# Implicit Memory Processing in the Formation of a Shared Communication System

**Junya Morita (j-morita@inf.shizuoka.ac.jp)[1], Takeshi Konno (konno-tks@neptune.kanazawa-it.ac.jp)[2], Jiro Okuda (jokuda@cc.kyoto-su.ac.jp)[3],Kazuyuki Samejima (samejima@tamagawa.ac.jp)[4], Guanhong li (adam.li@jaist.ac.jp)[5], Masayuki Fujiwara (m-fujiw@jaist.ac.jp)[5], Takashi Hashimoto (hash@jaist.ac.jp)[5]**

[1]Faculty of Informatics, Shizuoka University, 3-5-1 Johoku, Naka-ku, Hamamatsu, Shizuoka, Japan
[2]Information and Communication Engineering, Kanazawa Institute of Technology, 7-1 Ohgigaoka, Nonoichi, Ishikawa, Japan
[3]Faculty of Computer Science and Engineering, Kyoto Sangyo University, 458 Koyama, Kamigamo, Kita-Ku, Kyoto, Japan
[4]Graduate School of Brain Science, Tamagawa University, 6-1-1, Tamagawa-Gakuen, Machida, Tokyo, Japan
[5]School of Knowledge Science, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi, Ishikawa, Japan

## Abstract

This paper presents a simulation study focusing on implicit memory in the formation of a new communication system. In the models presented here, two agents aim to achieve their common goal by exchanging messages composed of two figures, whose meanings are not defined in advance. The effect of implicit memory has been studied with two different symbolic processes, implemented in ACT-R. Our results indicate that the difference caused by symbolic processes reduces when implicit memory is incorporated into the model. We have also found the effect of implicit memory on the creation of an isomorphic communication system, shared among agents. These findings suggest that implicit memory has some roles in the formation of a human communication system.

**Keywords:** Communication; imitation; implicit process; ACT-R

## Introduction

People try to communicate with others even when they do not share a common language. They also understand others' intentions through repeated interactions. It has previously been speculated that humans have the ability to develop a new communication system, where only limited common ground is shared, in advance. Addressing the types of cognitive functions involved in such a process will contribute to understand the origins of human communication.

Some researchers have examined this question by designing communication environments in laboratories (for a review Galantucci & Garrod, 2011; Scott-Phillips & Kirby, 2010). For example, Galantucci (2005) conducted an experiment to observe the formation of communication systems, wherein, a pair of participants communicated through a medium that restricted the use of standard communication means, such as utterances and letters. He observed the process of forming a new communication system, and discussed that implicit information was conveyed through routine behavior.

Related studies have also been conducted in the field of language acquisition. Most human infants naturally acquire languages, while a few experience difficulty. From the observations of such a developmental process, some behavioral characteristics that lead to language learning have been found. For example, Tomasello (1999) argued that a type of imitation, called "role-reversal", in which the child aligns himself/herself with the adult speaker, is essential for producing communicative symbols. The cognitive modules behind this behavior have also been discussed. Baron-Cohen (1997) hypothesized the Theory of Mind Module (ToMM) used for imitations of intentional behaviors in others. Rizzolatti and Arbib (1998) also suggests the origins of language from a viewpoint of the mirror neuron system.

For the cognitive modeling community, the challenging questions are: (1) how such modules are computationally represented, and (2) how are these integrated to a cognitive architecture that holds human-level goal management, and memory systems. Concerning these questions, several researchers have constructed a model of language evolution (Reitter & Lebiere, 2011), and an agent including the ToMM, (Stevens, Weerd, Cnossen, & Taatgen, 2016) in the general cognitive architecture.

In our previous study, we also developed a model of sharing communication systems (Morita, Konno, & Hashimoto, 2012). In our model, agents were implemented in the ACT-R cognitive architecture that posses general learning mechanisms such as reinforcement learning, and instance-based learning (Lebiere, Gonzalez, & Martin, 2007). By incorporating imitative learning into these mechanisms, Morita et al. (2012) investigated the role of imitation in the process of forming a new communication system. The results of the study indicated the importance of imitation to simulate the formation process of a human communication system.

However, in our previous work, the production rules executing imitative learning were coded manually. This does not provide the answer as to how these emerge from the human memory system. To overcome this limitation, our present study examines the process that substitutes the manually coded imitation process. This study especially focuses on the role of implicit memory processes in forming a communication system. Before presenting the details of our present study, we recapitulate concepts from our previous study.

## Task

This research simulates the experiment reported in Konno, Morita, and Hashimoto (2013), where the authors modified, and used a coordination game taken from Galantucci (2005). As in Galantucci's study, the game environment contained two characters, each controlled by a player, and four intercommunicating rooms. The game was composed of several

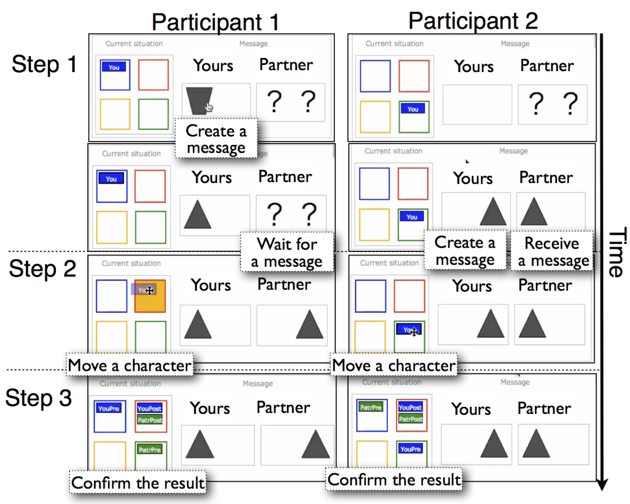


Figure 1: A single round of the coordination game consisted of three steps. In step 1, to create a message, participants selected figures by clicking the segments indicated by “Yours”. In step 2, a character (blue boxes indicated by “You”) was moved by drag-and-drop. In step 3, the result of the movement was shown to participants. Blue boxes (“You-Pre” and “You-Post”) and green boxes (“Pat-Pre” and “Pat-Post”) represent the movements made by the participant, and the partner, respectively.

repeated rounds. At the beginning of each round, characters were randomly placed in two different rooms. Players were unaware of each others’ locations, and aimed to bring their characters to the same room. The characters could not move to rooms that were located diagonally. Owing to this constraint, players needed to communicate before moving their characters.

Figure 1 presents the flow of each round, consisting of three steps: step 1 for exchanging messages; step 2 for moving characters; and step 3 for confirming the result of their movement. Among these steps, step 1 is the most crucial for the success of this task. In this step, the two players construct their own messages, composed of two figures such as “▲□”, using six available figures: ■, ◢, ▼, ◆, ▲, and □. The meanings and usages of the figures were not shared with the participants in advance. Each player could send only one message per round, but they could take turns in exchanging messages. A message sent by the first sender instantly appeared on the screen of the other player. The second sender could compose her/his message after observing the message of her/his partner (see participant 2 in Figure 1). Through such turn-taking, the first sender could transmit her/his current room location, and the second sender could transmit the destination, while taking into account the current room location of her/his partner. Participants were not assigned their roles by the experimenter; instead, they were required to self-assign their roles.

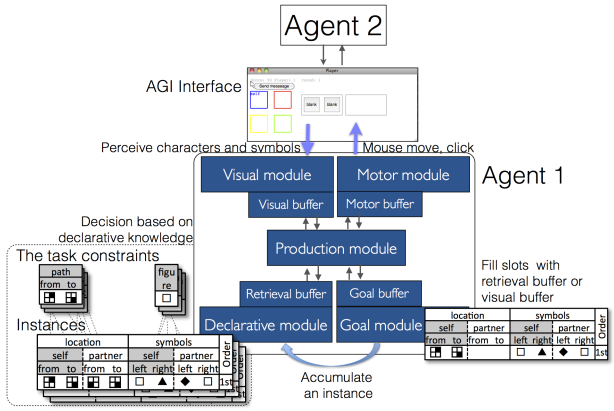


Figure 2: Schema of the model.

In the experiment reported in Konno et al. (2013), participants (21 pairs) attempted to develop a communication system within the stipulated one-hour time limit. When characters moved to the same room, players received two points, otherwise, they lost one point; although, the scores did not drop below zero. The session was terminated when the score reached 50 points. As a result, 66 percent of the participants (14 pairs) successfully reached the points in 48.42 averaged rounds. The models presented in the following section are intended to simulate the behavior of such successful pairs.

## Model

### Architecture

The task presented in the previous section requires symbolic learning for constructing a new symbol system. In addition, according to Galantucci (2005)’s report, implicit learning, which is not present in symbolic systems, possibly plays a role in this task. Morita et. al (2012) constructed a model using ACT-R (Anderson, 2007), which integrates symbolic and sub-symbolic learning mechanisms. This section illustrates how our previous study constructed a model for sharing the communication system.

ACT-R is composed of several independent modules. The modules used in our study are presented in Figure 2. Except for the production module, each module has a buffer to temporarily store information, called a chunk (a set of slot-value pairs). The production module integrates the other modules using production rules, which consist of a condition-action pair that is used in sequence with other productions to perform a task. The conditions and actions in the production rules are specified, along with the buffer contents of each module.

In the model presented in Morita et al. (2012), two independent agents interact through a simulated task environment developed in the ACT-R graphical user interface (AGI). AGI provides screens that hold visual information as chunks. In this study, the locations of the characters, and messages associated with each agent are displayed on the screen. An agent’s visual module searches for a character and stores its location

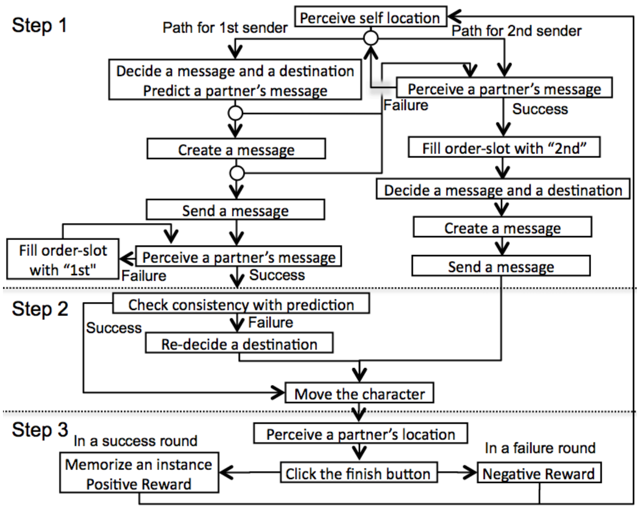


Figure 3: Process of the model. Circles indicate decisions based on conflict resolution.

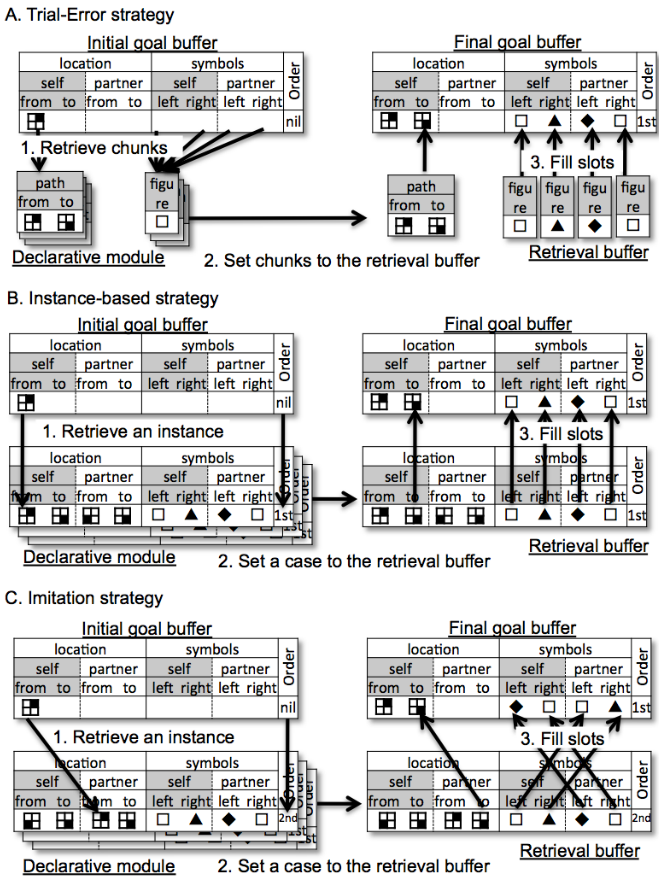


Figure 4: Three types of decision strategies.

(room) in a visual buffer. The visual buffer also stores the symbols that compose a message, attending to the screen locations where the figures appear. The simulated task environment also provides a virtual mouse to change the figures and move the characters on the screen.

Visual information stored in the visual buffer is transferred to the goal buffer through the production module. The goal buffer holds the goal of the current task, and other task-related information. Specifically, our model has nine slots for the goal buffer: four slots for storing room locations (initial (from)-destination (to) $\times$ self-partner), four slots for storing symbols (left-right $\times$ self-partner), and a slot for encoding the order for the exchanged messages.

The declarative module stores past states of the goal buffer, as instances. It also stores chunks representing task constraints such as path information indicating a room that the characters can move to (e.g., $from\_\boxplus\_to\_\boxplus\_isa\_path$), or figures the agent can use to construct a message (e.g., $\square\_isa\_figure$). An agent uses these chunks (i.e., declarative knowledge) to choose its destination and construct a message.

The productions of the model construct the process presented in Figure 3. This process is divided into three steps, just as in the original experiment (Figure 1). There are two paths in this process. The left path is for the first sender, and the right path is for the second sender. The choice of path is made by conflict resolution, which is a comparison of two conflicting productions, with noise added utilities. In each phase of the path of the first sender, there is a conflict (indicated by circles) between keeping the path of the first sender, and changing to the path of the second sender. If in any of these the agent selects the path of the second sender, the agent tries to perceive the message of her/his partner from the screen. When the agent obtains the message from her/his partner, s/he realizes that s/he is the second sender (fills the order slot with "2nd"). Otherwise, s/he resolves a conflict by waiting for the message of her/his partner and changing to the path of the first sender. This conflict loop continues until one of the agents sends a message.

## Explicit decision process

In step 1, regardless of the contents of the order slot, both agents make decisions about their destinations, and their messages. Concurrently, the first sender predicts the message that s/he will receive from her/his partner. The predicted message is checked against the message received in step 2. When the received message is inconsistent with the predicted message, the agent makes a new decision about her/his destination.

In summary, there are three situations where agents make decisions: the first sender in step 1, the first sender in step 2, and the second sender in step 2. In these situations, agents apply one of the three decision strategies shown in Figure 4. Every decision strategy begins by retrieving chunks from the declarative module, by using the current goal buffer as a cue. In the trial-error strategy, chunks concerning task constraints (chunks representing a path and symbols) are retrieved, and are used to fill in the blank goal slots. In the instance-based strategy, the agent retrieves an instance that is consistent with the current goal buffer. The retrieved instance is used to fill slots concerning the destination, and symbols. The imitation strategy also uses an instance, but the roles of an agent, and her/his partner are reversed when retrieving and filling slots.

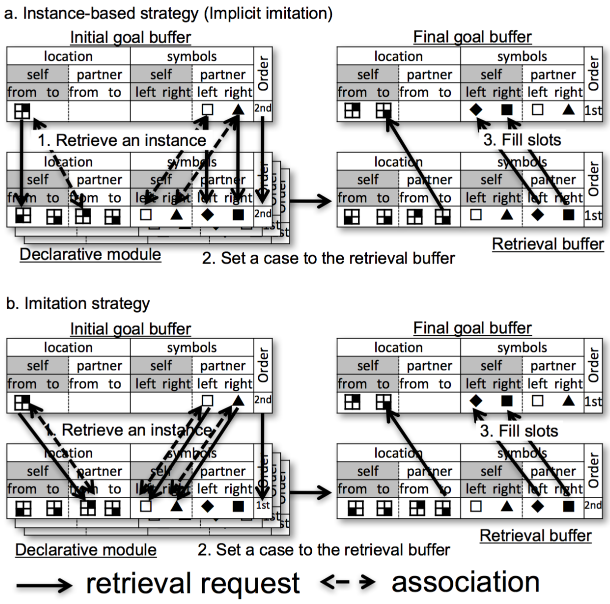


Figure 5: The implicit process of the use of instance.

## The implicit decision process

The decision strategies presented above follow a purely symbolic process. Each production rule explicitly holds the mapping of slots from the goal buffer to memorized chunks. Such a process needs different rules, which correspond specifically to each decision situation. Figure 4 only shows an example of the first sender in step 1, where the *location-self-from slot* is used as a retrieval cue. In addition to this slot, a partner's message, (the *symbols-partner-left slot*, and the *symbols-partner-right slot*) can be used as retrieval cues by the second sender in step 1. In the case of the first sender in step 2, where the goal buffer contains the message sent by the agent, more complex retrieval cues are available.

In order to maximize information sent through the exchanged message, it would be better to use instances wherein the slots are either perfectly, or partially matched to the current goal buffer. In the model constructed by Morita et al. (2012), there are rules concerning each combination of buffer slots and matching states for the two decision strategies in all the situations. However, this approach will face difficulty when the model is applied to open communication tasks, where the number of signals, or the number of turns are not decided in advance. Apparently, the model needs to have abstract mechanisms that permit the acquisition of such symbolic processes.

We have not yet solved this problem perfectly. However, in this paper, we propose another process possibly involved in forming a shared communication system, and show the behavior of this for future model development. The proposed mechanism tries to represent unintentional processes in imitation. People sometimes copy others' ideas even when it is not their intention to do so. Those phenomena have been studied in the context of source monitoring error (Johnson, Hashtroudi, & Lindsay, 1993) or deficits in self-other diffrenciations (Baron-Cohen, 1985).

We consider that spreading activation and partial matching are useful to realize such unintentional imitation. These are part of ACT-R sub-symbolic computation, which controls the activation of chunks. The spreading activation represents contextual effects caused by chunks, held by the goal buffer. The same chunks stored in the declarative module receive activation from the goal buffer. The ACT-R memory process usually retrieves chunks having the highest activation within the constraint of the retrieval cues, made by the production rule. When the partial matching process is enabled, it is possible for a chunk that is not a perfect match to the retrieval cues to be the one that is retrieved (Bothell, n.d.).

The combination of the two mechanisms characterizes the process presented in Figure 5, which presents an application of the implicit process to the two strategies by using an example of the second sender in step 1. The solid one-directed arrows connecting the goal buffer with the declarative module indicate the symbolic process noted in the production rule (retrieval cues / fill slots). The dotted two-directed arrows indicate the association connected by the spreading activation.

Importantly, in this figure, the instance-based strategy and the imitation strategy reach the same conclusion. Although the retrieval cues made by the instance-based strategy do not match the instance in the declarative module, the values stored in the slots other than the requested ones accidentally match to the state of the goal buffer. Consequently, this instance receives the high activation, and is retrieved from the declarative module. The retrieved instance is applied with a filling rule used in the imitation strategy.

The benefit of such an implicit imitative process involves reducing the complexity of symbolic processes. However, it is unknown whether the ACT-R sub-symbolic computation actually generates such imitative effects. To explore the role in the formation of human communication systems, it is needed to examine the behavior of this mechanism in a controlled simulation experiment.

# Simulation

## Simulation conditions

We first set up the following two models controlling the decision strategies presented in Figure 4.

- Instance model: In this model, the agent first tries the instance-based strategy. If the instance-based strategy fails, the agent chooses her/his destination and message based on the trial-error strategy.
- Imitation model: This model extends the instance model by adding the imitation strategy. The agent first tries to choose her/his destination and message using the instance-based strategy. If the agent fails to retrieve an instance, the imitation strategy is applied. When all other decision strategies fail, the agent uses the trial-error strategy.

In our previous study, the imitation model indicated better performance and better fitting to the human data. The imitation strategy gives the model the benefit of using instances in

Table 1: The performance indices. The numbers in parentheses indicate standard deviation.

| | Data | ExIns | ExImi | ImpIns | ImpImi |
|---|---|---|---|---|---|
| Success rates | 0.66 | 1.00 | 1.00 | 0.97 | 0.98 |
| Round | 48.42 | 70.74 | 60.21 | 115.48 | 117.33 |
| | (13.36) | (17.63) | (13.14) | (35.51) | (37.13) |

different ways. Therefore, the success rates of the imitation model are higher than the instance model in the early round.

With respect to the implicit process, we also set up the following two conditions.

- Explicit process: This model does not have the spreading activation, and partial matching mechanisms. As sub-symbolic parameters, only the activation noises, and the expected gains are set ($blc = 2, ans = 0.5, egs = 1$) to make the behaviors of the two agents differ. Except for this parameter setting, this model is same with the model presented in Morita et al. (2012)
- Implicit process: This model includes the implicit process presented in Figure 5. In addition to the sub-symbolic parameters noted in the explicit process, the matching penalty, and the maximized associative strength are set ($mp = 2$, $mas = 10$, $blc = 2$). This model also has several supplemental production rules to deal with memory errors, caused by partial matching.

Combining the symbolic, and sub-symbolic conditions, we prepared four models: ExpImi (the imitation model with the explicit process), ImpImi (the imitation model with the implicit process), ExpIns (the instance model with the explicit process), and ImpIns (the instance model with the implicit process). By comparing these, we try to identify the role of implicit processes in forming a shared communication system.

In this simulation, each model runs 100 times. In each run, the model continues the trial session for 3,600 sec[1], or until the scores reach 50 points. Following the trial session, the model is engaged in three test sessions similar to the experiment presented in section 2.

## Results

**Performance** Table 1 shows the proportion of runs/pairs whose scores reached 50 points, which is a termination condition for the session. It also presents the numbers of rounds required to reach the termination condition. Some runs utilizing the implicit model failed to form a communication system; whereas, all runs utilizing the explicit model succeeded in completing the session. Even though there were pairs that did not reach the termination condition, the number of rounds required to complete the session in the experiment (data) was smaller than that in all other models. Compared to the implicit models, the explicit models finished the session in fewer rounds. The effect of the decision strategy is only observed in the explicit models, where the imitation model finished faster than when using the instance-only strategy.

[1]We used the simulation time estimated by ACT-R.

Table 2: Fitting of the model performance to the human data.

| | ExIns | ExImi | ImpIns | ImpImi |
|---|---|---|---|---|
| RMSE | 0.11 | 0.10 | 0.20 | 0.20 |
| $R^2$ | 0.71 | 0.78 | 0.74 | 0.74 |

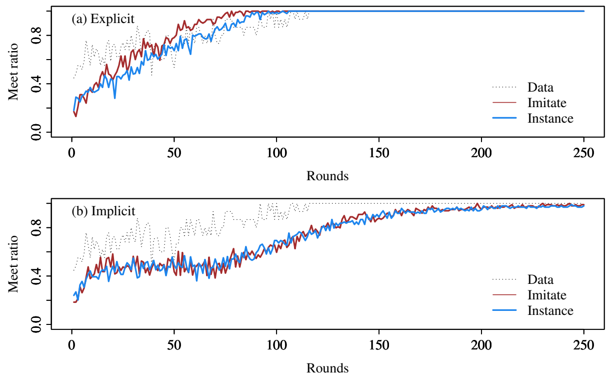


Figure 6: The ratio of sucess at each round.

The detailed processes are presented in Figure 6, which indicates the proportion of runs/pairs who met in the same room for each round. Table 2 also shows fitting indices calculated from the figure. Although the explicit models have a smaller absolute distance to the human data (RMSE) than the implicit models, there are no remarkable differences of an overall trend ($R^2$) between the four models.

**Messages** People usually try to share the same communication system even when their first languages are different. To model such characteristics of human communication, we examine the similarity of the constructed message system, as indicated by the following index.

$$Sim = \vec{M}_{player1} \cdot \vec{M}_{player2} \quad (1)$$

where $\vec{M}$ indicates a vector whose element corresponds to the use frequency of the 36 combination of figures. A dot product of the two vectors represents the degree of symbol sharing among agents.

Figure 7 indicates the moving scores of similarity with the window size of 20 rounds. Table 3 summarizes the fitting to human data, which is calculated from Figure 7. Among the four models, ExImi shows the best fit to human Data, consistent with the finding in Morita et al. (2012). It is noteworthy that models with the implicit process replicate the temporal trend of the similarity score, even without the explicit imitation strategy. The difference between the instance and the imitation models is also quite small in the implicit process.

Table 3: Fitting of the similarity score.

| | ExIns | ExImi | ImpIns | ImpImi |
|---|---|---|---|---|
| RMSE | 0.38 | 0.16 | 0.29 | 0.29 |
| $R^2$ | 0.06 | 0.72 | 0.57 | 0.64 |

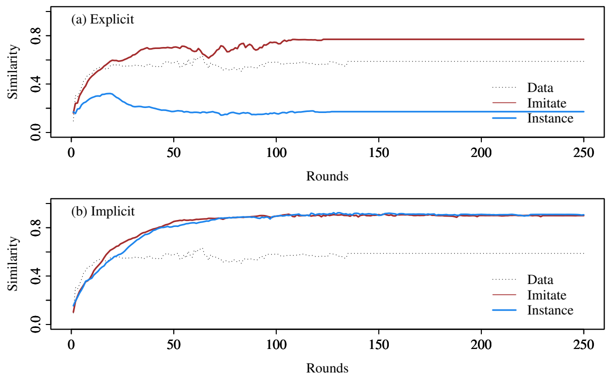


Figure 7: Similarity of messages at each round

## Discussion and Conclusions

This study constructed a model that forms a new communication system through interactive coordination. To date, many models for language evolution have been developed (for a review Steels, 2011). In addition, there exists a research that uses ACT-R to simulate experiments of forming a communication system (Reitter & Lebiere, 2011). However, such studies have not dealt with a situation with spontaneous turn-taking, or role-setting operations. Most of the previous models assign roles to agents, including being a director, or matcher, using simulation parameters.

Setting such an interactive situation, this paper examined the effect of implicit processes in forming a shared communication system. The results indicate a clear influence of the process on both the performance, and the similarity of messages. Importantly, adding the implicit process into the model, the difference caused by the explicit process almost disappeared. Although these findings alone are not enough to draw a concrete conclusion, this study shows that an isomorphic symbol system can be made without hand-coded imitations.

However, compared to human data, the implicit process results in a slower forming process, as presented in Table 1. Several explanations can be considered for this difference. The first explanation is about heuristics, utilized by human participants. Some participants in the experiment used ▲ to indicate the upper-rooms based on the shape similarity to the upper arrows. If such a pre-existing common ground is used in the model, the performance will undoubtedly increase. The other possible explanation relates to individual differences. As suggested by the failure pairs in the experiment, there are large variations in the formation process of the symbol communication system, in the collected human data. The literatures in the field of developmental psychology also indicate that children on the autism spectrum exhibit a unique language acquisition process (Baron-Cohen, 1985, 1997). Considering these factors, we can hypothesize that cognitive functions involved in forming a communication system are not determined uniquely, and the variations of the ACT-R model presented here might represent such individual differences. To examine this hypothesis, our future study will analyze the detailed behavior characteristics involved in this task. Especially, we will improve the similarity score used in this study to include characteristics of the syntax (combination rules of symbols), and symbol-meanings mappings.

## Acknowledgment

This work was supported by JSPS KAKENHI 26240037, and 16K16113.

## References

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York: Oxford University Press.

Baron-Cohen, S. (1985). *Social Cognition and Pretend Play in Autism*. Doctoral thesis, University of London.

Baron-Cohen, S. (1997). *Mindblindness: An Essay on Autism and Theory of Mind*. The MIT Press.

Bothell, D. (n.d.). Act-r 6.0 reference manual working draft [Computer software manual].

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science*, *29*(5), 737–767.

Galantucci, B., & Garrod, S. (2011). Experimental Semiotics: A Review. *Frontiers in Human Neuroscience*.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. (1993). Source Monitoring. *Psychological Bulletin*, *114*(1), 3–28.

Konno, T., Morita, J., & Hashimoto, T. (2013). Symbol communication systems integrate implicit information in coordination tasks. In *Advances in cognitive neurodynamics (iii)*. Springer.

Lebiere, C., Gonzalez, C., & Martin, M. (2007). Instance-based decision making model of repeated binary choice. *Proceedings of the 8th International Conference on Cognitive Modeling*.

Morita, J., Konno, T., & Hashimoto, T. (2012). The Role of Imitation in Generating a Shared Communication System. *Proceedings of the 34th Annual Conference of the Cognitive Science Society*.

Reitter, D., & Lebiere, C. (2011). How groups develop a specialized domain vocabulary: A cognitive multi-agent model. *Cognitive Systems Research*, *12*, 175–185.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neuroscience*, *21*, 188–194.

Scott-Phillips, T. C., & Kirby, S. (2010, January). Language evolution in the laboratory. *Trends in Cognitive Sciences*, *14*(9), 411–417.

Steels, L. (2011). Modeling the cultural evolution of language. *Physics of Life Reviews*, *8*(4), 339–356.

Stevens, C., Weerd, H. d., Cnossen, F., & Taatgen, N. (2016). A metacognitive agent for training negotiation skills. In *Proceedings of the 14th International Conference on Cognitive Modeling* (pp. 27–32).

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