Advances in Complex Systems
Vol. 15, Nos. 3 & 4 (2012) 1150020 (20 pages)

DOI: 10.1142/S0219525911500202

World Scientific
www.worldscientific.com

AN EXPERIMENTAL APPROACH TO SPEAKERS' PERSPECTIVES IN A PAIRED WALL GAME

RYOKO UNO

Division of Language and Culture Studies,
Institute of Engineering,
Tokyo University of Agriculture and Technology,
2-24-16 Naka-cho, Koganei-shi, Tokyo 184-8588, Japan
ryokouno@cc.tuat.ac.jp

KEISUKE SUZUKI

Sackler Centre for Consciousness Science,
University of Sussex, Falmer, Brighton, BN1 9QJ, UK
k.suzuki@sussex.ac.uk

TAKASHI IKEGAMI

Division of General Sciences,
The Graduate School of Arts and Sciences,
The University of Tokyo 3-8-1 Meguro,
Tokyo 153-8902, Japan
ikeg@sacral.c.u-tokyo.ac.jp

Received 4 May 2011
Revised 3 November 2011
Accepted 8 November 2011
Published 13 March 2012

A new interactive "wall game" is proposed to study the emergence of rules and symbols in interaction dynamics. In this game, two human players alternately configure a pattern on a board to communicate with each other. Distinct from related studies, players in this game have no explicit game scores or tasks to optimize. Any dynamics occurring in this game are therefore ad-hoc and on-going processes. There were three major findings in this paper. (i) The subjects mainly interacted in two modes: a dynamic mode where players proceed through the game without assigning any meanings to the pattern, and a metaphoric mode, where players process with narrative reflection. (ii) Subjects spontaneously switch between the two modes, but this switching is suppressed when playing alone. (iii) A transition diagram of the board pattern can be used to label the two modes, e.g. linearity of the diagram is correlated with the metaphoric mode. One of the main features of grammar is to display subjects' intentionality in a systematic way. We argue that the switching between the two modes observed in our experiment can be taken as a grammatical aspect that emerged in the process. These modes express the speaker's perspective in the same manner as grammatical elements do in natural language. The switching behavior should be seen as a process that embodies a player's intention using

the medium (in this case, the patterns in the wall game), and a player's exploration of the medium is a necessary step before generating a grammar structure.

Keywords: Communication; grammar; speaker's perspective; evolutionary linguistics; media; mental spaces.

1. Introduction

Many cognitive linguists, following Langacker's [13] symbolic view of grammar, investigate meanings expressed with grammatical elements (such as auxiliaries or articles) rather than those expressed with lexical items (such as nouns and verbs). For example, Cutrer [3] argues that among grammatical elements, tense and epistemic modality markers express the speaker's perspective or construal. (This can be viewed as a new interpretation of classical analysis of tense by Reichenbach [15]). For example:

(1) I am writing a paper on language and movement.
(2) If I were a linguist, I would write a paper on language and movement.

Cutrer's analysis is based on the terminology of Mental Space Theory Fauconnier [4]. In brief, a "mental space" is a subjective version of a "possible world". Mental spaces are used to describe how the speaker subjectively manages the type of information and perspective shifting during the discourse. In sentence (1), the event "my writing a paper" is located in "now" and "here" space by the progressive form of the verb. On the other hand, the event "my writing a paper" in sentence (2) is located in a mental space which is in a domain of irreality as demarcated by the epistemic modality "would". In both cases, what the grammatical items indicate are the locations of the events in the speaker's mind. In other words, grammatical elements indicate how the speaker locates the events, i.e. the speaker's point of view.

We can also indicate the speaker's perspective by the usage of language *per se*. The sentence, "Because the sky is blue, I feel like crying" can be taken as an example that shifts a listener's attention from the content of the sentence to the speaker's intention. An association is established between blue and sadness not by the grammatical structure, but by the speaker's intention. This happens because in this example the word "because" does not express reason or causation as it does in typical sentences. The decision making process of the speaker stands out more than the literal content of the sentence.

What has to be noted is that a speaker's perspective can be indicated not only by grammatical elements, but also by usage of linguistic items. We assume that at the beginning of language, what is now expressed with tense and epistemic modality had been expressed with usage. What we aim in this paper is to see the emergence of this proto-tense and proto-epistemic modality as a mode of communication.

In this paper, we will therefore analyze how speakers manipulate symbols and meanings even within a simple board game situation. We took an experimental

approach to see how the structures of communication emerge during interpersonal interaction. We asked subjects to communicate using our communication tool, which is called a "wall game". The analyses of the results show that there are two modes of communication. What emerged from the subjects' communication is not a set of labels, but the way in which intentions are embodied in actions. We argue that these two modes relate to the distinction between the two types of speaker perspectives which are linguistically represented in examples (1) and (2). The modes we discovered indicate the types of speakers' perspective through usages and may thus be considered to be proto-grammatical structures.

2. Related Studies

Artificial life studies provide a test bed for exploring how symbols and grammar emerge in minimally interacting systems through computer simulation. For the last 10–15 years, artificial life studies have contributed greatly to this, and the origin and evolution of language has become the target of many scientific studies (see e.g. [2, 8, 16, 19, 22, 26]). For example, Steels and Kaplan [21] have developed a platform for studying the interaction between two artificial agents acting as speaker or listener. In this approach, a population of robots develops a shared vocabulary and a corresponding ontology while playing language games (i.e. ritualized social interactions that follow a specific script).

More recently, as a new approach to evolutionary linguistics, there has been a great deal of research based on experiments using human subjects (e.g. [18, 20]). Subjects in this type of study communicate through a communication tool and a structured system emerges. Some of these studies testify to a hypothesis that is raised by computational simulation studies. For example, the "iterated learning model", which is a model of vertical cultural transmission, was proposed by Kirby [9]. The iterated learning model was originally studied as a computational model and later it was adjusted to experiments using humans [10].

Among many studies of "language evolution in the laboratory", Galantucci [5] introduced one of the most influential experiments. In his experiment, two subjects (who are staying in different rooms) play a video game together over a monitor. They have to be cooperative to get a high score. They are allowed to communicate using a special communication tool. This tool allows the subjects to draw graphics, but not letters. As the experiment proceeds, the difficulty of the video game increases. The pairs that ended the game successfully shared many signs for rooms and enemy, which were drawn with the communication tool.

In most of the research that adopts an experimental approach, the final outcome usually consists of lexical items (labels of entities or events). This stems from the fact that in most experiments, subjects are asked to perform a task together to make them communicate with each other (see Table 1, column A). Not only is lexicon an integral part of a linguistic system, but so too is grammar. To get more variations

in results, the communication observed in the experiments should not be limited to those that are task oriented. For example, we assume that "communication without purpose" can be important in triggering proto-language with both grammar and lexicon in experiments. As noted in the previous section, this is based on the idea that a speaker's perspective can be displayed when the atypical use of symbols occurs. In this case, rather than transferring information, sharing the speaker's perspective with the listener becomes more important.

This position is also supported by research in developmental psychology: infants are known to be engaged in two types of proto-linguistic communication. The first is communication with an aim, such as those that are task oriented. The second is communication without an aim, in other words, communication whose aim is communication itself [1]. Gómez, Sarria and Tamarit [7] argue the importance of the second type of communication. It is pointed out that the ability to communicate without purpose is an indicator for the ability called "theory of mind". With "theory of mind", one can infer other people's minds, which are different from one's own. And this ability is known to be integral to the proper use of grammar [23].

Uno, Marocco, Nolfi and Ikegami [25] made an attempt to use the artificial life approach to explore the relationship between communication without purpose and the emergence of grammar. The agents were supposed to stay together in the target area using signals. However, when agents were given uncertain information regarding the target area, they started staying together outside the target area using newly created signals, which was argued to have proto-grammar: a sentence used to share intentionality.

Most of the research mentioned above have an interest in the emergence of grammar, not only in the lexical items (see Table 1, column B). However, what they imagine grammar to be, differs. This paper aims to see the emergence of the display of a speaker's perspective as one part of grammar. Uno, Marocco, Nolfi and Ikegami have the same view of grammar. In the case of Steels and Kaplan, the intended grammatical element has to do with perspective encoding, which is also very close to the view held in this paper. Galantucci does not clearly state his grammatical view, however, some of the symbols seem to have an embodied grammatical property and this is brought about by the embodied nature of the task played by the subjects. For example, symbols that mean "go left" rather than "left" have almost the same function as an imperative sentence. On the other hand, Kirby regards the emergence of rules as the emergence of grammar. For example, a rule which produces regular and irregular past tense forms. In this view, the grammar is clearly separated from embodiment and is close to Chomsky's position.

Embodiment therefore becomes another factor that characterizes each study (see Table 1, column C). In most of the simulation studies, evolved signals are disembodied, but in Uno, Marocco, Nolfi and Ikegami's study, the movement of the robots and the way the signals are structured are tightly coupled (also see [14]). It is easier to include embodied aspects in experimental approaches. Not only the task, but also the media used to communicate with, has some characteristics that

connect embodiment and symbols. This is not significantly discussed in this paper, but Galantucci's unique communication tool must have had a huge impact on the way symbols have emerged (see Table 1, column D).

The final point to be considered is the sustainability of communication (see Table 1, column E); that is why and how communication can last for a certain amount of time. We assume that the natural interest to others, which is an unpredictable existence [6], contributes to sustainability.

Whether the existence of two people in communication contributes to the emergence of grammar is one of the issues that is not often discussed in the field of evolutionary linguistics. In this paper, to reveal the relationship between embodiment and grammar, we are going to take an experimental approach to see how communication gets structured. In our experiment, no special task to solve (such as playing games) is assigned to the subjects.[a] They are just asked to communicate using our communication tool, which is called a "wall game". Our main observation is focused on how the subjects interact with the tool and with each other. We explore the characteristics of human communication (which might possibly be implemented in artificial systems) when individuals are just having fun.

The results show that there are two modes of communication. What emerged from the subjects' communication is not a set of lexical items. We argue that

Table 1. Features of related research. We put ◯ when the research investigates the topic, × when it does not. The symbol △ is used when the research is relevant to the topic, but when the topic is not explicitly investigated. For the meaning of each column, refer to the main text.

	A		B	C	D	E
	Label	Task	Embodied grammatical view	Embodiment	Media	Sustainability
Steels and Kaplan (2001) [Ref. 21]	◯	◯	◯	×	×	×
Kirby (2002) [Ref. 9]	◯	◯	×	×	×	×
Galantucci (2005) [Ref. 5]	◯	◯	△	◯	◯	×
Uno, Marocco, Nolfi and Ikegami (2011) [Ref. 25]	△	△	◯	◯	×	△
This paper	×	×	◯	◯	◯	◯

[a]We admit that in any experiment, subjects are constrained to some extent (e.g. subjects are required to use a communication tool or write a report). The expression "without a task to solve" does not mean that subjects are unconstrained. It means that they are not forced to use communication to convey a certain type of information because of the restricted communication goal (such as getting a high score in the game cooperatively). Even if one of the participants wants to communicate for a restricted purpose which is not given by us, it is almost impossible to share the purpose with the other participant through the communication system in this experiment.

the modes we got can be seen as a proto-linguistic grammar, since they have the primitive function of tense and epistemic modality.

Section 3 explains the basic design of our experiment. Sections 4 and 5 show the results of the experiments. Finally, Sec. 6 analyzes and discusses the results.

3. Description of the Experiment

Subjects were asked to communicate using an artificial communication system, where the expressions were spatial patterns of triplets in a 3-by-3 bit square. Subjects were allowed to rewrite the pattern alternately. We call this pattern a "message". Our instruction to the subjects was: "Please communicate with your partner using this tool. Please look at the *message* (= pattern) sent by your partner and compose your *message* as an answer to the preceding *message*".

The number of the subjects was twenty-six (13 pairs). Among them were 10 females and 16 males. The ages of the subjects ranged from 21 to 36.

For the first 9 pairs (Pair 1 to Pair 9), each subject sent 8 messages in turn, which is 16 messages in total. For the next 4 pairs (Pair 10 to Pair 13), in one set each subject sent 15 messages making 30 in all. These 4 pairs played two sets, that is, they exchanged 60 messages. In addition 4 players from Pair 9 to 13 are asked to play the game by themselves. They made an additional 30 messages by themselves.

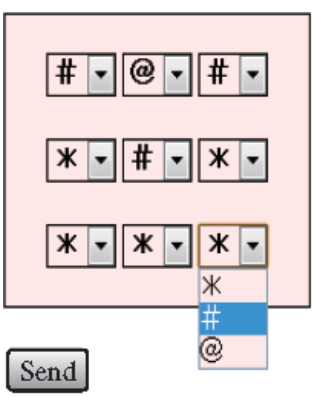
We asked the subjects to report their intentions behind the sent messages, and their interpretations of the received messages in natural language (we call this data the "intention report"). All pairs (for the first set if they played two sets) were asked to write intention reports after the messages were exchanged. In this case, during the game, subjects were not informed that they were going to write intention reports afterward. For the last 4 pairs, in the second set, we asked them to write an intention report every time they exchange messages.

We conducted the experiment in Japanese. The reports shown in this paper were translated into English by the authors.

The two subjects of each pair stayed in different rooms. The messages were sent to each other over the Internet. The left-hand figure in Fig. 1 shows a screen where one can compose messages. All the messages that were sent and received were shown to the subjects so that they could compose their messages based on their communication history.[b] The right-hand figure in Fig. 1 shows how the history of exchanged messages was displayed to the subjects.

Here are some examples of exchanged messages from our data in Table 2. Player A sent (1) and Player B answered with (2). Then Player A replied with (3). Finally, (4) is an answer to (3) by Player B.

[b]We think it is important to show the history of messages in this experiment. How the new message differs from the preceding one is the most important factor in composing a message when the timeline is shown, as in our experiment.



1	2010/7/24 13:18:06	✳✳✳ ✳✳✳ ✳✳✳
2	2010/7/24 13:25:32	@✳@ ✳✳✳ @✳@
3	2010/7/24 13:26:45	@✳# ✳✳✳ #✳@
4	2010/7/24 13:27:39	@✳# ✳#✳ #✳@

Fig. 1. Two screen shots. The left is a message composer. The right is the timeline of exchanged messages.

Table 2. Exchanged messages between Player A and Player B.

(1) From A to B	(2) From B to A	(3) From A to B	(4) From B to A
@@@	@@@	###	###
@*@	@*#	###	##@
@@@	@@@	###	#@@

We performed a linguistic analysis of the intention report and a mathematical analysis to the patterns. The results are given in Secs. 4 and 5.

In summary, we performed the experiment under three conditions:

Condition 1. (Appears in Experiment 1, 2 and 3)

- Two players. They exchanged messages.
- Intention reports were written after exchanging all messages.

Condition 2. (Appears mainly in Experiment 2)

- Two players. They exchanged messages.
- Intention reports were written every time they sent or received a message during the game.

Condition 3. (Appears in Experiment 3)

- One player. The player composed messages by himself or herself.
- Intention reports were written after composing all the messages.

The basic game is Condition 1. In Condition 2, the timing of writing the report differed from that in Condition 1. In Condition 3, the game was played by a single player.

4. Experiment 1: Two Modes of Communication

4.1. *Linguistic analysis*

After collecting and interpreting the data from the 13 pairs playing this game, we noticed that there were several patterns that we could categorize from a linguistic point of view. We had three categories of reports: dynamic reports, metaphoric reports, and others.

What we call a "dynamic report" was a literal description of the patterns in the messages. For example:

(D1) All kinds of symbols are used.
(D2) The pattern is scrolled from left to right.

In these reports, the patterns are described just as they were. The messages that these reports are made for, are shown below or in Table 3. For example, the report (D1) (D2) above is an intention report of the message shown in (D1-M) (D2-M) or (D1) (D2) in Table 3 below.

(D1-M)
*#@
*#@
*#@

(D2-M)
@#*
@#*
@#*

On the other hand, in what we call a "metaphoric report", the subjects created a story based on the symbols inside the pattern and were not describing the pattern as it was. Instead, they were using metaphors (in the sense of [12]) to update patterns for exchange. They described the symbols or a string of symbols appearing on the board as metaphors for stories that the players spontaneously produced. Here are some examples:

(M1) A rabbit is in a cage.
(M2) The rabbit made a hole in the cage to escape.

Table 3. Examples of messages that the three types of reports are made for. (D1-2) had dynamic interpretations, (M1-2) had metaphoric ones, and interpretations of (O1-2) were categorized as others.

(D1)	(D2)	(M1)	(M2)	(O1)	(O2)
*#@	@#*	@@@	@@@	@*@	@##
*#@	@#*	@*@	@*#	*@*	#*#
*#@	@#*	@@@	@@@	###	***

And here are the messages:

(M1-M)
@@@
@*@
@@@

(M2-M)
@@@
@*#
@@@

Here the player saw the symbol "*" as a rabbit and a sequence of "@" as the cage, and "#" as a hole.

In this system, there is no way for one player to transmit his/her story to the other player. For example, while Player A intended to express a rabbit using the message shown as (M1) in Table 3, Player B made the following intention report for the same pattern:

(D3) * is surrounded by @.

(D3-M)
@@@
@*@
@@@

In the category called "others", the reports were not strongly connected to the texture of the patterns. For example, we have emphatic expressions such as (O1) or feelings of the players, which are irrelevant to the patterns such as (O2):

(O1) Hello. Nice to meet you.
(O2) This experiment is difficult.

(O1-M)
@*@
@
###

(O2-M)
@##
#*#
***

Figure 2 shows the ratio of reports in each category used by each pair. It is shown that most of the reports are either dynamic or metaphoric. We try to quantify these categories with mutual entropy and a transition table analysis. The interpretation is in the following sections.

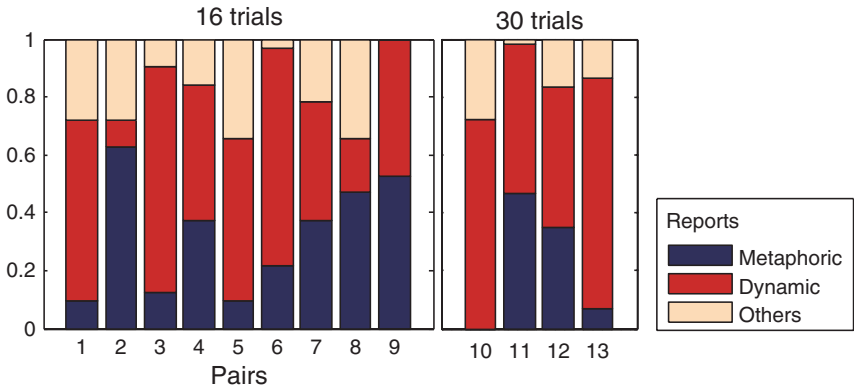


Fig. 2. Ratio of report types in Experiment 1. This is calculated from the accumulated reports of two players who exchanged messages. The ratio between metaphoric and dynamic reports varies over pairs. The total number of game iterations is different for each pair.

The following (Fig. 3) shows the interpretation used by the same individuals across the trial. Only metaphorical (blue) and dynamical (red) interpretations are shown to see the contrast between the two interpretations.

We will be discussing the data in Fig. 3, in Secs. 5.1 and 5.2.

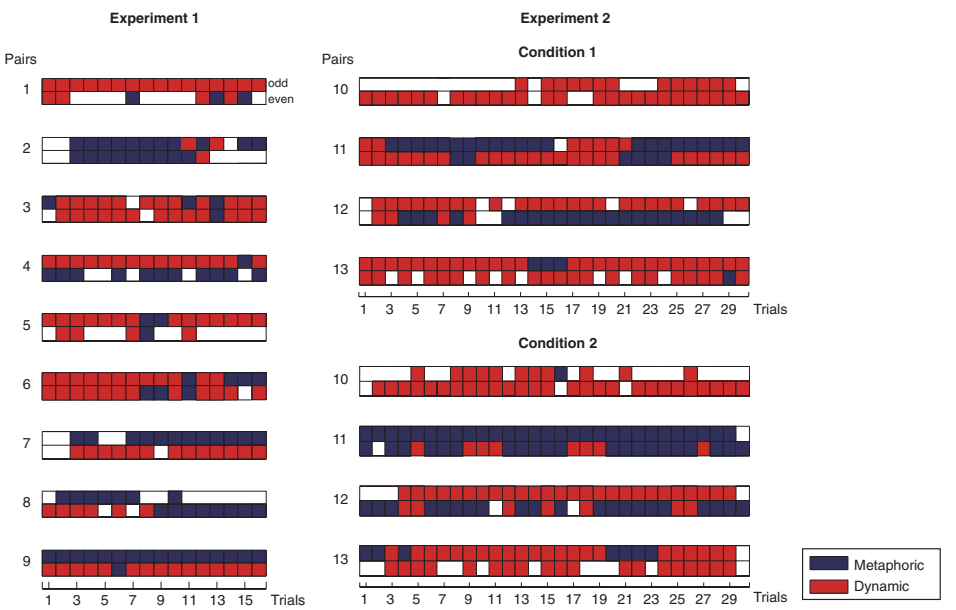


Fig. 3. (Color online) Temporal changes of report types in Experiment 1. The color of the boxes represents the types of reports accumulated from the both players. The types of reports are classified into "Metaphoric", "Dynamic" and "Others" as we have noted in Fig. 2. In this figure, we show only the former two types to see the contrast between them. The upper boxes express the usage of Player 1 and the lower boxes express that of Player 2.

4.2. *Mathematical analysis 1*

In order to see the characteristics of the wall patterns in the two different report categories, we calculated the correlation between the Hamming distance of adjacent patterns and the frequency of each type of report. The Hamming distance is defined as the number of changes required to match one character string with another string. Therefore, we regarded the wall patterns as linear character strings to calculate it. For example, (D1-M) below from Table 3 is regarded as "*#@*#@*#@".

(D1-M)
*#@
*#@
*#@

The larger the Hamming distance of a pair of patterns, the less they are similar. In order to treat the report under mathematical analysis, the two categories of the report are indexed by counting the number of them in each turn (i.e. the metaphoric index was scored 2 for when both subjects interpreted metaphorically, 1 for when one subject did, and 0 for when neither did). We calculated the Pearson correlation coefficient between the Hamming distances of the successive patterns and the mode indices.

The results are shown in Fig. 4. We found that a small Hamming distance correlates with the use of metaphoric reports. On the other hand, a larger Hamming

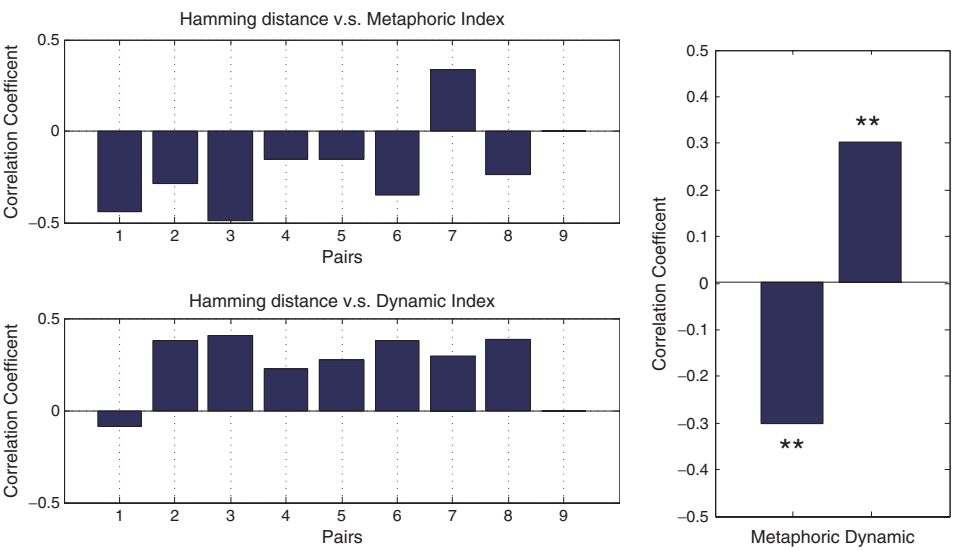


Fig. 4. (a) Correlation coefficient between the Hamming distance and the number of the two categories in reports for each pair. (b) The same evaluation across all the pairs. The Hamming distance has significant ($p < 0.01$) positive correlation with the dynamic report, and significant ($p < 0.01$) negative correlation with the metaphoric report.

distance correlates with dynamic reports usage. We will interpret this characteristic correlation in the following sections.

4.3. *Mathematical analysis 2*

We also drew a state transition graph between successive patterns. To analyze the transition state of the patterns, 16 messages were not enough to have statistically valid results. We therefore focused on Pair 10 to Pair 13 who exchanged 30 messages (15 each) in one trial. They performed two trials under two different conditions (i.e. Conditions 1 and 2. See Sec. 3 for an explanation). We are going to discuss the contrast between Condition 1 and 2 in detail in Sec. 5. The point is that under Condition 2, more metaphoric reports tended to be used compared to Condition 1.

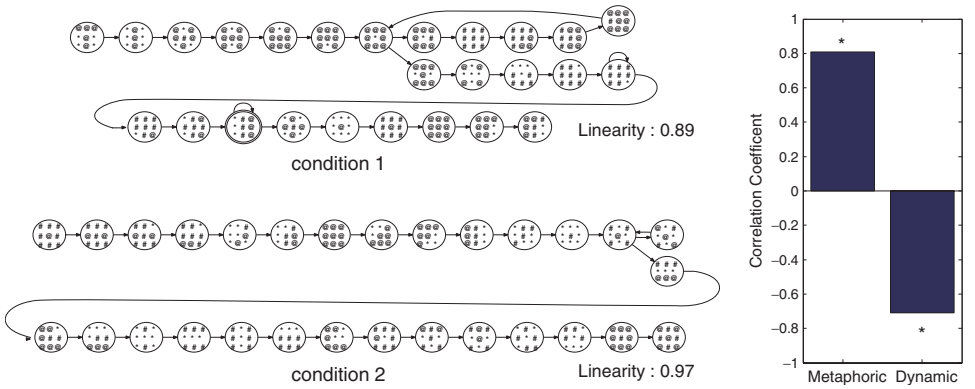
In order to create the transition graph, we grouped the patterns used in a game by the numbers of symbols the pair used. We first separated the patterns into three rows, and grouped each row using only the constituent ratio of the symbols (e.g. "*@*" is grouped into "210". "210" is derived from the fact that there are 2× "*", 1× "@" and 0× "#". In the same way, "###" is grouped into "003", etc.). The total number of possible combinations is 10, thus, each row is sorted into 1 of 10 groups (0 for "012" ...9 for "210"). We then assigned the groups a triple-digit number (e.g. 091 for "#@#/*@*/@#@"). Finally, we grouped all the patterns used in the game from the number, and calculated the transition between them.

Figure 5(a) shows the state transition graph calculated for Pair 11. The linearity of the transition graphs is defined as follows: "the number of nodes minus 1 divided by the number of edges of a graph". The linearity of the Pair 11 under Condition 1, whose main communication mode is dynamic, is calculated as 0.89, and the linearity of Pair 11 under Condition 2, whose main communication mode is metaphoric, is 0.97. The analysis of this pair suggests that the dynamic mode has a tendency to have a lower linearity than the metaphoric mode.

Figure 5(b) shows the correlation coefficient between the linearity and the number of the two report modes in all the trials by four pairs. The result shows that the same types of transition are used repetitively in the dynamic mode but not in the metaphoric mode.

4.4. *Summary*

We found and characterized two major communication modes: the metaphoric mode and the dynamic mode. During the game, subjects enjoyed processing patterns and trying to assign meanings to them. The dynamic mode is associated with a report that contains literal descriptions of dynamic patterns, and the metaphoric mode is associated with a report that uses a metaphor and narrates a story based on this metaphor. This difference can be measured by the Hamming distance and the linearity of the transition diagram.



(a) (b)

Fig. 5. (a) Examples of state transition graphs obtained from results of the Pair 11. Each node represents group of patterns that have same ratio of symbols, that is, "***/###/@@@" is in the same group of "###/@@@/***". The pattern images drawn in the nodes depict examples of the pattern for the groups. Edges between the nodes represent transitions between those groups of patterns. Here, the linearity of the state transition graph is defined as (the number of nodes − 1)/(the number of edges). A higher Linearity is observed in Condition 2, compared with Condition 1. (b) The correlation coefficient between the linearity and the number of the two report modes in the eight games. A positive correlation can be seen between the linearity and the metaphoric mode, and a negative correlation between the linearity and the dynamic mode.

5. Experiments 2 and 3

To further investigate the two modes of communication pointed out in the last section, we designed two additional experiments as below.

5.1. *Experiment 2: Message-by-message report*

We asked the four pairs who had exchanged 30 messages in Experiment 1 (which we call Condition 1) to exchange an additional 30 messages in the new trial (which we call Condition 2). For descriptions of the conditions see Sec. 3. This experiment is to make the intentions behind the messages clear and to see the effects on the subjects' behavior. In Experiment 1, whilst exchanging messages (i.e. patterns), the players were not informed that they would have to report the intentions afterward. In Experiment 2, every time the player received or sent a message, he had to write an intention report on the spot. Because of this difference, we assume that the player becomes more conscious about the content of their messages.

Compare the ratios of the metaphoric and dynamic reports of Conditions 1 and 2 shown in Fig. 6. This result suggests that when subjects are more conscious of the intention of the message, they tend to be engaged in the metaphoric mode rather than the dynamic mode.

The variability of style by the same individuals is shown in Figs. 3 in Sec. 3, which shows that the increase of metaphoric reporting can be observed in all pairs.

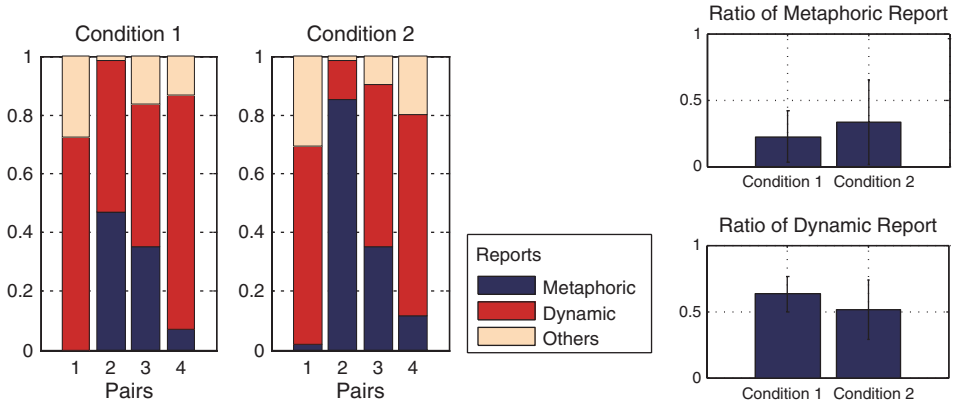


Fig. 6. (a) Ratio of reports in each category under Conditions 1 and 2. In Condition 1, the players did not have to make "intention reports" during the game. In Condition 2, the players had to make "intention reports" every time they sent or received messages. It is calculated in the same manner as in Fig. 2. (b) The ratios of the metaphoric reports and dynamic reports are averaged across all the pairs. There seemed to be a tendency for dynamic reports to be more prevalent than metaphoric reports under Condition 1, while the opposite tendency can be seen under Condition 2.

5 players out of 8 had more metaphoric reports under Condition 2 than under Condition 1.

5.2. *Experiment 3: Solitary play*

In the third experiment, we asked each subject to play the game by him/herself (we call this Condition 3). We asked one of the players in each pair who experienced Experiments 1 and 2 to make 30 messages by him/herself without having another player "behind the wall" asking him/her to report his/her intentions. Compare Condition 1 in Fig. 5 and Condition 3 in Fig. 7. The result reveals that subjects tend to use either one of the modes, not both of them, when they have no one to communicate with.

We think that the players show variations in behavior when they are communicating with others. To get a better idea of the results in Experiment 2 let us look at a real exchange of messages in that experiment. Table 4 shows an exchange between two players (Pair 9, Messages 3–6, in Fig. 3. A is Player 2 and B is Player 1).

Player A interprets the whole exchange in a metaphoric mode. Below are the intention reports by Player A from (1) to (4).

(A1) @ is me and * is a cherry blossom. Shall I go out by myself? (Report of intention behind A's own message).
(A2) I am also alone (Report of A's interpretation of B's message).
(A3) It is more fun if we stay together (Report of intention behind A's own message).

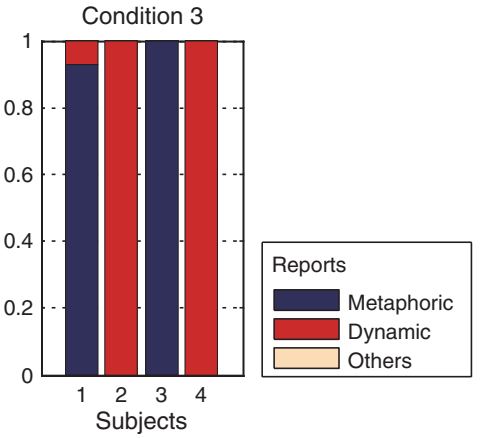


Fig. 7. Ratio of reports in each category under Condition 3. Under this condition, subjects play the wall game alone. The reports are extremely biased into either metaphoric or dynamic.

Table 4. Exchanged messages between Player A and Player B.

(1) From A to B	(2) From B to A	(3) From A to B	(4) From B to A
#**	###	***	**@
#@*	#@#	*@@	**@
###	###	###	###

(A4) Different scene. Here * and # are people. @ and @ joined them (Report of A's interpretation of B's message).

On the other hand, Player B is communicating with the dynamic mode from (1) to (3). At (4) he starts to use the metaphoric mode. Here are the intention reports by Player B for (1) to (4):

(B1) More #s (Report of B's interpretation of A's message).
(B2) I added more #s (Report of intention behind B's own message).
(B3) @ was added. * appeared again (Report of B's interpretation of A's message).
(B4) @ looks like a cute animal. # and * are environments. So I moved @ (Report of intention behind B's own message).

Player B changed the mode of communication after communicating with Player A.[c] As shown in this sample, an interaction between two players facilitates

[c]After (B4) which corresponds to the 6th message in Pair 9 in Fig. 3, all the reports of B is categorized as dynamic mode. However, we can observe a clear change in a tone of messages before and after the 6th message. For example, while the central report of the message is dynamic, Player B adds his comments that symbols can be seen as a living entity which has a flavor of metaphoric report. This observation shows that more than the data in Figs. 3 and 6 show Player B is influenced by the metaphoric mode used by Player A.

Table 5. Summary of experimental results.

	1			2	3
Experiment	Intention report	Hamming distance	Transition state	On-spot report	Alone
Metaphoric mode	Metaphoric	Smaller	More linear	Increase	N/A
Dynamic mode	Dynamic (literal)	Larger	Less linear	Decrease	N/A

switching between two modes. Namely, when a player shifts to, for example, the metaphoric mode, the other player will also use the same mode.

5.3. *Summary*

All of the experiments show the characteristics of two modes of communication. We summarize the findings in the table above (Table 5). We interpret the results in the next section.

6. Analysis and Discussion

6.1. *Interpretation of two modes*

Distinct from previous studies on evolutionary language, players of the wall communication game in the present experiment were asked to freely communicate with each other. No explicit purpose of the game was assigned. Therefore, we expected the outcome to be self-organized and self-sustainable.

In this game, the easiest way to compose a message is to mimic what the other player did in the previous round. However, this strategy is discarded as the communication becomes monotonous and predictable, so that the players cannot sustain the game itself. Accordingly, we assumed that players would try to avoid mimicking each other and that they would nee d to send messages that invite novelty for the other player.

Two modes of communication can be understood from this point of view. With a dynamic mode, the player interprets messages as textures rather than as content. For example, symmetrical patterns, rather than asymmetric ones, are more preferred. Therefore, the reports are literal descriptions of the patterns (dynamic report). To make an interesting change in messages only with patterns, there must be a distinct change at each time step or over several time steps. This explains why the Hamming distance calculated in Experiment 1 is relatively large. The frequently used patterns that can make interesting transitions are limited. For example, patterns with three lines are often used, as is shown in the transition from (D1-M) to (D2-M) in Sec. 4.1.

(D1-M)
*#@
*#@
*#@

(D2-M)
@#*
@#*
@#*

This explains the result of Experiment 1 in Sec. 4.2 which shows that the transition graph of the dynamic mode was less linear, which means that the same pattern was frequently used.

Turning to the metaphoric mode, the players make their own stories based on the transition of the patterns. The story itself cannot be transmitted to the other player in this game. So for Player B, who does not share the story, the message by Player A in the metaphoric mode is unpredictable and novel. It has been pointed out that metaphor helps people extend their understanding [12] and make inferences [24]. In addition, we want to point out that by using metaphor, people can come up with a message that is unpredictable for the receiver. This can be seen as a case where metaphor helps people behave in an autonomous way in communication.

In the metaphoric mode, as shown in Experiment 1, the Hamming distance between the successive patterns is relatively small. This is understandable when we realize that even small changes can be meaningful in a story. Let us compare (M1-M) and (M2-M) in Sec. 4.1.

(M1-M)
@@@
@*@
@@@

(M2-M)
@@@
@*#
@@@

As shown in Sec. 4.2, the linearity of the transition graph is relatively large, that is, the same patterns are rarely used. This can be explained by the fact that in metaphoric mode, what is meaningful is the difference between the current diagram and the last one. This means that there is no particular pattern that has to be used in the metaphoric mode.

When the players are alone, they tend to use either of the modes. When two players are together, both modes occur in communication. This suggests that coexistence of the two modes is enhanced by communication.

6.2. *Grammatical element, mental spaces, and dynamic patterns*

The outcome of the wall game experiment is that players switch behavior between the two modes. These two modes together form a procedure of taking our inner thoughts and feelings and then expressing them externally through the media, i.e.

in this case, the wall game. What we got was apparently not lexical items (labels of entities or events) but two modes of communication. We believe it can be seen as a primitive form of grammatical elements. It helps us display the perspectives to the other player in a structured way.

Let us go back to the examples we referred to in the introduction.

(1) I am writing a paper on language and movement.
(2) If I were a linguist, I would write a paper on language and movement.

In the first sentence, using the progressive form the speaker indicates that the event "my writing a paper" belong to "now–here" space. And in the second sentence, "would" indicates that the event belongs to a space in irreality domain.

Our argument is that the dynamic mode is close to the mental space construction in (1) and the metaphoric mode to that in (2) in Fig. 8. That is, in the dynamic mode, players handle patterns without associating concrete meanings to them. Rather, players pay more attention to the texture of patterns themselves. The interpretation is literal, which means what is interpreted is what is shared between the subjects. What is conveyed through the dynamic mode is the speaker's attention to the "now" and "here" of the interaction. On the other hand, in the metaphoric mode, people associate meaning to the patterns in the message. This meaning is not shared by the other player. For example, people associate the symbol "*" with a rabbit (see M1-2 in Sec. 4). What is shared with the partner is the symbol "*" which belongs to "now" and "here". And the meaning "rabbit" does not exist in reality, but in the irreality created by the player. What is different from the function of "would" is that, in the case of "would", the irreality domain is shared by the listener. Still as we have seen with the examples in Table 4, sometimes it can be hinted that the other player is associating meaning with the pattern (i.e. constructing an irreality space in communication).

The interesting observation here is that the "grammatical process" observed includes the exploration of the media. Players try to find the possibility of the pattern, and what kind of patterns can be used to make distinctive messages. Based on our findings, we propose that the exploration of the nature of the media is an

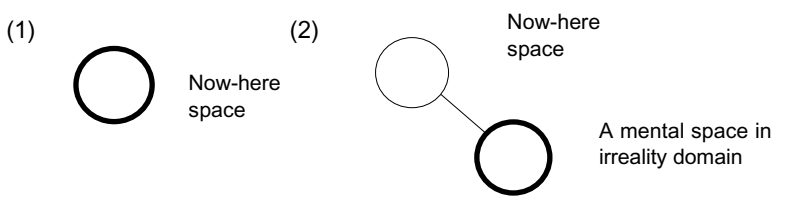


Fig. 8. A simplified notation for mental spaces in sentences (1) and (2). The circle expresses a mental space. The type of the space is written next to it. The progressive form and the auxiliary "would" show that the event "my writing a paper" is located in a space with a thick line. A line between the two spaces in (2) shows that the lower space is opened based on the upper space.

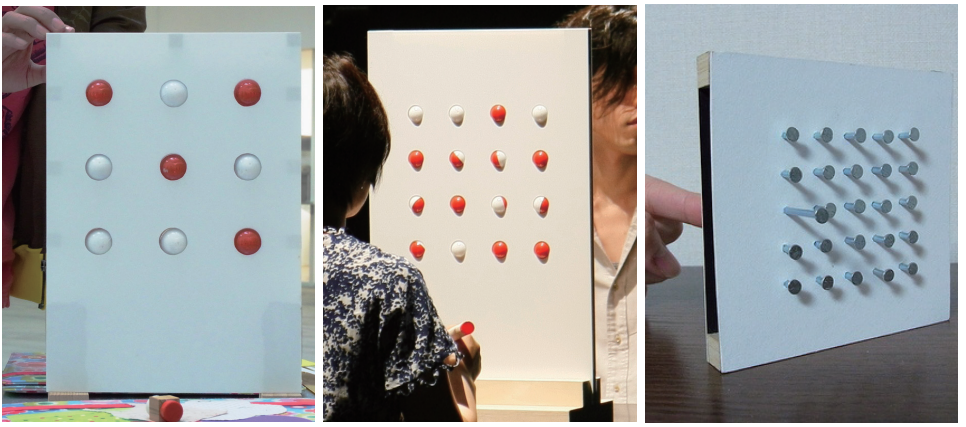


Fig. 9. We are now making various types of wall games to analyze the exploratory behavior of players in playing with the wall. (Designed by Seara Ishiyama.)

integral part of the emergence of grammar.[d] Just by looking at natural language, whose media is already transparent to the users, it is difficult to see whether or not this is true. As shown in Fig. 9, we are currently making wall games with various textures so that we can observe how the players explore the media until the game became consciously transparent, until the game board became a part of the body schema [17]. This will give us a way to look into all kinds of theoretically possible languages.

Acknowledgment

We would like to thank Tom Froese, Akira Honda and Masashi Okamoto for insightful comments; Seara Ishiyama and Miyuki Kawamura for technical support.

References

[1] Bates, E., *Language and Context: The Acquisition of Pragmatics* (Academic Press, New York, 1976).
[2] Cangelosi, A. and Harnad, S., The adaptive advantage of symbolic theft over sensorimotor toil: Grounding language in perceptual categories, *Evol. Commun.* **4**(1) (2000) 117–142.
[3] Cutrer, L. M., *Time and Tense in Narrative and in Everyday Language* (Ph.D. dissertation, University of California, San Diego, 1994).

[d]There might be a Japanese writing style influence on the outcome of the wall game, but as for the present paper we do not have enough space to go into that direction. One speculation that we made is that Eastern people (such as the Japanese) might have a more dynamic interpretation than Western people based on Kiyokawa, Dienes, Tanaka and Yamada's (2010) study [11] which points out that with respect to pattern recognition, Eastern people prefer holistic processing (paying attention to the global shape) and Western people analytical processing (paying attention to the components).

[4] Fauconnier, G., *Mental Spaces: Aspects of Meaning Construction in Natural Language* (MIT Press, Cambridge, 1985, 1994).
[5] Galantucci, B., An experimental study of the emergence of human communication systems, *Cognitive Sci.* **29** (2005) 737–767.
[6] Gergely, G. and Watson, J. S., Early socio-emotional development: Contingency perception and the social biofeedback model, in *Early Social Cognition*, Rochat, P. (ed.) (Lawrence Erlbaum Associates, Hillsdale, NJ, 1999), pp. 101–137.
[7] Gómez, J. C., Sarria, E. and Tamarit, J., The comparative study of early communication and theories of mind: Ontogeny, phylogeny and pathology, in *Understanding Other Minds: Perspectives from Autism*, Baron-Cohen, S., Tager-Flusberg, H. and Cohen, D. J. (eds.) (Oxford University Press, Oxford, 1993), pp. 195–207.
[8] Hashimoto, T. and Ikegami, T., Emergence of net-grammar in communicating agents, *BioSystems* **38** (1996) 1–14.
[9] Kirby, S., Natural language from artificial life, *Artif. Life* **8**(2) (2002) 185–215.
[10] Kirby, S., Cornish, H. and Smith, K., Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language, *Proc. Natl. Acad. Sci. USA* **105** (2008) 10681–10686.
[11] Kiyokawa, S., Dienes, Z., Tanaka, D. and Yamada, A., Cross cultural differences in implicit learning, *CogSci 2010 Proc.* (2010) 2206–2211.
[12] Lakoff, G. and Johnson, M., *Metaphors We Live By* (University of Chicago Press, Chicago, 1980).
[13] Langacker, R. W., *Concept, Image, and Symbol* (Mouton de Gruyter, Berlin, 1991).
[14] Marocco, D. and Nolfi, S., Emergence of communication in embodied agents evolved for the ability to solve a collective navigation problem, *Connect. Sci.* **19** (2007) 53–74.
[15] Reichenbach, H., *Elements of Symbolic Logic* (Macmillan, New York, 1947).
[16] Rizzolatti, G. and Arbib, M. A., Language within our grasp, *Trends in Neurosciences* **21**(5) (1998) 188–194.
[17] Sato, Y., Iizuka, H. and Ikegami, T., Generation of body images in an active/passive perception model, *Proc. 5th Asia-Pacific Computing and Philosophy Conf.* (Tokyo, Japan, 1–2 October 2009), pp. 100–105.
[18] Scott-Phillips, T. C. and Kirby, S., Language evolution in the laboratory, *Trends Cognitive Sci.* **14**(9) (2010) 411–417.
[19] Steels, L., The emergence and evolution of linguistic structure: From lexical to grammatical communication systems, *Connect. Sci.* **17**(3) (2005) 213–230.
[20] Steels, L., Experiments on the emergence of human communication, *Trends Cogn. Sci.* **10**(8) (2006) 347–349.
[21] Steels, L. and Kaplan, F., AIBO's first words: The social learning of language and meaning, *Evol. Commun.* **4**(1) (2001) 3–32.
[22] Sugita, Y. and Tani, J., Learning semantic combinatoriality from the interaction between linguistic and behavioral processes, *Adapt. Behav.* **13**(1) (2005) 133–152.
[23] Tager-Flusberg, H., Language and understanding minds: Connections in autism, *Understanding Other Minds: Perspectives from Autism*, Baron-Cohen, S., Tager-Flusberg, H. and Cohen, D. J. (eds.) (Oxford University Press, Oxford, 2000), pp. 124–149.
[24] Thibodeau, P. H. and Boroditsky, L., Metaphors we think with: The role of metaphor in reasoning, *PLoS ONE* **6**(2) (2011) e16782, doi:10.1371/journal.pone.0016782.
[25] Uno, R., Marocco, D., Nofi, S. and Ikegami, T., Emergence of protosentences in artificial communicating system, *IEEE Trans. Auton. Ment. Dev.* **3**(2) (2011) 146–153.
[26] Vogt, P., Perceptual grounding in robot, in *Proc. 6th European Workshop on Learning Robot*, Birk, A. and Demiris, J. (eds.) (Springer-Verlag, 1998), pp. 126–141.