Mirror neuron framework yields representations for robot interaction

Emilia I. Barakova [a,*], Tino Lourens [b]

[a] Eindhoven University of Technology, P.O. Box 513, 5600MB Eindhoven, The Netherlands
[b] Philips Healthcare, P.O. Box 10000, 5680DA Best, The Netherlands

ARTICLE INFO

ABSTRACT

Common coding is a functional principle that underlies the mirror neuron paradigm. It insures actual parity between perception and action, since the perceived and performed actions are equivalently and simultaneously represented within the mirror neuron system. Based on the parity of this representation we show how the mirror neuron system may facilitate the interaction between two robots. Synchronization between neuron groups in different structures of the mirror neuron system are on the basis of the interaction behavior. The robotic simulation is used to illustrate several interactions. The resulting synchronization and turn taking behaviors show the potential of the mirror neuron paradigm for designing of socially meaningful behaviors.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Recent neurophysiological, cognitive, and developmental research clearly shows that there are shared representations in the brain between perceived and generated actions, between actions produced by oneself and others (see for instance [14,15,19,25]). These shared representations, conveyed by the mirror neuron system (MNS), underlie the process of imitation, social learning, and prediction of the behavior of conspecifics. Many attempts have been made to model the imitation process, for review see [16,24]. However, the imitation that has been modeled so far does not go further than one directional demonstrator–imitator interaction. In this paper, we want to make an attempt to show the potential of the mirror neuron paradigm for social interaction, in particular for movement synchronization, entrainment, and interchangeable turn taking between two agents.

Entrainment of timing of social interaction has been investigated in multidisciplinary research on conversation. Conversation is an exchange of speech between two or more individuals. Although at first glance it looks like a chaotic process, conversation usually proceeds smoothly, by having the two parties take well timed turns. A number of authors have proposed that the listeners anticipate an upcoming end of a turn by perceiving eye gaze, body movement, or other semantic, syntactic, or prosodic queues from the speaker, for reviews see [8,9]. Conversely, listeners indicate their desire for turn ending. Speech is, in its essence, a motor act and it is likely that the mechanisms of speech and turn taking co-evolved, perhaps building on the same preexisting structures and mechanisms for motor expression [27].

In their theoretical study Wilson and Wilson [27] argue that turn taking is likely to be successfully modeled by entrainment of endogenous oscillators.

Mutual entrainment of rhythmic activities has been theoretically studied as the basic mechanism of the organization of temporal order by Pavlidis [17]. Endogenous oscillators have been implicated in a range of cognitive processes, including perception, motor control, attention, memory, and consciousness [5].

In a robotic setup turn taking behavior is discussed in [6,11]. The turn taking behavior in these studies takes place as a result of interaction of two dynamical recognizers—Elman type of recurrent neural networks that have widely been used to model dynamic systems. The training has been replaced by a genetic algorithm, which aims to produce 'genetically different' agents. This will prevent from the low reliability of the interaction process based on neural learning [5].

We base our interaction behavior on synchronization between neural lattices that together simulate the mirror neuron-like function. The neuron firing in every lattice of neurons is modeled by an oscillatory model.

This paper is organized as follows. In Section 2, we propose the biological background of the mirror neuron model and the common coding paradigm. Section 3 connects the biological modeling to concrete computational framework and shows how it is applied in a robot setting. The experimental setting and results of following and turn taking simulations are shown in Section 4. The discussion (Section 5) summarizes the results and puts this work into perspective.

2. MNS model for inter agent interaction

The common coding paradigm postulates parity between perception and action, i.e. the action and perception arise

* Corresponding author.
 E-mail addresses: e.i.barakova@tue.nl (E.I. Barakova), tino.lourens@philips.com (T. Lourens).

simultaneously [10]. A core assumption of the common coding paradigm is that actions are coded in terms of the perceivable effects (i.e. the distal perceptual events) that they should generate. It has the advantages to an information-processing paradigm which is unable to explain perception in many cases related to direct action [1,3]. Common coding paradigm has more solid foundations than selection paradigm and Gibson's theory of direct perception [7] which fail to explain another group of phenomena like memory and imagination that can certainly originate an action by themselves [1,2].

A growing body of behavioral and neurophysiological studies support the grounding principles of the common coding paradigm. As first evidence for direct matching between action perception and action execution came the discovery of 'mirror neurons' in the ventral pre-motor cortex (PMv) of the macaque monkey [20–22]. Mirror neurons fire both when monkey carries out a goal-directed action and when it observes the same action performed by another individual [23], i.e. the perception and the action are likely coded in the same way, by the same structures. More recently, it was found that a subset of these mirror neurons also responds when the final part of a previously seen action is hidden and can only be inferred [26]. Therefore, the observation of an action activates action representations to the degree that the perceived action and the represented action are similar [13]. Specific neurons in this region respond to the representation of an action rather than to the action itself.

It can be inferred that the sensory and the motor activations that represent the same action or intention are related to the activation of the same area in the brain. One such area is the PMv. Since the observed, executed, and imagined actions are related to an activation in a common representation, we schematically show this phenomenon like the activation from the three events is projected to the common representation (Fig. 1, the scheme of the individual agent). Actually, in case of an executed action the activation in the pre-motor and motor areas occurs in a very short time for the behavioral time-scale interval, i.e. practically co-occurs. In the case of two agents that share perceptual space the common representation for perception and action for each agent will create a basis for an interaction behavior, as shown in Fig. 1.

Actually, there is more than one representational structure that gets active by the same event encountered by the sensory and the motor states. Most of the frontal motor areas receive robust sensory input (visual and somatosensory) from the parietal lobe. This pattern of connectivity supports relatively specialized
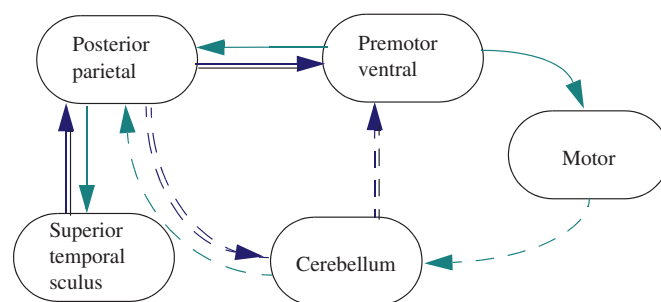

Fig. 2. The information flow by the observation and the imitation of the same action in an individual agent. Double line arrows mark the path of the inverse model that signifies action observation. Single line arrows show the forward path for execution of imitated actions.

fronto-parietal area for sensorimotor integration. A posterior area with mirror neuron properties is located in the rostral part of the inferior parietal lobule (IPL). Both areas form the MNS. The main visual input to the MNS originates from the posterior sector of the superior temporal sulcus (STS). Together, these three areas form a 'core circuit' for imitation, one of the basic building components of social behavior. The information flow from the parietal MNS, which is mostly concerned with the motoric description of an action, reaches back to the STS. By macaque, STS and the equivalent of IPL share patchy connections that overlap particularly well with the locations in which neurons respond specifically to complex body movements. The STS although considered to be a part of the 'mirror system' [18] do not show any motor activation itself. In spite of lacking mirror properties, STS neurons seem to 'understand' actions quite well, and it is plausible to assume that they send (via IPL) preprocessed signals about actions to the pre-motor areas that include information about the goal or the meaning of the observed action.

To construct a computational model that can facilitate the imitation and interaction functionality, we have modeled the three interconnected structures with lattices of neurons [4]. The direction of the connectivity between the structures differs while different actions take place. Fig. 2 denotes the information flow by the observation and the imitation of the same action. The solid arrows show the part that has been considered in our model for achieving the imitation functionality.

3. Oscillatory neural dynamics of the MNS and experimental results

From the framework proposed in the previous section it becomes apparent that the mirror neuron model that materializes the common coding paradigm is a useful tool for modeling interactive behavior. Inter agent interaction is initiated from the representation of the movement of each robot within the neural structures of the partner robot. To achieve the imitation functionality and create a model that is suitable for robotics, we have to make some simplification. We base our core scheme for imitation learning on conceptual model of [12], whose experimental work has shown that there are anatomical connections between the macaque analogous of STS, IPL, and PF areas, and therefore a Hebbian learning rule can be applied. The particular network that has been used to simulate the imitation functionality is shown in Fig. 3.

In this scheme the role of the STS neurons have the function to transfer the sensory (visual) stimuli and to account for the influence of the inhibitory neurons. For a robotic setting, modeling of the STS area can be reduced to the influence of the inhibitory neurons. Therefore, the sensory signals project directly to the
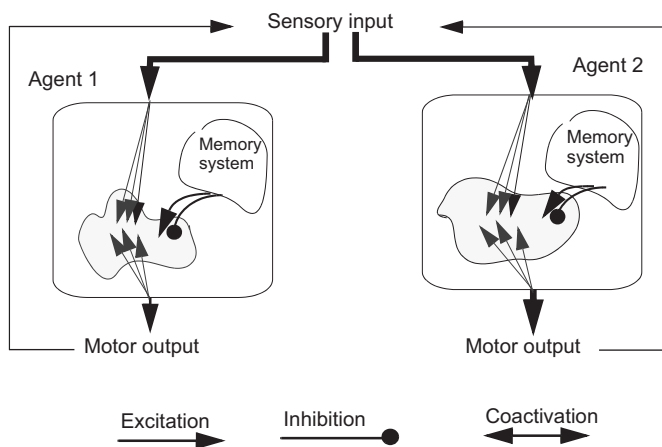

Fig. 1. The common coding theory suggests that sensing and action is related to the activation in the same internal representations. Moreover they can be activated by endogenous factors. Common coding for two agents that share perceptual space as a basis for modeling interaction behavior.
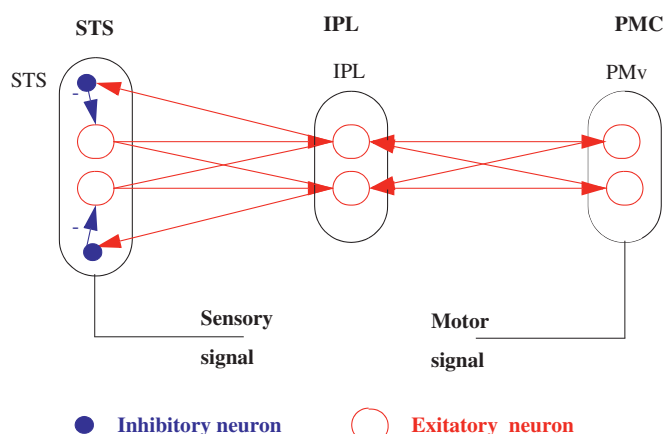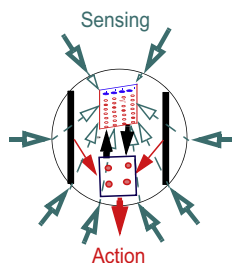
Fig. 3. Core circuit for imitation.

Fig. 4. Robot architecture with mirror circuit.

IPL area which is associated with multi-sensory integration. The motor information or the information from the movement of the wheels is co-activated in the simulated PMv area, which has sensorimotor integration functionality. The bidirectional projections between the two areas will insure that both areas represent the sensory and the motor signals.

The embodied implementation of this model is shown in Fig. 4. The eight range sensors of each robot project to the sensory integration area that resembles the functionality of the joined STS–IPL areas. The two wheels project to the sensorimotor integration area, which resembles the PMv, as shown in Fig. 4.

Self-organization of rhythmic activity is a fundamental characteristic of biological systems. In addition, rhythmic activities are found in any level of the hierarchical structure, i.e. from the biochemical to the socio-biological level. At neuronal level, single neurons and networks respond with sequences of spike bursts to a strong input. From a dynamical point of view, this means that such neurons generate a limit cycle oscillation in response to constant stimulation. The natural frequency or eigenfrequency of the damped oscillation is a result of two opposing effects, often modeled by the combined effect of excitatory and inhibitory neurons. Modeling of interacting non-linear oscillators is a complex problem. Cohen et al. [28] proposed a simplification of this problem, by making explicit that only the phase of each oscillator needs to be considered when the coupling between the oscillators is weak.

We suggest to use entrainment of endogenous oscillators for modeling turn taking behavior. The mirror neuron paradigm that allows the behavior of each robot to be represented in the neuronal structures of its partner makes possible the oscillatory dynamics of the turn taking process to be modeled through the individual agents.

The mutual interaction between two robots has to emerge through self-organizing entrainment of oscillatory neurons.

To check this hypothesis, we construct experimental setting that consists of two parts. In the first part each robot had to build sensory-motor experiences by exploring an environment that comprises a circular arena inhabited by another robot, see Fig. 5a. In the second part, interactive turn taking behavior emerged, based on the established oscillatory sensorimotor couplings.

Initially, we constructed the experimental scenario that performs following behavior for the training phase, see Fig. 5a. Both robots were consequently taking the role of the follower, in order to establish adequate 'mirroring' couplings between the lattices that resemble the IPL and PMv areas. IPL served as multi-sensory integration area and PMv is primarily sensorimotor integration area.

In the initial experiment the robot-follower determined its shortest distance reading, which signaled the presence of the partner robot, as shown in Fig. 5b. The angle between the heading direction of the follower robot and the placement of the sensor that detects the shortest distance to the partner robot served as a sensory input signal. The angle of the change of direction of the same robot calculated by the detected torque on its wheels is the motor output value that is fed as a motor input to the PMv area.

If there are two neurons that are activated by the sensory and motor signals, they will synchronize in a time slot dependent on the input activation values and the connection strength. In Fig. 6 it is shown that the stronger the coupling between the neurons, the earlier synchronization takes place. In this simulation, for
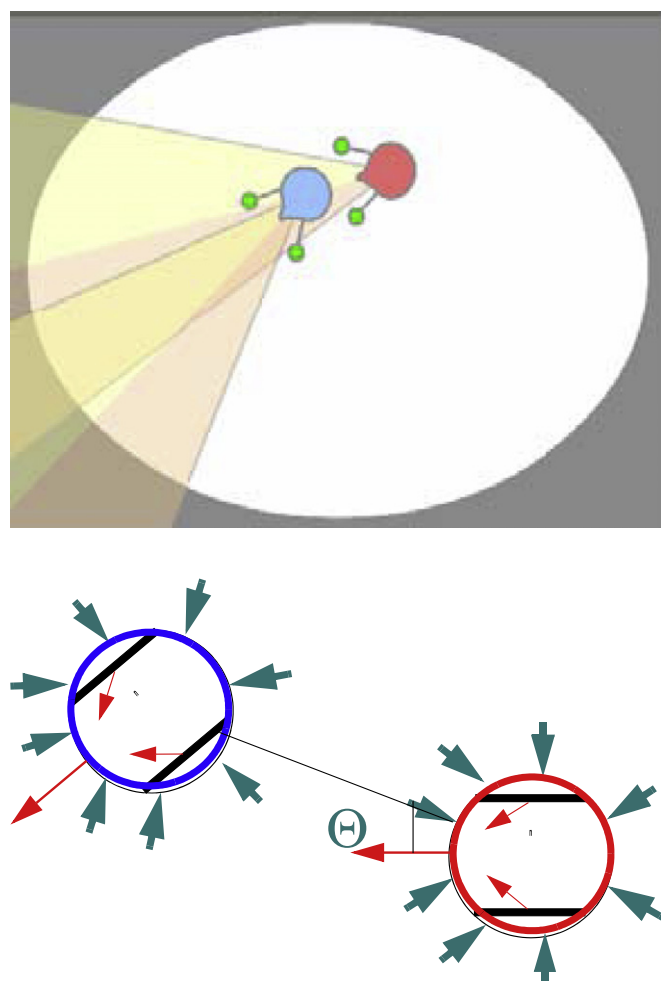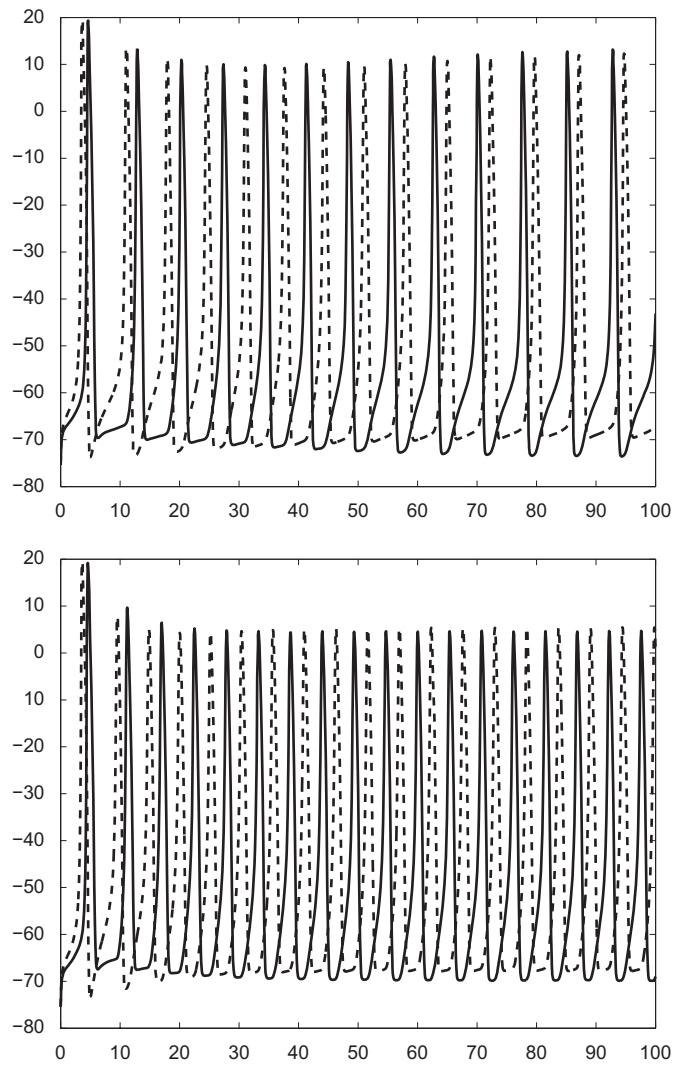

Fig. 5. (a) Scenario for following and turn taking behaviors based on the tag game. (b) The shortest sensor measurement determines the relative position and direction of the partner robot for the following interaction.
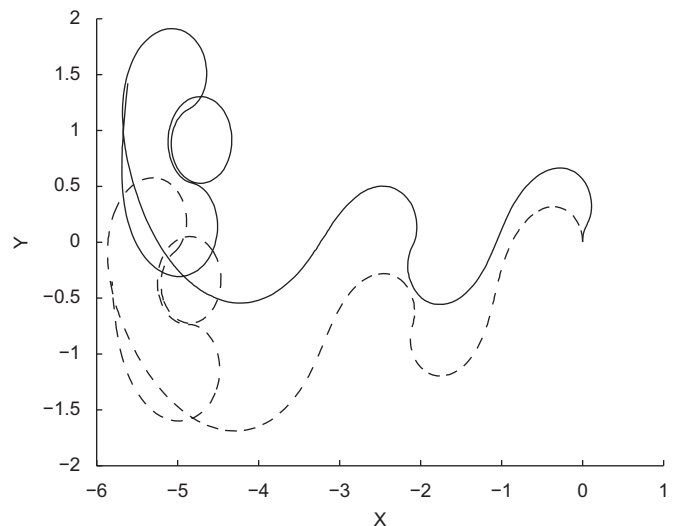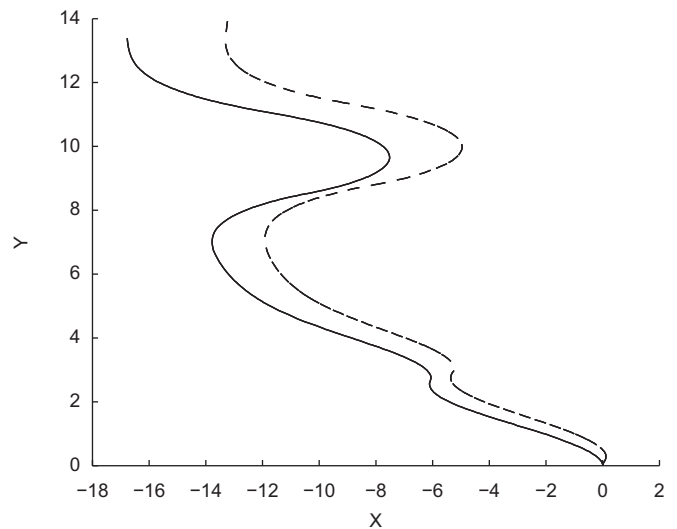
We used lattices of neurons with

$$\alpha_i = \beta_i - I_i(t) \tag{5}$$

where I_i denoted the sensory input and β_i represented the native frequency and the stabilization coefficient, respectively. The connections between the neurons that receive sensory stimulation and the neurons that receive motor stimulation are adapted by Hebbian rule. As illustrated in the experiment depicted in Fig. 6, the lower synaptic strengths do not result in synchronization, and the higher the value of the synaptic strength gets, the synchrony occurs faster. The adaptation of the connection strengths between these lattices are adapted in a way that after training both areas will reflect the common sensorimotor representation that is the basis for interaction behavior:

$$\Delta w_{lk}^{PM-IPL} = \alpha(PM_l - \overline{PM_l})(IPL_k - \overline{IPL_k}) \tag{6}$$

where $\overline{PM_l}$ and $\overline{IPL_k}$ are the average activation values of units l and k over a certain time interval. IPL–PMv synaptic plasticity has the following dynamics: the connection between them is strengthened if both of them are simultaneously active and weakened if the activation of one decreases.

In the first experiment the robot-follower tends to synchronize its motion direction with the motion direction of the leading robot. Since this is only a computer simulation, we assume that

Fig. 6. Synchronization between two neurons that are activated by sensory and motor signals, respectively. (a) Coupling strength is equal to 6. (b) Coupling strength is equal to 8.

demonstration purposes only, more realistic but relatively complex model of a neuron was used.

For the robot simulations the neurons of each robot's mirror system are simulated as oscillators:

$$\theta(t) = \omega t (\mathrm{mod}\, 2\pi) \tag{1}$$

The above equation determines the change of rate of the phase with the time. ω is the frequency of the limit cycle oscillation. The phase is periodic over the range 2π. If a synaptic coupling H connects two neurons, their phase equations will be represented in the following way:

$$\frac{\mathrm{d}\theta_1}{\mathrm{d}t} = \omega_1 + H_1(\theta_2 - \theta_1) \tag{2}$$

$$\frac{\mathrm{d}\theta_2}{\mathrm{d}t} = \omega_2 + H_2(\theta_1 - \theta_2) \tag{3}$$

where indexes 1 and 2 refer to the first and the second neurons, respectively. H_1 and H_2 are periodic with period 2π. We initiate all the neurons with equal coupling strengths. The rate of change of the phase by which the oscillators will synchronize is determined by the following equation:

$$\frac{\mathrm{d}\theta_1}{\mathrm{d}t} = \omega_1 + \alpha_1 \sin(\phi - \sigma) \tag{4}$$

Fig. 7. Movement imitation behaviors.

the distance sensors are positioned at an equidistant angle around the robot, so we can have smoother trajectory of the follower.

After neurons from the two lattices synchronize, the two simulated robots express a simple form of social behavior. The leader robot performs movements with different complexity, and the follower (dashed lines) imitates it from its movement perspective, as shown in Fig. 7.

At the second part of the experimental scenario, the emergent turn taking was observed. The role of the robot, being follower or leader at the present moment depends on which robot is 'within the visual field' of its partner. For the training phase, the tag game is simulated, by which the runner and the tagger functions change between the robots once the tagger reaches the runner.

After training, the emergent turn taking has to take place which is expressed by symmetry breaking process after a period of synchronization. This way the leading robot can become a follower and later take the lead again. The turn taking, in analogy to humans, takes place as a result of some subtle or explicit external stimulation. For the case of the tag game, the external stimulation is usually caused by losing the runner robot from the perceptual field, caused by reaching the end of the arena or other reason for escape of the runner robot.
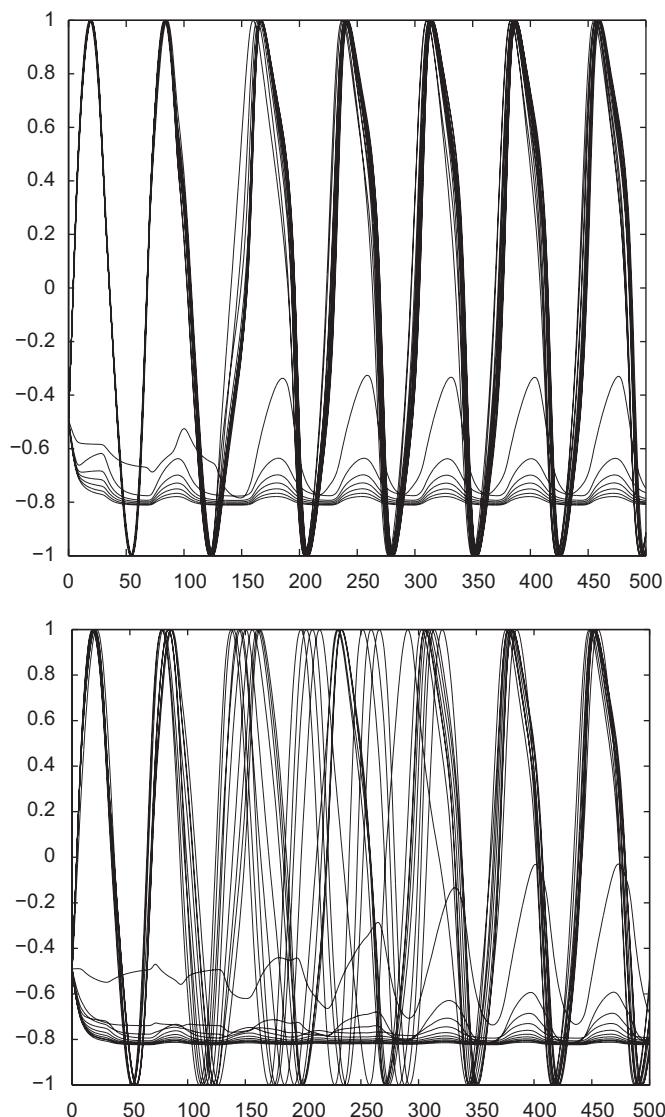

Fig. 8. (a) Neural activation during following behavior. (b) A typical re-synchronization in the case of turn taking.

Fig. 8 shows the neural activation during turn (the left plot) and in the period of turn taking (the plot on the right). The desynchronization of the neurons in central part of the right plot corresponds to the moment of losing the runner robot from the perceptual field. At that period previous follower changes its role to a runner, and vice versa.

4. Discussion

Social interaction has wide spectrum of expressions as synchronous movements, turn taking, gaze sharing, following, imitation and conversation. We have simulated simple interaction behaviors of following and turn taking. In the training phase, the simulated following and the simulated tag game helps to gather examples and establish the sensorimotor couplings between the two robots. In the test runs, there is not an external control that will cause the turn taking behavior. The turn taking is caused by changing of synchronous firing of the oscillatory neurons. Although external events are initiating turn change, turn taking does not take place only by the same conditions as during the training—turn taking has emergent properties due to nonlinear oscillations and their interaction. These results resemble turn taking in speech: an upcoming end of a turn is anticipated by perceiving eye gaze, body movement, or other queues by the speaker, or indicated in a subtle manner by the listener.

Important questions for designing a movement interaction lies in the respective computational role of each brain area that subserves the internal simulations and shared representations between self and others. We based our model on the simplified mirror neuron network, in which the mirroring functionality is obtained via the self-organization of synchronized neural firing in two robots that share perceptual space. The emergence is an important element, but better understanding of underlying processes and computations will increase the possibilities and reliability of the interaction behavior.

References

[1] E.I. Barakova, T. Lourens, Event based self-supervised temporal integration for multimodal sensor data, J. Integrative Neurosci. 4 (2) (2005) 265–282.

[2] E.I. Barakova, T. Lourens, Efficient episode encoding for spatial navigation, Int. J. Systems Science 36 (14) (2005) 887–895.

[3] E.I. Barakova, Social interaction through movement: concepts from perception-action interplay, in: Feijs, Kyffin, Young (Eds.), Proceedings of DESFORM Conference, November 25–27, 2006.

[4] E. Barakova, Y. Yamaguchi, Sensorimotor synchronization for action prediction: a mirror neuron model, in: the Proceedings of the Eighth BSI Retreat, October. 31–November 1 2005, Tochigi Japan.

[5] E.I. Barakova, Learning reability: a study on indecisiveness in sample selection, PrintPartners Ipskamp B.V., March 1999, ISBN: 90-367-0987-3.

[6] E.A. Di Paolo, Behavioral coordination structural congruence and entrainment in a simulation of acoustically coupled agents, Adaptive Behav. 8 (2000) 25–46.

[7] J.J. Gibson, The Senses Considered as Perceptual Systems, Houghton-Mifflin, Boston, 1966.

[8] C.E. Ford, S.A. Thompson. Interactional units in conversation. in: E. Ochs, E.A. Schegloff, S.A. Thompson (Eds.), Interaction and Grammar. New York, Cambridge University Press, pp. 134–184.

[9] J.E. Fox Tree, Coordinating spontaneous talk, in: L. Wheeldon (Ed.), Aspects of language production, 2000, Taylor & Francis, Philadelphia: pp. 375–406.

[10] B. Hommel, J. Musseler, G. Aschersleben, W. Prinz, The theory of event coding; a frame-work for perception and action, Behav. Brain Sci. 24 (2001) 849–878.

[11] H. Iizuka, T. Ikegami, Adaptability and diversity in simulated turn-taking behaviour, Artif. Life 10 (2004) 361–378.

[12] C. Keysers, D.I. Perrett, Demystifying social cognition: a Hebbian perspective, ends Cogn. Sci. 8 (2004) 501–507.

[13] G. Knoblich, R. Flach, Action identity: evidence from selfrecognition, prediction, and co-ordination, Conscious Cogn. 12 (2003) 620–632.

[14] A.N. Meltzoff, M.K. Moore, Newborn infants imitate adult facial gestures, Child Dev. 54 (1983) 702–709.

[15] A.N. Meltzoff, M.K. Moore, Imitation in newborn infants: exploring the range of gestures imitated and the underlying mechanisms, Dev. Psychol. 25 (6) (1989) 954–962.

[16] E. Oztop, M. Kawato, M. Arbib, Mirror neurons and imitation: a computationally guided review, Neural Networks 19 (3) (2006) 254–271.

[17] T. Pavlidis, Biological oscillators: their Mathematical Analysis, Academic Press, New York, London 1973.

[18] G. Rizzolatti, L. Craighero, The mirror-neuron system, Annu. Rev. Neurosci. 27 (2004) 169–192.

[19] G. Rizzolatti, L. Fadiga, L. Fogassi, V. Gallese, (2002). From mirror neurons to imitation: facts and speculations, in: A.N. Meltzoff, W. Prinz (Eds.), The Imitative Mind: Development, Evolution and Brain Bases, Cambridge University Press, Cambridge, (pp. 247–266).

[20] G. Rizzolatti, Cortical mechanisms subserving object grasping and action recognition: a new view of the cortical motor functions, in: The New Cognitive Neuroscience, M. Gazzaniga, (Ed.), MIT Press, Cambridge, MA, 2000, pp. 539–552.

[21] G. Rizzolatti, A unifying view of the basis of social cognition, Trends Cogn. Sci. 8 (2004) 396–403.

[22] G. Rizzolatti, Localization of grasp representations in humans by PET: 1. Observation versus execution, Exp. Brain Res. 111 (1996) 246–252.

[23] G. Rizzolatti, L. Fadiga, V. Gallese, L. Fogass, Premotor cortex and the recognition of motor actions, Brain Res. Cogn. Brain Res. 3 (1996) 131–141.

[24] S. Schaal, A. Ijspeert, A. Billard, Computational approaches to motor learning by imitation, Philos. Trans. R. Soc. Lond. B 358 (2003) 537–547.

[25] G. Stoet, B. Hommel, Interaction between feature binding in perception and action, in: W. Prinz, B. Hommel (Eds.), Common mechanisms in perception and action: Attention & Performance XIX, Oxford University Press, Oxford: 2002, pp. 538–552.

[26] M.A. Umilta, E. Kohler, V. Gallese, L. Fogassi, L. Fadiga, C. Keysers, G. Rizzolatti, I know what your are doing: a neurophysiological study, Neuron 31 (2001) 155–165.

[27] M. Wilson, T.P. Wilson, Related articles, links an oscillator model of the timing of turn-taking, Psychon. Bull. Rev. 12 (6) (2005) 957–968.

[28] A.H. Cohen, P.J. Holmes, R.H. Rand, The nature of the coupling between segmental oscillators of the lamprey spinal generator for locomotion: a mathematical model, J. Math. Biol. 13(3) (1982) 345(H36(B9.

Emilia I. Barakova is affiliated with the Department of Industrial Design of the Technical University of Eindhoven, The Netherlands, and simultaneously holds a Visiting Researcher position at RIKEN Brain Science Institute in Japan. She has expertise in robotics and computational neuroscience, learning methods, and design. Currently she is working on social interaction of intelligent agents. She has Masters degree in Electronics and Automation from Technical University of Sofia (Bulgaria) and PhD in Computing Science from Groningen University (The Netherlands, 1999). She has worked at different research institutes, among which RIKEN Brain Science Institute (Japan), GMD-Japan Research Laboratory (Japan), Groningen University (The Netherlands), Eindhoven University of Technology (The Netherlands), and Bulgarian Academy of Science. She is a reviewer of many conferences and special issues of journals, and serves as book review editor of Journal of Integrative Neuroscience. She has been organizing several scientific workshops, special sessions and one IEEE conference and has participated in two Robocup competitions.

Tino Lourens received an MS degree in computer science in 1993 and a PhD degree in 1998, both from the University of Groningen, The Netherlands. He was a researcher at The Netherlands Organization for Applied Scientific Research at The Human Factors Research Institute (TNO HFRI, 1998–1999), Japan Science and Technology Corporation (JST), ERATO, Kitano Symbiotic Systems Project in Tokyo, Japan (1999–2001), GMD-JRL, Kitakyushu, Japan (2001–2002), and Honda Research Institute (HRI) in Wako, Japan (2003–2006). Currently he is with Philips Healthcare. Dr. Lourens has published over 50 papers in refereed journals and conferences and is the author of visual programming environment TiViPE. His research interests include object recognition from real world images, biological models of vision, visual programming, parallel and distributed processing, brain-like computing, multimodal integration, memory, learning, and intelligence.