parameters were used:  $N=2000$ ,  $w_1=1$ ,  $w_2=1$ ,  $w_3=30$ ,  $P_m=0.01$ ,  $Div_f=0.00003$ ,  $\beta, \gamma=0.001$  and  $f_{max}=0.00003$ . The initial values of  $r_A$  and  $r_L$  were picked at random from  $[0, 0.001]$ .  $\theta_A$  and  $\theta_L$  were picked at random from  $[0, 2\pi]$ .

Fig. 5 shows an example run of this experiment. We observed a typical evolution scenario, which we summarize in Fig. 6. From the initial population, both agents and languages are aggregated around the origin of the linguistic space. At this stage agents communicated successfully with each other using just a few languages at around the origin, because their innate linguistic abilities were quite similar. From there onward, we observed that (0) the number of languages rapidly increased until the 250th generation, reached to around 40 languages. This could be interpreted as a "Linguistic burst". This is thought to be due to the high concentration of agents around the origin easily leads to opposing cultural pull and hence frequent language division.

After the 250th generation, we observed cyclic coevolution processes of languages and agents. Let us look at the evolution process from the 7500th to the 13500th generation (i-iii) as example. Around 7500th generation, we see agents with smaller plasticity fields clustered densely together. In



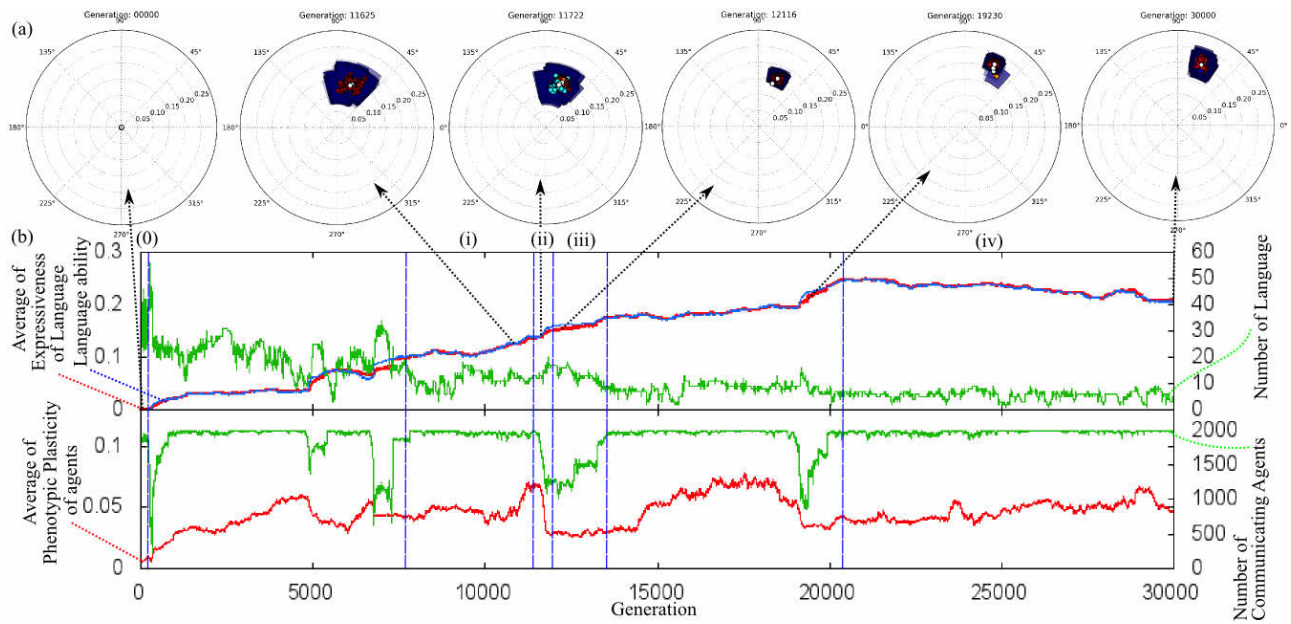


Figure 5: An example evolution process. (a) Several snapshots of coordinate distributions of languages and agents. Each agent is shown as a colored circle, with its color representing its fitness, and a fan-shaped field representing its linguistic plasticity. Languages are shown as white circles. (b) Evolution plots. The x-axis shows the generation number.

this situation, there is only weak selection pressure on innate language ability, because agents can already communicate successfully. This lack of selection leads innate language abilities to scattered by neutral evolution. This in turn leads to a gradual increase in phenotypic plasticity until around the 11500th generation (i), because additional plasticity became necessary for the agents to keep their communication successful. Then, around the 11500th generation, some agents with more expressive innate language ability and lower phenotypic plasticity appeared (ii), and occupied the population quickly. Instead of communicating with many agents in less expressive languages while incurring high plasticity costs, these agents communicate with a limited number of neighbors using more expressive languages while incurring only a small plasticity cost, which results in a net relative fitness gain. Thus, the average expressiveness of the innate language ability became larger than that of the existing languages, and the number of successful communications decreased drastically. Note that the number of languages increased because languages were dragged by two groups: the group of agents with more expressive ability of language and the group of the agents with less expressive ability of language.

After that, from the 11800th generation until the 14700th generation, the language population evolved toward the languages used by these adaptive agents, via a process of cultural evolution arising from the increased use of the more

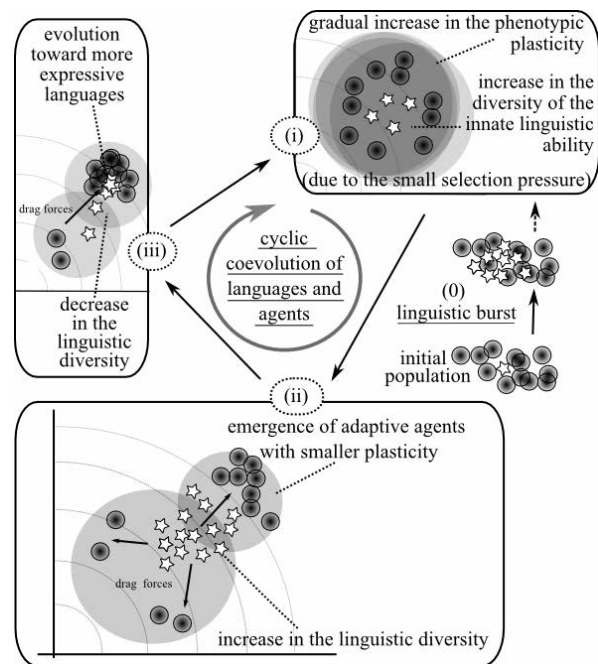


Figure 6: Typical scenario of evolution process.

expressive languages (iii), increasing the number of successful communications among agents again. Languages now

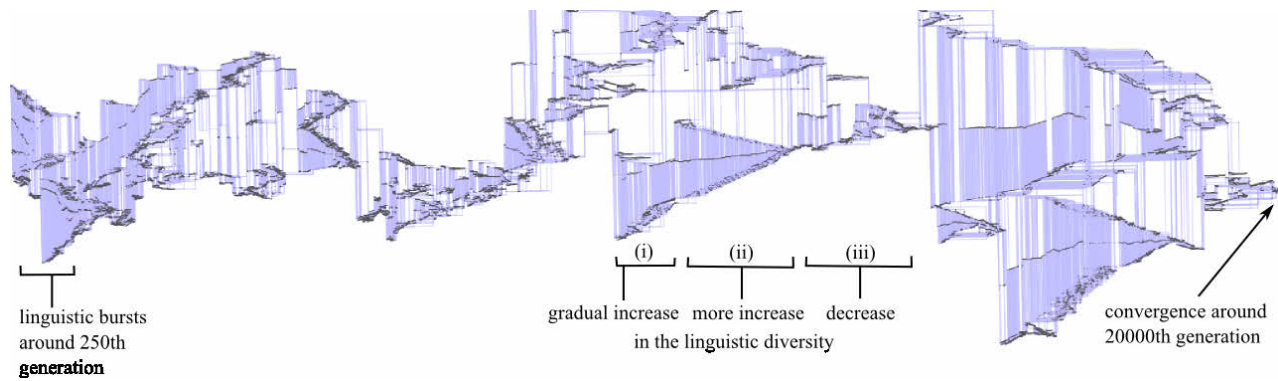


Figure 7: The generated phylogenetic network of languages (from the 250th to the 20000th generations). Each black node represents a language, and each link between the left node and the right node represents the genealogical connection between the ancestral language and its offspring language. There are 64389 nodes and 91976 links in this tree. The language division occurred 45838 times and fusion of language occurred 23802 times. The death of language occurred 22036 times.

distant from the agents' (shifted) innate language abilities go extinct, leading to a gradual decrease in the number of languages. As a result, language expressivity caught up with the expressivity of innate language ability, i.e. both the agent population and language population moved outward from the origin, and the overall system state was back at the initial state of a cycle. As this cyclic coevolution processes repeated, the expressiveness of languages and agents' innate language ability increased, while overall the number of languages decreased. This could be interpreted as the emergence of major dominant languages (Abrams and Strogatz, 2003). The cyclic mechanism can be summarized as follows: (i) The phenotypic plasticity of the agents increases gradually due to the selection for the robustness of successful communications against the increased genetic variation of the innate language ability by a genetic drift (caused by the previous step). (ii) Some agents with the more expressive innate language ability and the smaller plasticity occupy the population quickly because they can communicate using more expressive languages than other agents with the larger plasticity. (iii) The expressiveness of languages increases and the diversity of languages decreases because the languages are dragged by the agents with more expressive language ability. This brings back to the process to the beginning, because the small fitness differences due to the very few number of different languages creates the variation of the innate language ability of agents.

However, (iv) the evolution process eventually halts once the expressiveness of languages reached the high value of 0.2 after the 20000th generation. This is thought to be due to the fact that it became increasingly difficult for agents with relatively high expressiveness to maintain enough plasticity for successful communications, as the increasing cost cancels out the benefit of their expressivity.

In addition to the above analysis, we show the phyloge-

netic network of languages for our basic simulation experiment (Fig. 7). We found that a process of diversification and unification of languages emerged through repetition of the interaction processes between genetic and cultural evolution that we described in the previous section. In previous studies of language change, the phylogenetic relations of language families is generally represented as a tree (Gray and Atkinson, 2003). Fusion of languages cannot be captured using a tree representation. However the generated network showed that language fusion occurred frequently. This shows the important role the cultural processes of fusion plays in the evolution of language within our model.

Finally, we conducted experiments to study the effects of the model's parameters on the evolution process. First, to investigate the effect of learning cost, we conducted experiments with various settings of the weight on the learning cost,  $w_3$ . We found that the duration until the population reached the coevolution phase increased as  $w_3$  increased. A higher cost of learning puts the population under stronger selection pressure for low plasticity. Because individuals with the low plasticity were less robust against mutations and often failed to leave offspring, evolution speed dropped. Also, the speed of the increase in the expressiveness of languages was inversely proportional to  $w_3$ , due to the increased duration until the start of the coevolutionary phase. For example, in the case of no cost ( $w_3 = 0$ ), the duration was quite short: the coevolutionary phase started after about one hundred generations. In the case of a huge learning cost ( $w_3 = 100000$ ), the evolution of language and population stagnated around the origin, because individuals could not increase their plasticity at all. It also should be noted that higher values of  $w_3$  lead to shorter cycle period. This is thought to be due to the fact that the rapid decrease in phenotypic plasticity (ii) tends to occur more often as the cost

of plasticity increases.

We also investigated the effects of  $f_{max}$  which determines the strength of the attraction force that pulls a language toward its users. Because this parameter is used in the processes of cultural change and division of languages, we assumed the condition in which the threshold for the division  $Div_f$  was proportional to  $f_{max}$  ( $Div_f = f_{max} * 300$ ) in order to mainly focus on the effects of change in  $f_{max}$  on the process of cultural change. Experiments with different settings of  $f_{max}$  (from  $1.0 * 10^{-7}$  to  $1.0 * 10^{-4}$ ) showed that the chances of all languages dying off during the early generations increase with increasing  $f_{max}$ . This is because at large  $f_{max}$ , when there are many individuals pulling on a language, the resultant of the attraction forces tend to be so large that it displaces the language outside the plasticity ranges of the agent population. However, once the initial increase in the number of languages (0) has started successfully, we see a more rapid increase in the expressiveness of languages in the cyclic processes, because the larger pulling force facilitates rapid adaptation of in the language population. Especially in trials at high  $f_{max}$  ( $1.0 * 10^{-5}$ ), successful evolution was only observed when by chance the initial population had high plasticity. At extremely high  $f_{max}$  ( $1.0 * 10^{-4}$ ), all trials failed in about 10 generations.

## Discussion

Cultural linguistic change is often assumed to be much faster than biological change. Chater et al. (2009) showed, using a computational model, that genetic natural selection may not keep pace with a language change. Conversely, Számadó and Szathmáry (2012) argued that there are many ways for organisms to adapt to quickly changing targets, and showed numerous examples of rapid evolutionary change. For example, the rate of biological adaptation depends on the population size and genetic variation. This means that there is a possibility that biological adaptation too can be quick. Also, the phenotypic plasticity of a genotype effectively could facilitate adaptation, via the Baldwin effect. On the other hand, they also pointed out the possibility that the rate of language change could slow down. Historically the rate could have been much slower than it is now, due to smaller population sizes, slower rates of technological innovations, more limited contexts of language use, and much smaller vocabulary. Furthermore, the rate of linguistic change depends on frequency of use: words and rules used more frequently evolve far slower.

The experimental results obtained using our framework of genetic and cultural evolution of language demonstrates that the rate of language evolution could change through cyclic coevolutionary processes. This partly support Számadó et al.'s claims, especially with regards to the way phenotypic plasticity promotes adaptation. In addition to Számadó et al.'s claims, we obtained the following insights from our simulation: 1) Diversity across language groups increases

fitness variance, which accelerates the rate of biological evolution. 2) The rate of cultural evolution tends to be restricted by the plasticity of individuals, as languages cannot survive outside of the linguistic plasticity range of individuals. 3) The rate of cultural change can be slow, especially when individuals reduce their learning cost as they cluster around existing languages with sufficient expressiveness for communication. In contrast to situations with no linguistic conventions among speakers, this tends to lead language evolution to stagnate. We think that the rate of cultural change may be faster when there are no linguistic conventions among speakers, and slower when some shared conventions exist among them.

## Conclusion

This paper proposed an integrated framework for investigating genetic and cultural evolution of language. On basis of this framework, we constructed an agent-based model capturing both cultural evolution of languages and biological evolution of linguistic faculties, expressed on a two-dimensional linguistic space. Our evolutionary experiments showed that, after an initial rapid increase in the number of languages, a cyclic coevolution process occurs in which biological evolution and cultural evolution proceed in alternation. Here we observed genetic assimilation of language into innate linguistic ability. Eventually, the population reached languages with high expressiveness. Our model can be regarded as an "emergent computational thought experiments" (Bedau, 1999), or "opaque thought experiments" in which the consequences follow from the premises in such a non-obvious manner that the consequences can only understood through systematic enquiry (Di Paolo et al., 2000). We believe our model can also be extended to function as a realistic simulacra type of simulation model. For this purpose, we should further investigate what parameter settings correspond best to reality, especially in regard to the relative speeds of cultural and biological evolution.

## References

- Abrams, D. M. and Strogatz, S. H. (2003). Modelling the dynamics of language death. *Nature*, 424:900.
- Atkinson, Q. D. (2011). Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science*, 332(6027):346–349.
- Atkinson, Q. D., Meade, A., Venditti, C., J. Greenhill, S., and Pagel, M. (2008). Languages evolve in punctuational bursts. *Science*, 319(5863):588.
- Azumagakito, T., Suzuki, R., and Arita, T. (2011). Visualizing language evolution as an emergent phenomenon based on biological evolution and learning. *Artificial Life and Robotics*, 16(3):366–372.
- Azumagakito, T., Suzuki, R., and Arita, T. (2012). A simple integrated framework for investigating genetic and cultural evolution of language. In Scott-Phillips, T. C., Tamariz, M., Cartmill, E., and Hurford, J., editors, *The evolution of language* :

- proceedings of the 9th international conference on the evolution of language (EVOLANG 9)*, pages 405–406. World Scientific Publishing, Singapore.
- Bedau, M. A. (1999). Can unrealistic computer models illuminate theoretical biology? In *Proceedings of the 1999 Genetic and Evolutionary Computation Conference Workshop Program*, pages 20–23.
- Chater, N., Reali, F., and Christiansen, M. H. (2009). Restriction on biological adaptation in language evolution. *Proceedings of the National Academy of Sciences*, 106:1015–1020.
- Darwin, C. (1871). *The Descent of Man, and Selection in Relation to Sex*. John Murray.
- Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and brain*. W. W. Norton & Company, Inc.
- Di Paolo, E. A., Noble, J., and Bullock, S. (2000). Simulation models as opaque thought experiments. In *Proceedings of the Seventh International Conference on Artificial Life (ALife XI)*, pages 497–506.
- Fitch, W. T. (2011). Biological versus cultural evolution beyond a false dichotomy comment on "modeling the cultural evolution of language" by Luc Steels. *Physics of Life Reviews*, 8(4):357–358.
- Gilbert, S. F. and Epel, D. (2009). *Ecological Developmental Biology: Integrating Epigenetics, Medicine, and Evolution*. Sinauer Associates, Sunderland, MA.
- Glackin, S. N. (2010). Universal grammar and the Baldwin effect: a hypothesis and some philosophical consequences. *Biology and Philosophy*, 26(2):201–222.
- Gray, R. D. and Atkinson, Q. D. (2003). Language-tree divergence times support the anatolian theory of Indo-European origin. *Nature*, 426:435–439.
- Levinson, S. C. and Gray, R. D. (2012). Tools from evolutionary biology shed new light on the diversification of languages. *Trends in Cognitive Sciences*, 16(3):167–173.
- Lupyan, G. and Dale, R. (2010). Language structure is partly determined by social structure. *PLoS ONE*, 5(1):e8559.
- Mesoudi, A., Mcelligott, A. G., and Adger, D. (2011). Introduction: Integrating genetic and cultural evolutionary approaches to language. *Human Biology*, 83(2):141–151.
- Pagel, M. (2009). Human language as a culturally transmitted replicator. *Nature Reviews Genetics*, 10(6):405–415.
- Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine A*, 2(6):559–572.
- Peter Turney, D. W. and Anderson, R. (1996). Evolution, learning, and instinct: 100 years of the Baldwin effect. *Evolutionary Computation*, 4(3):4–8.
- Pinker, S. and Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, 13(4):707–784.
- Sayfarth, R. M., Cheney, D. L., and Marler, P. (1980). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science*, 210:801–803.
- Steels, L. (2011). Modeling the cultural evolution of language. *Physics of Life Reviews*, 8(4):339–356.
- Suzuki, R. and Arita, T. (2008). How learning can guide evolution of communication. In *Proceedings of the Eleventh International Conference on Artificial Life (ALife XI)*, pages 608–615. MIT Press.
- Suzuki, R. and Arita, T. (2012). Reconsidering language evolution from coevolution of learning and niche construction using a concept of dynamic fitness landscape. In McCrohon, L., Fujimura, T., Okanoya, K., Fujita, K., Suzuki, R., Martin, R., and Yusa, N., editors, *Five Approaches to Language Evolution (Proceedings of the Workshops of the 9th International Conference on the Evolution of Language (Evolang IX))*, pages 104–113. Evolang9 Organizing Committee.
- Suzuki, R. and Arita, T. (2013). A simple computational model of the evolution of a communicative trait and its phenotypic plasticity. *Journal of Theoretical Biology*, 330(7):37–44.
- Számádó, S. and Szathmáry, E. (2012). Evolutionary biological foundations of the origin of language: the co-evolution of language and brain. In Tallerman, M. and Gibson, K. R., editors, *The Oxford Handbook of Language Evolution*. Oxford university press, Oxford.
- West-Eberhard, M. J. (2003). *Developmental plasticity and evolution*. Oxford University Press.
- West-Eberhard, M. J. (2005). Phenotypic accommodation: adaptive innovation due to developmental plasticity. *Journal of Experimental Biology*, 304(6):610–618.
- Wund, M. A. (2012). Assessing the impacts of phenotypic plasticity on evolution. *Integrative and Comparative Biology*, 52(1):5–15.
- Zollman, K. J. S. and Smead, R. (2010). Plasticity and language: an example of the Baldwin effect? *Philosophical Studies*, 147(1):7–21.