Transition from imperatives to declaratives in artificial communicating systems

Ryoko Uno(a), Davide Marocco(b), Stefano Nolfi(b) and Takashi Ikegami(a)

a) **The Graduate School of Arts and Sciences, University of Tokyo, 3-8-1 Komaba, Tokyo 153-8902, Tokyo , Japan.**
b) Institute of Cognitive Science and Technologies, CNR, Via San Martino della Battaglia 44, Rome, 00185, Italy

Abstract

This paper investigates the relationship between embodied interaction and symbolic communication. We refer to works by Iizuka & Ikegami and Marroco & Nolfi as the examples of simulating EC (embodied communicating) agents, and argue their differences in terms of joint attention, a class of communication between cognitive agents. Then we introduce a new simulation to bridge the gap between the two models; with the new model we demonstrate the two pathways to establishing agents' coordinating behaviors. Based on the simulation results, we explain the typology of sentences (such as 'declarative', 'imperative' and 'exclamative' sentences) from a communicative point of view, which challenges the traditional views of formalizing grammar.

1. Introduction

What are the precursory variations in preverbal communication before language systems emerge? Chomsky could argue that no such precursory system does exist; however, recent developments in cross-disciplinary studies have revealed that there are ample examples showing the embodied groundings for language communication. Kita studied Japanese mimetics (1997, 2002) and argues that the meaning of Japanese mimetics is primarily represented in an *affect-imagistic* dimension, where "language has direct contact with sensory motor and affective information" (Kita 1997) and "vivid imagery of perceptual and physiological experiences" (Kita 2001). Glenberg (2002) shows how action execution interferes with understanding meanings of sentences. He proposed an effect of action-sentence compatibility to demonstrate that even sentences which do not express concrete movement describe the transfer of abstract objects and concepts. For example, "She told you the story" will interfere with your body schema. The relationship between language and embodiment has been intensively sought after since the discovery of mirror neurons (Rizzolati, 1996). Now the so-called mirror neural system (MNS) is defined as a neural subsystem responsible for matching the self-action with the other's same intentional action. Rizzolati and Arbib (1998) have argued that a ventral premotor cortex called F5 and AIP found in Makaque monkeys is theoretically relevant for evolving a language system. They aren't only reporting the functional similarities between F5 and the human Broca area, but they also hypothesize the unique scenario of initiating a communication system with MNS, which they call an observation/ execution matching system.

Artificial life study provides a test bed for examining how symbols and grammars emerge in minimal interacting systems. For the last 10-15 years, artificial life studies have contributed greatly to this direction, and the origin and evolution of language has become a target of scientific study (see e.g. Steels[2005], Kirby[2002], Cangelosi & Harnad [2000], Vogt [1998], Hashimoto & Ikegami [1996], Sugita & Tani [2005], Sasahara & Ikegami [2007]. etc.). For example, Steels & Kaplan (2000) have developed a rigid platform for studying interacting speaking/listening agents. Steels' basic approach is to set a plausible gaming environment to examine what sorts of grammatical rules to evolve in a community. In the case of the "naming" game, agents in a community are trying to associate consistent labels of objects in the environment, or in the case of the guessing game, agents guess the objects by communicating with other agents. In his recent studies, more sophisticated grammatical structures that can incorporate perspectives have been analyzed with embodied agents (AIBO, for example). Steels' approach assumes we have primitive symbols and communication channels, and then he examines how symbols and rules would evolve. On the other hand Marocco and Nolfi's approach (2006) investigates the less symbolic model to see how symbolic/behavioral communication emerges. They also assume a simple gaming environment where mobile agents on a two-dimensional arena get rewarded, if two of them can simultaneously get to the same circle in the arena. Those agents can communicate by passive local infra-red sensors and active local/global acoustic signals. They have reported a qualitative transition of acoustic signals in an evolutionary pathway.

Fully embodied interacting agents with no symbolic communication have been studied more intensively (see e.g. Baldassarre et al., 2003). When seeking to bridge the gap between nonverbal and verbal communicating system, IIzuka and Ikegami's simulation study (2004, 2007) deserves attention. They have studied how turn-taking between two agents can develop without using any explicit cue but with some subtle action variations for exchanging turns. It reminds us of Rizzolati and Arbib's observation/matching theory.

In this paper, we will compare Iizuka and Ikegami's turn-taking study with Marroco and Nolfi's signaling game from linguistic points of view. Since the two research studies have many things in common but stress different viewpoints, it is useful to compare their studies to clarify the problems lying between verbal and non-verbal communicating systems. Based on that, we will examine a variation of Marroco and Nolfi's model to see the effect of symbolic communication with embodiments. A typological classification from linguistics and the concept of joint attention will help us to sort out the problem, which we will briefly describe in the next section. In particular, we propose a classification of joint attention: instrumental and participatory JA. In section 3, we compare two communication models to clarify the important issue. By developing concepts and ideas, we newly interpret and generate variational studies on Marocco and Nolfi's modeling in section 4. In the discussion part, we reconsider the organization of sentence typology from interaction perspectives based on the new simulation results.

2 Joint attention and Inter-subjectivity

2.1 Joint attention

Joint attention (henceforth "JA") is a coordinated preverbal behavior among two or more persons. A simple example is a children's pointing behavior under the attention of the mother. It is a process of sharing one's experience of observing an object or events with others by following pointing gestures or eye gazing. We distinguish two types of joint attention. If a person uses joint attention as a tool to achieve a goal (e.g. establishing the joint attention to let your dog pick up a ball), we call it "instrumental joint attention". But if a person takes joint attention itself as a goal, we call it "participatory joint attention". For example, two people looking at the same sunset establish the participatory joint attention as it doesn't require further achievements. We assume that participatory and instrumental joint attentions form continuous spectra, as shown in Figure 1.

We extend the meaning of JA and call any performance such as language or dance that would make two people pay attention to the same thing (except for each other) JA in a broad sense.

Participatory joint attention can be regarded as an example of inter-subjectivity, which is to share mental states and intentions with others. Murray & Trevarthen (1993), who first noticed its importance, classified two types of inter-subjectivity. A primary inter-subjectivity is found in between an infant and a caretaker. It is thought that there exists an infant's innate ability to organize inter-subjectivity between she and the mother. From around 7 months old, a secondary inter-subjectivity develops. It starts to involve a third object or event besides the person-person relationship. So, more specifically, instrumental JA is related with the secondary inter-subjectivity.

In order to clarify the interrelation between JA, inter-subjectivity, and turn-takings which we are going to see in the following section, we have to refer to another notion: novelty.

JA, not always but usually, has something to do with surprise and expectation. A kind of "novelty" attracts another person to engage in communication. It is known (and used) in developmental studies that infants watch new/surprising events longer. Namely, infants are attracted to new events and thus interaction between infants and the given event will be established. Nadel (1999, 2002) gave some anecdotal reports on synchronous imitation between infants at the age of two and a half years. She reports that synchronous imitation or perception of imitation/being imitated is an important cognitive skill that promotes an ability to find topics and turn-taking in a preverbal interaction. Further, by using synchronous imitation or recognizing mutual action patterns, infants can be ready for correctly inferring others' intentions. Nadel argues that un-affordant ways of using objects can initiate synchronous imitation.

We use a novelty as unpredictability as a necessary factor to start and maintain interaction. Novelty or unpredictability should be continuously generated in order to sustain interactions, but at the same time it is also true that novelty cannot be prepared beforehand and too much novelty breaks up the interaction. As such, JA infers the continuous and adequate introduction of novelty into the interacting field.

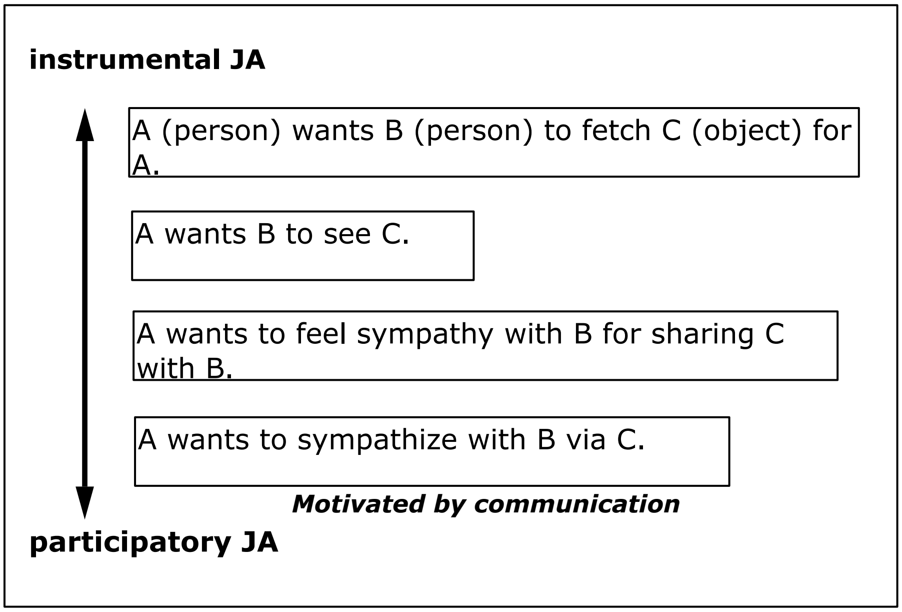

Figure 1 A degree of using JA as a purpose of using JA is interpreted and used it as a measure for ordering from a participatory to instrumental JA. A participatory JA is typically taking JA as it purpose.

3. Two Simulation Studies

First we introduce two previous simulation studies which can be seen as attempts to evolve JA mechanisms. The first one models turn-taking and the second models communication with signals. Then, we argue the differences between them.

3.1 Simulated Turn-taking (Ikegami & Iizuka, 2004)

A task of the simulation is the maintenance of turns and spontaneous switching of turns between two mobile agents on the two-dimensional field (Figure 2). When one agent gets behind the other, by definition the other agent has its turn. Here ‘getting behind’ means entering the region in the other's rear scope. An agent gets the apart distance and the heading of the other agent from the input sensors then it computes the motor outputs by the internal neural circuit. Here a small noise was added to the sensory input so that the agents have limited sensory information.

The agent’s internal neural circuits were evolved by a genetic algorithm. There are two populations (each contains 15 individuals). The performance of all paired agents from the separated populations is evaluated at each generation to choose the best performing pairs. The performance is measured by 1) How good is the prediction of the other agent’s future navigation, and 2) How much time can an agent take for its own turn? By multiplying the fitness of two agents, we define it as a performance of the selected pair. According to this fitness function, agents are selected and breed with some mutations in the neural connections. The result of the simulation is briefly summarized as follows.

1) There are two phases along the evolution timeline. In the early stage of the evolution, the turn-taking is geometrical and regular in space and time. Agents are "automatically" taking turns similar to two pendulums adjusting their phases. But in the later stage of the evolution the pattern becomes more chaotic and dynamic. Agents are changing their positions temporally and its timing can vary from time to time. A remarkable property of those chaotic dynamics is that the agents can cope with agents from different generations.
2) When switching turns, agents' prediction accuracy is significantly lowered and has not been improved by a genetic algorithm; nevertheless the performance of the turn-taking is getting better. Therefore, agents can switch without prediction.

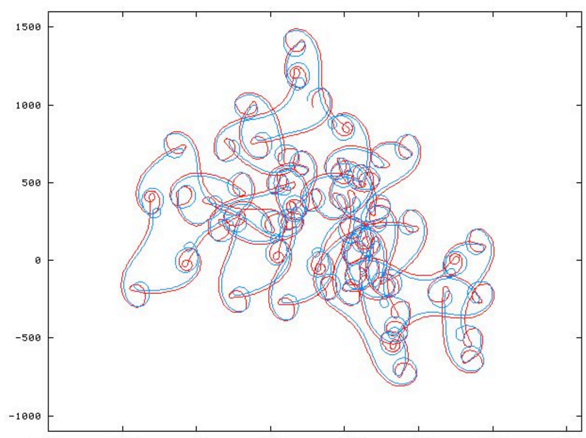

Figure 2 Two agents' trajectories were overlaid, where the turn-taking dynamics becomes chaotic in space and time.

3.2. Simulated Signaling Cooperation (Marocco & Nolfi, 2006)

A task of the simulation work is to solve a collective navigation task by evolving four mobile robots. To achieve the task, agents go to a target area, making a pair and staying there (Figure 3). Therefore, four robots are spontaneously subdivided into two pairs of robots. The important point is that robots can emit signals in this simulation.

A robot (of a circular body) has 14 sensory neurons. By the eight infrared sensors, the robot can detect obstacles and other robots nearby. By the ground sensor, the robot can detect the color of the ground. The floor is colored white except for the target area which is colored black. By the four communication sensors, the robot can perceive the signals emitted by the other robot. By the self-monitoring sensor, the robot can recognize the signal emitted by itself a step before. Here the communication by the signal is spatially limited; a robot can hear the signal emitted by the nearby robot. The agent's internal neural circuits were developed by a genetic algorithm.

The result of the simulation of the best agents is briefly summarized as follows:

In the early time phase, the robots explore the environment. In this phase their signals reflect the environment and the signal of the other robots. However, they do not exploit the signals of each other. In the second phase, one of the agents gets into a target area colored black. Since the ground sensor recognizes the black color of the floor and the voice of the robots inside the target area becomes distinctively intensive. During this period the agent outside of the target area lowers its voice. Here, the speaker-hearer distinction appears. Using the signal of the agent inside the target area the other one can come into the target area

In the last stage of the trial, both agents come inside the target area and their signals start to synchronize. Because of this synchronization they can remain inside the target area, and the collective navigation task is completed.

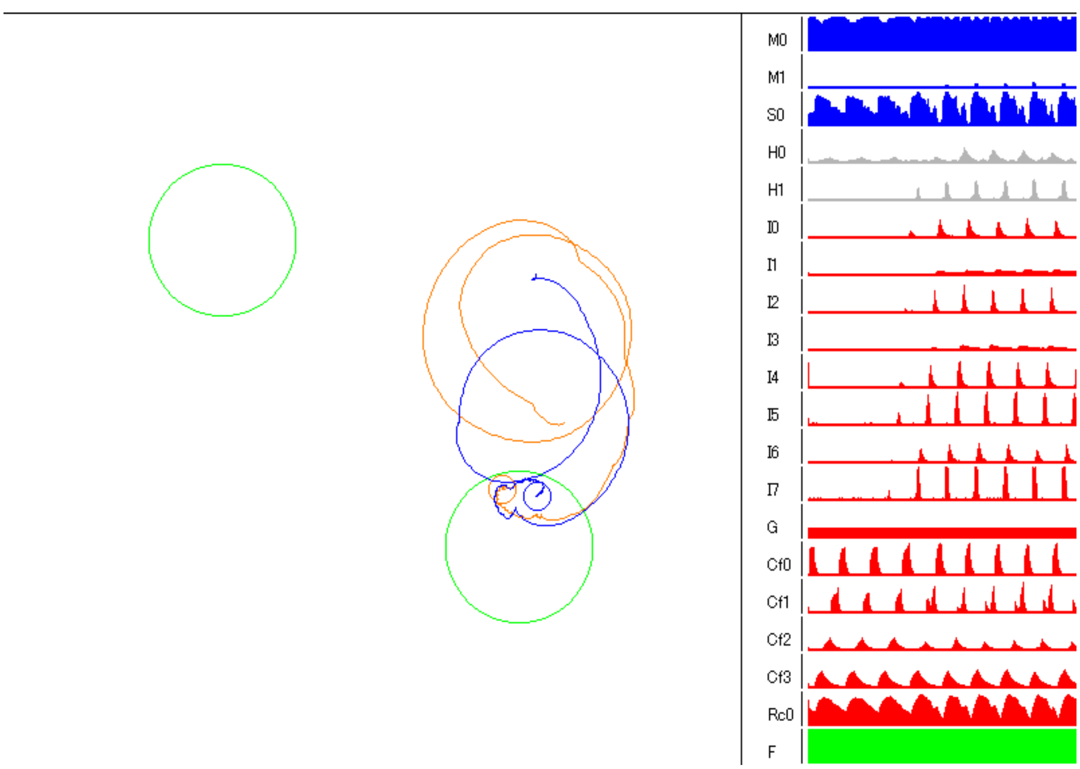

Figure 3 Two agents' trajectories on the field with two target areas (left) and the associated sensory time series(right).

3.3 Comparison of the two simulations

Two models are focusing on the different kinds of communication and both of them use synchronization in different ways. In Iizuka & Ikegami's model, turn-taking is established by alternatively switching positions. As the result, their spatio-temporal patterns of sequential turn-taking look like the same patterng (see Figure 2). But their internal states or headings aren't synchronized in the rigid sense. In the case of Marocco & Nolfi's model, a synchronized oscillation is observed in several sensory channels. When two agents come to stay in a target area, they are communicating with each other by having synchronized signaling patterns.

On the other hand, Iizuka & Ikegami's model shows more dynamic interaction with each other. That is, turn-taking is maintained by creating chaotic but correlated dynamics, which is generated by the interaction between the agents. In Marroco & Nolfi's model, agents communicate in rather static ways. Their signaling pattern is already separated from their embodiment and becomes a function of the task to solve.

Lastly, agents in Iizuka & Ikegami's model always aim for the participatory JA, whereas Marroco & Nolfi's model demonstrates the transition from the instrumental to participatory JA. This transition is interesting and useful for classifying the possible interaction between the internal states of the agents. The other critical issue here is that Iizuka & Ikegami put the stress on the difference between mere synchronization and the inter-subjectivity achieved by participatory JA. Shared intentionality requires detecting the intentionality of the other agents, which is translated to the sensitivity of the styles of turn-taking in their models. Without having the variety of styles of motion, it is difficult to examine the JA in fruitful ways. Thus, they used co-creativity and cooperation instead of synchrony.

As we see above, there are advantages and disadvantages in those two models. To consider the evolution of communication styles, it is important to pose a new model that is a sort of compromise between them. Agents in Iizuka & Ikegami's model don't care where to take turns, whereas agents in Marroco & Nolfi have the intrinsic motivations to go to a target area as the expected task. When the agents come to the target area, they have the ground sensors in an active state. Now we train the agents with unreliable ground sensors. Therefore, agents should coordinate their behaviors with the infrared sensors and communication channels. Since agents in Iizuka & Ikegami's model are merely using sophisticated IR sensors to establish turn-taking, this new setup will link the two models. Also by having this new setup, the linguistic comparison will be made much clearer, as will be shown in the discussion section.

4. Variations of the Signaling model

4.1 New setup

In modifying Marocco and Nolfi (2007), three changes are made. First, in the training phase the target area is not only in black but it can be any shade of gray. Second we only have two agents in the field. Thirdly, now robots can hear each other's signals all over the environment (in their first model, the signaling communication was limited to the local neighborhoods). These changes resultsin the different "strengths" of symbols mediating the collective behavior. Because the signals for the target domain become uncertain, the agents try to use other sensory channels to enter the target areas. Secondly, the infrared sensors change their meanings. Since there are only two agents in this new environment, when the infrared sensors are activated within a target area, it means either that the other agent or a wall exists in the proximity. In the previous setup, it was not able to use the infrared sensor to detect the agent who is in the target area because there were many robots. Since the signal for the target area is made uncertain by temporally being varied in target color and since there are only two agents, detecting agents by the infrared sensor now becomes possible. Finally, to discriminate the wall and the other robots, synchronization becomes useful, which is also true for the previous experiments. It has to be emphasized that the signal for the infrared sensor sometimes overrides the signal for the target area. In this case, synchronization is used only to find out whether the robot is interacting with another robot or not. Signals are used rather to build up a ground of interaction between the two.

4.2 Two forms of collective behaviors

There two collective navigation forms observed in this new setup. In the first type, a synchronization of all channels is organized inside the target area (We call this "JA inside a target area"). In the second type, a synchronization of infra-red (IR) and communication (C) channels is organized but not the ground channel (G) outside the target area (We name this "JA outside the target area"). The following diagrams in Figure 4 show how the interrelation between robots changes over time.

In the case of “JA inside a target area”, both robots explore inside the field in the beginning stage. Then one of them gets into the target area by chance and emits intensive signals, which the other can hear. Here their signals have no correlation at all. But when the one outside the target area can get inside (sometimes guided by the first agent) they then get close to each other. When both agents’ IR sensors are turned on, the C sensors show a strong synchronization. But this collective state isn’t that stable compared to the previous model, and the collective state breaks up by losing the IR sensory patterns due to the fact the agents move apart within the target area. When this happens, the synchronization of the C sensory pattern is weakened (i.e. only showing anti-phase synchrony). Eventually, one of the agents leaves the target area.

On the other hand in the case of “JA without a target area”, the trial starts from exploration and then they find each other using the IR sensors. Once both IR sensors are turned on, then their C signals get into synchronization.

Namely, the trigger of the synchronization is given by the communication channel, but only after the infrared sensor's synchronization can the communicative channel synchronize. The potential cue for establishing JA is in the order of ground sensor < infrared sensor < communicative channel.

We interpret the results in the following way using the states of three channels. At the exploration phase, all the three sensors are off, which we express as a state (-, -,-). The three positions stand for the ground sensor, the infrared sensor and the communication channel, respectively. Then in the case of JA inside the target area, sensors end up with the value of (+, +, +). In the case of JA outside the target area, it starts from (-,-,-) and ends up with (-, +, +). Based on the typology of joint attention and inter-subjectivity section 2, we can say that, primary inter-subjectivity is present when we have the value (-, -, -) and secondary inter-subjectivity or participatory JA as (*, +, +). And, instrumental JA is present when one has (-, -, +) while the other has (+, -, -).

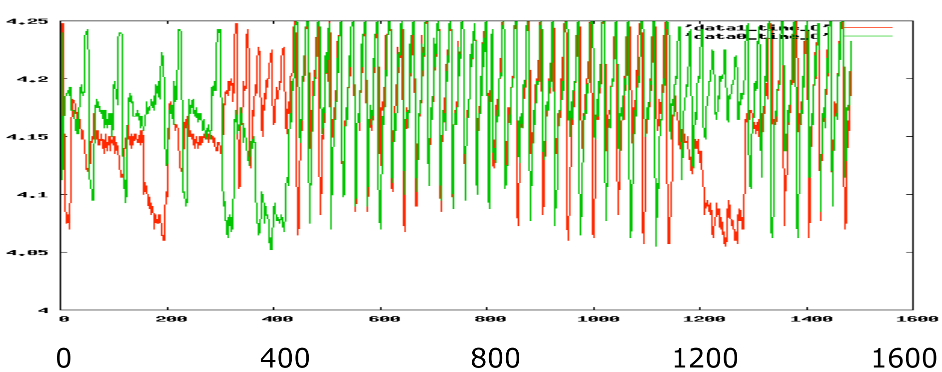

Figure 4 Communication sensory data of two agents as a function of time in case of JA inside a target area. They demonstrate synchronization around 400-800 and 1300-1500 time steps. Anti-phase oscillation is observed around 800-1100 time steps.

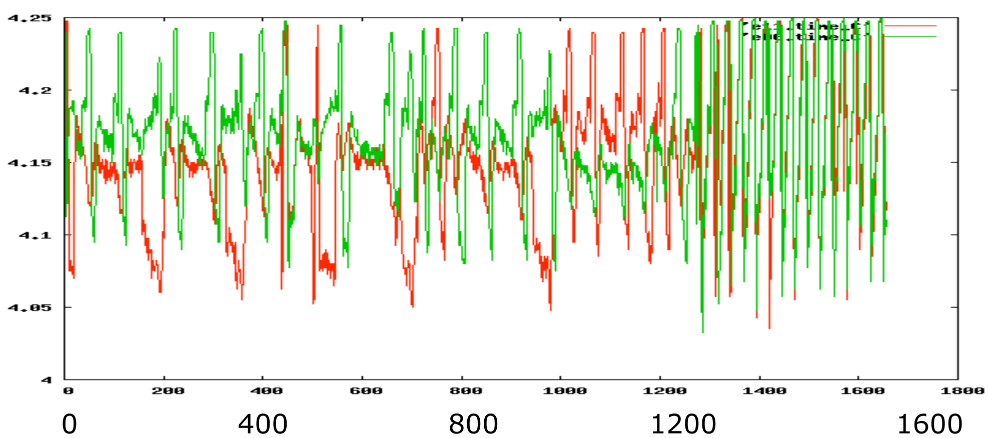

Figure 5 Communication sensory data of two agents as a function of time in case of JA outside a target area. They demonstrate synchronization around after 1300 time steps .

The following points should be noted. Close analysis of synchronization reveals that phase synchronization comes first and amplitude synchronization comes afterwards in the case of JA within/without a target area. Since the infrared sensors are passive, those patterns are easily lost while agents are moving around. But the communication channels can be actively synchronized by modifying the communication channels.

An implication from this itinerant behavior observed in this new setup is that the infrared and communication channel synchrony is more unstable than in the case of reliable ground sensing. Because agents can't trust the ground sensor pattern, they organize imaginary synchrony without a target area, which we consider an essential function of language, as will be discussed in the discussion section.

Section 5. Discussion: Sentence typology

Up to this point we have observed three types of interaction from the perspective of JA. First we compared the models of turn-taking and communication with signaling. While turn-taking only realizes one type of JA, namely participatory JA, communication of signaling realizes various types of JA that ranges from instructive JA to those which have some characteristics of participatory JA. In this paper we have modified the communication with a signaling model to see whether we can observe the emergence of a pure participatory JA continuously from instrumental JA. We have focused on the nature of the coupling by analyzing the behavior of the sensors.

We propose that the results of modeling interaction can be linked to linguistic structure via the typology of JA. Traditionally, what a speaker is going to do with linguistic expression and the linguistic structure does not have to be mutually correlated. The speech act is conventionally linked with linguistic structure called sentence-type.

Sentence typology is originally based on the intention of the speaker. Based on typological research, Sadock and Zwicky (1985) argue that the major forms of sentences are declaratives, imperatives and interrogatives. Declarative sentences are forms which are combined with the speech act of making statements. Imperative sentences are combined with commanding. And interrogative sentences are combined with questioning. Exclamative sentences also are related to surprises.

Reinterpretation of sentence typology from the interaction pattern rather than the speaker's intentional stance provides us a novel view on the classification of sentences. Among sentences, declaratives are thought to be the most typical ones. Most of linguistic analysis is based on declarative sentences, and usually declaratives are thought to be used for information transmission. However, as we discussed, "a full declarative" is defined as a special way to share the ground of speech between the speaker and the hearer. If we take a declarative sentence such as "A cat chased the rat" it is informative and it elicits a particular type of ground sharing. It has to be noted that declarative sentences are usually thought redundant or less optimized from the perspective of information. For instance, when we say "The moon is beautiful" to a person who is looking at the moon with us, there is no information transmission but there is a sharing of ground. A complementary category of sentences is "full imperative". Sentences in this category use the sharing of ground of speech as a tool to achieve some purpose. The speaker's use of a sentence affects the behavior of the hearer (e.g. "come here!

This idea of declarative sentences points out that language may help establishing JA in a broad sense. The simplest case is a one-word sentence. There are mainly two cases of this. The first case is to make a request (e.g."Water!"). The second case is to share the context. (E.g. By saying "Snow!", you can direct the attention of others to the snow and make them feel the same way you do.)

Using the term JA we can restate the argument in the previous section as follows: Full declaratives are sentences that are used for participatory JA and full imperatives are used for instrumental JA. Exclamatives are the special case used as a primary inter-subjectivity, which is often established between a baby and his or her mother. It is not necessarily true that the intention of the speaker exists before a communication starts. Intention is a co-product of a communication itself, which was also true in the case of simulated turn-takings. The problem is that if we only focus on the speaker's intention we do not know why only these speech acts are grammaticalized and not the others. Instead we aimed to show what two people sharing the ground of speech has to do with the grammaticalization of speech acts. By applying the notion "joint attention" we might be able to understand why we have these types of sentences.

Types of sentences	**Speech Act**	**Joint attention**	
Exclamative	Exclamation	Readiness for JA	N/A
Imperative	Command	Tool JA	Asymmetry
Interrogative	Question	Tool JA	Asymmetry
Declarative	Statement	Goal JA	Symmetry

Table I. A mapping between types of sentences and that of Joint attention.

Acknowledgments

This work is partially supported by The 21st Century COE (Center of Excellence) program (Research Center for Integrated Science) of the Ministry of Education, Culture, Sports, Science, and Technology, Japan, and the EC Agent project, sponsored by the Future and Emerging Technologies program of the European Community (IST-1940).

References

Baldassarre G., Nolfi S., Parisi D. (2003). Evolving mobile robots able to display collective behavior. *Artificial Life, 9,* 255-267.

Cangelosi A. & Harnad S. (2000). The adaptive advantage of symbolic theft over sensorimotor toil: Grounding language in perceptual categories. *Evolution of Communication* 4(1), 117-142

Giacomo Rizzolatti et al. (1996). *Premotor cortex and the recognition of motor actions*, Cognitive Brain Research 3 131-141

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. Psychonomic Bulletin and Review, 9, 558-565.

Hashimoto, T. and Ikegami, T. (1996) Emergence of Net-grammar in Communicationg Agents BioSystems 38 1–14. 1996

Iizuka, H. and T.Ikegami (2004) Adaptability and Diversity in Simulated Turn-taking Behavior Artificial Life 10. 361-378.

Ikegami, T and Zlatev, J.(2007) From pre-representational cognition to language, In Body, Language and Mind, vol. 2. Zlatev, Jordan; Ziemke, Tom; Frank, Roz; Dirven, Ren (eds.). Berlin: Mouton de Gruyter. 241-283.

Ikegami, T. and Iizuka, H. (2007) Turn-taking Interaction as a Cooperative and Cocreative Process, Infant Behavior and Development 30.278-288.

Kirby, S. (2002b). Natural language from artificial life. *Artificial Life*, 8(2):185-215.

Kita, Sotaro (1997) Two-dimensional semantic analysis of Japanese mimetics. *Linguistics* 35: 379–415.

Kita, Sotaro (2001) Semantic schism and interpretive integration in Japanese sentences with a mimetic: a reply to Tsujimura. *Linguistics* 39: 419–436.

Marocco, D. and Nolfi, S. (2006) Self-Organization of Communication in Evolving Robots. In Luis M. Rocha et al., editor, *Artificial Life X*. MIT Press. 178-184.

Melzoff & W. Prinz (Eds.), The imitative mind: Development, evolution, and brain-bases (pp.42–62). Cambridge: Cambridge University Press. 244

Murray, L., &Trevarthen, C. (1993). Emotional regulations of interactions between two-month olds and their mothers. In T. M. Field, &F. N. A.(Eds.), Social perception in infants(p.177-197).Norwood,NJ:Ablex.242

Nadel,J.,Carchon,I.,Kervella,C.,Marcelli,D.,&ReserbatPlantey,D.(1999). Expectancies for social contingency in 2-month-olds. Developmental 245 Science,2(2), 164–173.246

Nadel, J. (2002) Imitation and imitation recognition: Functional use in preverbal infants and nonverbal children with autism. InA.

Rizzolatti G, Arbib MA (1998) Language within our grasp. Trends Neurosci 21: 188-194

Sadock J. & A.M.Zwicky (1985) Speech act distinctions in syntax. Timothy Shopen (ed.) Language Typology and Syntactic Description, Vol. 1. Cambridge: Cambridge University Press. 155-196.

Sasahara, K., Ikegami, T. (2005) Evolution of Song Comu-nication in a 2D Space. Advances in Artificial Life: 8th European Conference ECAL2005 Proceedings. 835-844. Springer-Verlag.

Sasahara, K and Ikegami, T., Evolution of Birdsong Syntax by Interjection Communication, Artificial Life 13 (2007) pp. 1-19.

Steels, L. and Kaplan, F. (2000) AIBO's first words: The social learning of language and meaning. Evolution of Communication, 4(1):3-32.

Steels, L., (2005). The Emergence and Evolution of Linguistic Structure: From Lexical to Grammatical Communication Systems. *Connection Science, 17(3-4)*, September-December. 213-230.

Steels, L. (2006) Experiments on the emergence of human communication. *Trends in Cognitive Sciences* 10(8), 347-349.

Steels, L. (2006) Semiotic Dynamics for Embodied Agents. *IEEE Intelligent Systems* 21(3), 32-38.

Sugita Y. and J. Tani (2005) Learning semantic combinatoriality from the interaction between linguistic and behavioral processes", Adaptive Behavior, Vol.13, No.133-52.

Trevathen, C.(1999). Musicality and the intrinsic motivepulse: Evidence fromhuman psychology and infant communication. MusicaeScientiae, SpecialIssue,155–215.259

Uno,R.,&Ikegami,T.(2003). Joint attention / prediction and language: A mechanism to align intentionalities. Papers in Cognitive Linguistics, 2, 231–274.Hituzi.

Vogt, P. (1998) Perceptual grounding in robots. In (Birk, A. and J. Demiris eds.) Proceedings of the 6th European Workshop on Learning Robots 1997. *Lecture Notes on Artificial Intelligence.* Springer-Verlag 1998.