



HAL
open science

”Synchrony” as a Way to Choose an Interacting Partner

Syed Khursheed Hasnain, Philippe Gaussier, Ghiles Mostafaoui

► **To cite this version:**

Syed Khursheed Hasnain, Philippe Gaussier, Ghiles Mostafaoui. ”Synchrony” as a Way to Choose an Interacting Partner. IEEE Conference on Development and Learning and Epigenetic Robotics, Nov 2012, San Diego, California, United States. hal-00749002

HAL Id: hal-00749002

<https://hal.science/hal-00749002>

Submitted on 6 Nov 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

“Synchrony” as a Way to Choose an Interacting Partner

Syed Khursheed Hasnain

Neurocybernetic team, ETIS

ENSEA, University of Cergy-Pontoise
95302 France.

Email:syed-khursheed.hasnain at ensea.fr

Philippe Gaussier

Neurocybernetic team, ETIS

ENSEA, University of Cergy-Pontoise
95302 France.

Email:gaussier at ensea.fr

Ghiles Mostafaoui

Neurocybernetic team, ETIS

ENSEA, University of Cergy-Pontoise
95302 France.

Email:ghiles.mostafaoui at ensea.fr

Abstract—Robots are poised to fill a growing number of roles in today's society. In future, we could have robots expected to behave as companions in our homes, offices etc. Moving to social robotics imply to address several issues related to human-robot interactions for instance, how the robot can develop an attentional mechanism and select an interacting agent among several interactants. We took our inspiration from neurobiological and psychological studies suggesting that synchrony is an essential parameter for social interaction. We assumed that synchrony detection could be used for initiating human robot interaction. We present a neural network architecture able to focus the attention of the robot and to select an interacting partner on the basis of synchrony detection.

I. INTRODUCTION

In future years, robots may become a part of our social lives, consequently, they are supposed to live with human being and are expected to behave as companions in offices, factories, homes and other social gathering. Moving to social robots introduce new challenges to manage the development of Human-Robot interaction. For instance, how can a robot choose a partner for interaction and how its attentional mechanism works to focus its attention on regions of interest? In other words, how can a robot be able to discriminate the relevant stimulus from multiple ones.

In order to design a robot that can interact with humans, we need to understand how humans interact with each other. Psychological studies of dyadic interactions as well as neurobiological data in motor coordination suggest that synchrony is an important parameter for human-human-interaction.

An interesting aspect of synchronized behaviors in humans is its unconscious nature (unintentional synchronization among interactants). According to psychologists, during verbal interactions, several non-verbal behaviors for instance, head movements, expressions on face and many other gestures are also associated [1]. An important aspect of these non-verbal communications is their timing and synchrony according to the partner's behavior.

Studies of Interpersonal motor coordinations also point out these unconscious synchronizations or communications among people. For example when an adult and a child walk together, unconsciously, they try to maintain the same phase by constantly adjusting the length or the frequency of their steps [3]. Moreover, individuals have their own clapping frequencies

but people tends to applaud in a synchronized way [4]. In addition to the above examples, Issartel et al. studied interpersonal motor-coordination between two participants when they are instructed to not coordinate their movements between each other. Results showed that participants could not avoid unintentional coordination [3]. This reflects that when visual information is shared between two people they coordinate (unintentionally) with each other.

In recent researches on schizophrenia, Varlet and al. examined intentional and unintentional social motor coordination in participants oscillating hand-held pendulums from the wrist. Their results show that for a participant suffering from schizophrenia, intentional motor coordination remained impaired however, unintentional social motor coordination was preserved. This study demonstrate that unconscious communication sustains even though patients suffering from social interaction disorders [5].

Researches in neurobiology also acknowledge the synchronization of brain activities during social interactions. Several studies used fMRI and EEG to record the brain activities during social interaction. In an study carried out by Stephens et al. [6] they recorded the brain activities of a speaker (reciting a monologue) and then scanned the brain of a listener who was participating, it was found that there is a temporal and spatial coupling between listener and speaker. Recently, Dumas et al. [7] revealed, using hyperscanning, the emergence of inter-brain synchronization across multiple frequency bands during social interaction (with a millisecond synchrony).

Keeping in view the importance of synchrony in social interaction, it has also been widely studied and used in robotics. Andry et al. proposed synchrony as an internal reward for learning [9]. Prepin and al. also used the level of synchrony as a reinforcement signal for learning [10]. Blanchard and Canamero proposed a velocity detection system to synchronize the movements of two robots to improve the reactivity of agents to changes in their environment [11]. Marin et al. underlined that motor resonance between robots (humanoid) and humans could optimize the social competence of human-robot interactions [12].

Moreover, studies of developmental psychology also acknowledged synchrony as a prime requirement for interaction between a mother and her infant. An infant stops interacting

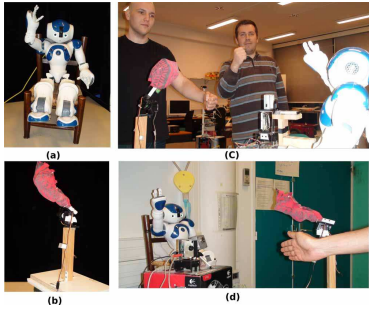


Fig. 1. Setup for our experiments. (a) Nao robot (b) Basic automaton (1 degree of freedom) (c) & (d) Overall setup for human-robot interaction.

with his mother when she stops synchronizing with it [2]. Inspired by these observations, we assumed that this notion of unconscious synchrony can be used as a starting point in Human-Robot interactions.

In this paper, we propose a developmental approach to study unconscious or unintentional synchronization during human-robot interaction. We developed a neural network model permitting the robot to: first, be aware of his own dynamic by learning (babbling step) to link his actions (proprioception) and the induced visual stimuli (optical flow), and then, be able to automatically select a partner among many interactants using immediate imitation and locate its focus of attention on the basis of synchrony detection.

II. EXPERIMENTAL SETUP

A minimal experimental setup is used to avoid complexities (figure 1), it includes a basic automaton (1 DoF), Nao humanoid robot, human partners and an external camera. In human robot interaction experiments, frame rate or sampling frequency of the system is an important parameter. Instead of Nao's camera (frame rate limited to 10 Hz through ethernet connection) an external camera is used to allow our architecture to work on the frame rate of 30 Hz.

Phase Locking Value (PLV) [13] is used to analyze synchrony between two signals. The PLV is defined by $PLV_{x,y} = \frac{1}{N} |\sum_{t=1}^N \exp(i(\phi_x - \phi_y))|$, where N is the number of samples and $\phi_x - \phi_y$ is the phase difference between two signals. The PLV value is close to 1 for synchronized signals and approaches 0 otherwise. Videos of our experiments can be found on : <http://www.etis.ensea.fr/neurocyber/Videos/synchro/>

III. SIMPLE HUMAN-ROBOT INTERACTION USING OPTICAL FLOW

As a first step towards human-robot interactions, we developed a dynamical interaction model where two agents (human and robot) synchronize dynamically by influencing each other. Specially, our designed architecture gives minimal abilities to Nao to adopt the phase and frequency of interacting agent. Velocity vectors of the perceived movements (Nao's visual field) are estimated by an optical flow algorithm. These velocity vectors represent the visual stimuli and inputs for our architecture.

The oscillator module in this dynamical interaction architecture (figure 2(a)) is identical to [14]. It consist of two neurons N_1 and N_2 inhibit each other proportionally to the variable β . The oscillating frequency is the function of the variables α_1 , α_2 and β :

$$N_1(t+1) = N_1(t) - \beta N_2(t) + \alpha_1 \quad (1)$$

$$N_2(t+1) = N_1(t) + \beta N_2(t) + \alpha_2 \quad (2)$$

As it is shown figure 2(a), the Nao's arm is linked with the oscillator and it oscillates according to the its default frequency. The visual activities in front of Nao are processed by an optical flow algorithm to estimate the velocity vectors which are then transformed into negative and positive activities. If an upward motion is realized, it is considered as a positive activity. On contrary, the downward motion is considered as a negative activity. Figure 3(f) is a real picture (visual activities in front of Nao) captured by the camera and Figure 3(e) shows the corresponding positive and negative activities. Two different agents (a human and an automaton) move in different direction. Human's movement direction is upward and induces positive activity (filled black color). The Automaton's direction is downward and induces negative activity (unfilled pixels). The equation 1 of the oscillator is rewritten as:

$$N_1(t+1) = N_1(t) - \beta N_2(t) + \alpha_1 + f' \quad (3)$$

Where f' is the coupling energy or the energy induced by the external visual activities (positive and negative activities extracted from optical flow). By using this model, the external stimuli induced by the interactant motion influence directly the oscillator controlling the robot arm making NAO modifying its own movements and adopting the interactant ones.

Figure 3(a) details the signals of both Nao's arm and the human (arm or hand) while trying to interact by imitating each other. In the onset of experiment, movements of both agents are not synchronized. Consequently, PLV (measure of synchrony) shows lowest value (see figure 3(b)) in region of asynchronous movements. As shown in Figure 3(a) and 3(b), during the interaction between Nao and a human both are synchronizing successively as the time passes. Increasing tendency in PLV values reflects the emerging synchrony between the two agents. Figures 3(a) and 3(b) also illustrate that, after a some time, both interactants are fully synchronized and the PLV values corresponds to synchronized interaction are at maximum values. For this experiment, Nao's standard frequency was 0.428 Hz and human oscillations were between 0.4615 Hz (7.8% higher than Nao's frequency) to 0.476 Hz (11% higher).

IV. SELECTION OF PARTNER

After developing a basic architecture initiating automatically a human-robot interaction by synchronizing agent's movements (in an imitating framework), we develop an architecture (Figure 2(b)) capable of choosing an interacting partner among various interacting agents. This architecture is based on

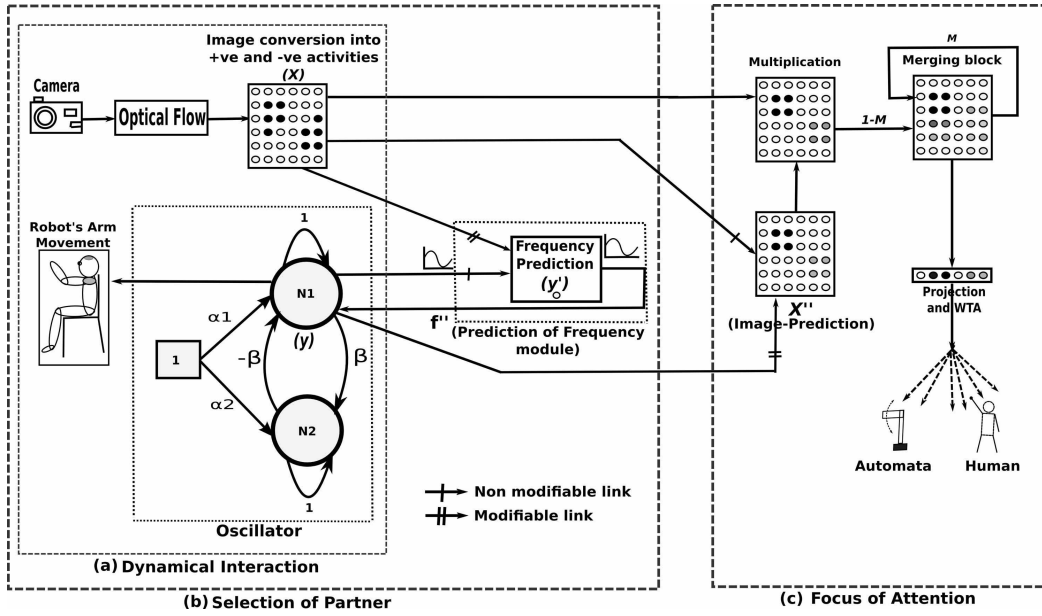


Fig. 2. (a) Dynamical Interaction model (b) Selection of Partner: select a interacting partner on the basis of synchrony detection among various interacting agents. (c) Shows attentional mechanism architecture.

synchrony detection and it can be segregated into two parts. The first one is dynamical interactions (stated in the previous section) and the other one is the frequency-prediction module. As stated in the previously section (Figure 2(a)), the oscillator that governs the Nao's motion was directly linked to external visual stimuli (positive and negative activities from optical flow). Now, the coupling activities are linked with frequency-prediction module (f'') (Figure 2(b)). The advantage of this indirect coupling from the frequency-prediction block is to make sure that our algorithm will choose an interacting agent that is approximately identical to its own dynamics (learned by the frequency-prediction block). The equation 3 can be rephrase as $N_1(t+1) = N_1(t) - \beta N_2(t) + \alpha 1 + f''$. Here, f'' is the coupling energy feed by the frequency-prediction block. The other variables remains unchanged.

The frequency-prediction module (y') is connected to the Nao's oscillator (y) with a unconditional (non modifiable) link and visual activities (X) is connected with a conditional link (modifiable link). The frequency-prediction (y') block learns the Nao's oscillating frequency as a weighted sum of visual stimulus (X). The activity of frequency-prediction (y') neuron in equation (4) can be computed by $X \rightarrow y'$ synapses only and corresponds to the prediction of future states.

$$y'_i(t) = \sum_{k \in X} W_{X_k - y'_i} X_k \quad (4)$$

The learning (equation 5) takes place in the $X \rightarrow y'$ synaptic weights and it is based on Normalized Least Mean Square algorithm (NLMS) [15] (Synaptic learning modulation η is added):

$$W_{X_j - y'_i}(t+dt) = W_{X_j - y'_i}(t) + \alpha \eta \cdot \frac{y_i(t) - y'_i(t)}{\sum_{k \in X} X_k(t)^2 + \sigma 1} \cdot X_j(t) \quad (5)$$

Where y' stands for frequency-prediction, X for the image of visual activities and y for the Nao's arm Oscillator, the learning rate is represented by α and $W_{X_j - y'_i}$ is the synaptic weight from visual activity (X) neuron j to frequency-prediction neuron i , y_i is the activity transmitted to neuron i by the oscillator, it is a target signal for the Least Mean Square (LMS) algorithm. To improve the LMS convergence in the online learning case, we introduced the learning modulation η . The particularity of the NLMS over the standard LMS is the normalization term $\sum_{k \in X} X_k(t)^2 + \sigma 1$. $\sigma 1$ is a small value used to avoid the divergence of the synaptic weights for low visual activities (X).

Now, we analyses the complete architecture. Our selection of partner algorithm works in two steps: learning and testing phase. In the learning stage, Nao moves according to his own dynamics (its default frequency). Firstly, the frequency-prediction block that was inactive due to absence of visual activities begins now predicting robot's oscillation as a weighted sum of its own visual activities. The frequency-prediction block learns the associations of Nao's motion (optical flow induced by Nao's arm). Consequently, Nao's oscillator is also modified (as described in section 3). This process of modifying, learning and adapting continues and converges after some time. This adjustment is similar to the infant's elementary behaviour by which infants gain self reflective abilities [16]. After this phase, Nao learns to predict oscillatory movements similar to his own movement. When a human or other agents interact with a frequency similar to NAO's own dynamics (or the learned one), weights (learned on modifiable link) are

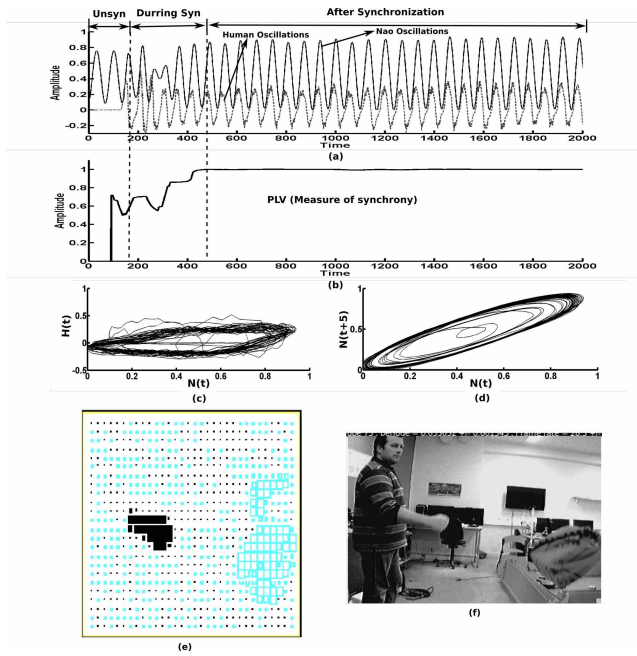


Fig. 3. Describes how the two agents are synchronized. (a) Shows two signals (human and Nao’s modifiable oscillator). (b) PLV measurement.

associated with the visual stimulus produced by the interacting partner and the robot adopts the phase and frequency of the interacting agent. However, if the interacting oscillations are not similar to its own dynamics (the learned frequency), the weights (modifiable link) could not be associated with the visual activities and Nao robot continue to oscillate at its standard frequency. The same holds for multiple agents case. If two agents interact with Nao and only the one of them having a frequency close to Nao’s learned frequency will be chosen by our architecture.

For this selection of partner experiment, we use two agents to interact with Nao robot, in addition to human partner, a basic automaton is introduced (Figure 1(d)). The results of selection of partner algorithm are shown in section VI. In this experiment Nao’s standard frequency was 0.407 Hz, automaton: 0.4318 Hz (6% higher than Nao’s default frequency) and the frequency of human partner varies between 0.36 Hz (11% low) to 0.38 Hz (6% low).

As it is described in the previous sections, Nao is able to select and interact with a partner having movements rhythmically similar or close to his inner dynamic represented by just one oscillator. One of the main goals of our work is to use this unintentional synchrony as an automatic starting point for human/robot interactions leading to a developmental learning of more complex tasks by imitation games. To do so, the robot must be able to select (synchronize) and interact with multiple agents having different dynamics of motion (or one agent with more complex movements). Nao must be capable of selecting partners in a larger band of frequencies (different dynamics of motion).

Using the same model described previously, we introduced

three different oscillators A, B and C with respectively the following frequencies $f_A = 0.441Hz$, $f_B = 0.83Hz$ and $f_C = 1.153Hz$. These oscillators determine the Nao’s current oscillating frequency and switch according to the visual stimulus. When there is no visual stimulus, one of the oscillators is randomly selected (every 4 sec). This step simulates in a very simple manner a babbling state (3 simple gestures randomly performed). As the interacting agent arrives and interacts with a given frequency, the oscillator prediction module which is near to this frequency synchronizes with it. Our architecture selects, as a current oscillator for Nao, the one having less error with the visual stimulus. Figure 4 illustrates the results of our experiment. Initially, there is no human signal (until 800 time steps or 26.6 sec) that results in random selection of Nao’s oscillating frequencies. After 800 time units human interacts with frequency of 0.81 Hz, which is near to oscillator B, consequently, oscillator B is selected as Nao’s current motion frequency. After 1750 time units (58 sec), human changes his frequency near to oscillator C, this results in selection of oscillator C. The same is true for the selection of oscillator A.

V. ATTENTIONAL MECHANISM

Using the previous architecture (selection of partner), if two agents interacts simultaneously with Nao (with different frequencies), the robot selects an ”interacting” partner moving accordingly with its frequency and synchronizes with it. However, Nao can not locate the spatial position of the selected partner, it is due to the fact that the selection of partner algorithm operates on the perceived energy, it does not take into account the agent location. In order to locate the correct interacting partner, a new Focus Of Attention (FOA) algorithm dynamically learns, using spatial predictions, to focus on the right interacting agent defined by the selection of partner algorithm. Here, synchrony prediction is use as way to draw the robot’s attention. In our experiments, the robot’s head direction is used to indicate the current FOA.

Figure 2(c) details the architecture of attentional mechanism. This module also works in two phases i.e learning phase and testing phase. In the absence of visual stimuli, FOA moves randomly according to the noise. As an agent (human) comes with a similar frequency (by moving arm / hand), the image-prediction module (X'') learns the position of this visual stimulus as a weighted sum of robot’s oscillations. This makes it possible to predict the location of the synchronized partner. After this learning, If some other agent comes and interacts with a different (asynchronous) frequency (as shown in figure 1(c)), X'' predicts strongly the spatial location of the synchronized agent as compared to the asynchronous one (because X'' predicts on the basis of the weighted sum of the learned oscillations).

To locate the synchronized agent and to distinguish it among multiple stimuli, our architecture multiply the image-prediction X'' and the current visual stimuli. Then, a memory block computes the weighted average of these resultant multiplication and the previous iterations. Obviously, in the merging block, higher correlation values are obtained for synchronous

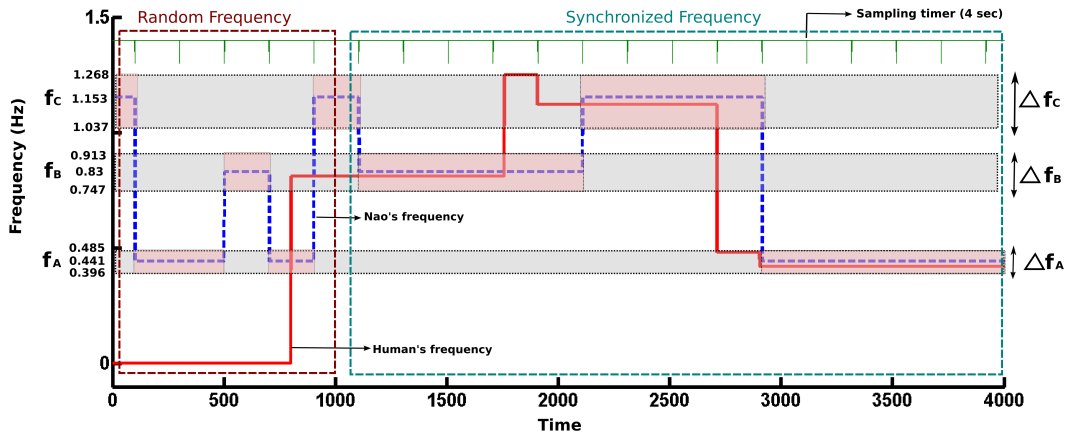


Fig. 4. Shows how the different oscillators (A, B and C) are selected according to the visual stimuli. Due to the initial lack of an interactant, a random selection of oscillator is observed. After 800 (26.6 sec) time units oscillators are chosen according to the frequency of the agents. Nao's oscillator selection (dotted line) according to human frequency (solid line)

movements. To determine the highest activated column (that corresponds the position of synchronized agent) all pixels of memory module are projected on x-axis and a Winner Takes All (WTA) finds the column of highest activity (Figure 2(c)). Finally, the architecture turns the robot's head towards the selected column.

The learning rule for movement-prediction (X'') is same as the frequency-prediction (stated above) except that the weights are normalized for smoothing the learning activity.

VI. EXPERIMENTAL RESULTS

In this section, we present the experimental results of our two algorithms: selection of partner and attentional mechanism. Results show that when two agents interact simultaneously with Nao and one of the agent moves similarly (in terms of frequency) as the Nao robot while other oscillates with a different frequency. Nao synchronizes with the agent having a similar frequency and selects it as a partner. Likewise, FOA architecture turns the robot's head towards the selected agent. Similarly, if both agents exchange the role of interaction (asynchronous agent becomes synchronous and synchronous agent become asynchronous) then, selection of partner and FOA architecture also switches the partner.

Figure 5 shows the experimental results. Figure 5(a1) sketches the curves of automaton and Nao's motion signal. In the start of the experiment (figure 5(a)) automaton interacts with Nao from the left side (about -20°). Both agents synchronizes after a short time by using our partner selection architecture. Figure 5(a3) describes the quality of synchrony between Nao and automaton in terms of PLV. In the beginning, PLV was at its lowest value but it increases slowly to the highest value as the interaction continues and the synchrony establishes between the two agents. Initially, there is no other agent in front of robot except Automaton therefore, FOA turns towards Automaton (figure 5(a4)). After 700 time units (23.33 seconds), a human starts interacting from the right side of

robot with different oscillations but he fails to disturb Nao-Automaton interaction (PLV continues to its higher values for automaton) and FOA continues to point out to automaton.

Now, the roles are switched. Human is advised to make similar movements as Nao while the automaton is adjust to a low frequency (figure 5(b1) and (b2)). Consequently, Nao also flips its role by synchronizing with human and select him as partner. PLV (measure of synchrony) also increases for human and decreases for automaton (figure 5(c3)). As the synchrony emerges between human and Nao, FOA also turns from automaton to human (figure 5(c4)). To validate our experiment we again switch the roles of two interacting agents after 2650 time units (88.3 sec). Consequently, this induces a switch of the focus of attention and the synchronized agent (figure 5(b)).

VII. CONCLUSION

In this paper, we presented a new model allowing the robots to select an interacting partner among multiple agent based on synchrony detection. We also demonstrated that prediction of synchrony (for spatial position) could be used as a tool to locate the Focus Of Attention. Our experimental results showed that when several agents interact with Nao and one of them moves in synchrony with the robot, Nao will select it as a partner.

From the psychological point of view, we were inspired by the unconscious communications between humans. The synchronous exchanges during social interactions are directly associated to the sensorimotor information of the two agents. These inter brain networks are "symmetric" in low frequency band while "asymmetric" in high frequency bands [7]. This could reflect the different processing levels of information. In our case, synchronization between two agents can be assumed as "symmetric" in low frequency band and Focus of attention can be associated with high frequency carrier.

Actually, we are studying three human-robot applications for synchrony detection. The first and most obvious one is

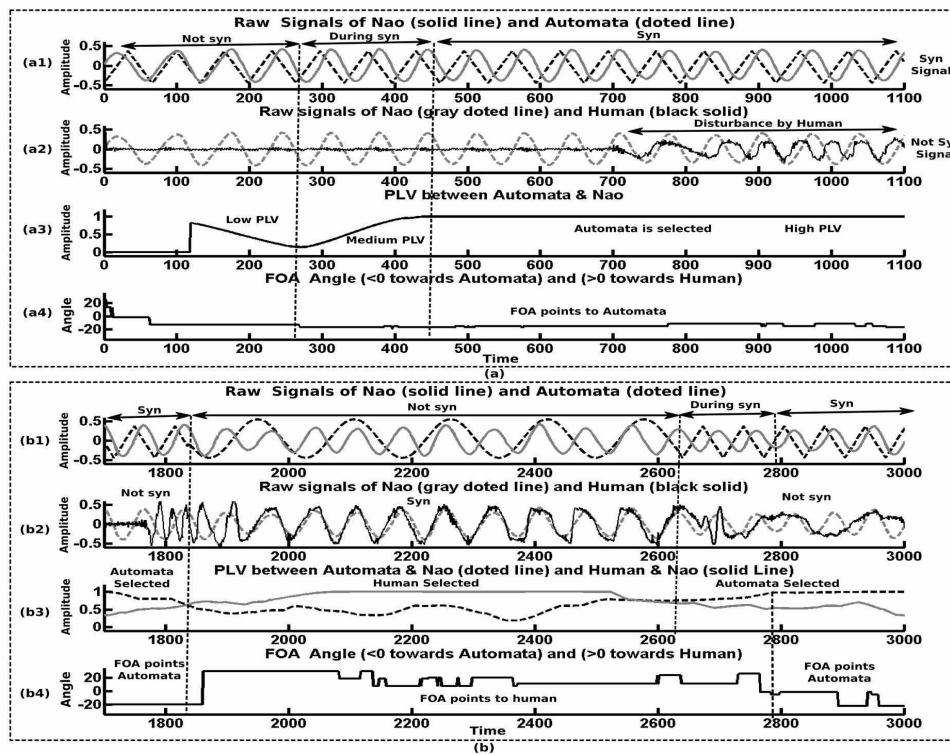


Fig. 5. Results: (a) shows start of experiment with single agent and then disturbed by the other agent. (b) Different frequency agents interact with Nao.

to extend the model to learn, in a developmental way, more complex interactions (complex gestures). Indeed, synchrony detection and selection of partner permit to maintain interaction with a partner moving synchronously with the robot in terms of low fundamental temporal frequency of interaction. As a result, more complex gestures (higher temporal frequencies) can be autonomously learnt by the robot while interacting and imitating the human partner. Similarly, we aim to use our architecture for navigation tasks. A mobile robot can choose a synchronous agent to interact with and consequently learn complex navigation tasks by keeping synchrony while moving with the selected partner. Finally and in a global point of view, we question the use of synchrony detection, focus of attention and selection of partner in turn-taking games during interaction. In fact, synchrony can not only be considered as a starting point for social interaction but also as a way to re-engage the interaction with a selected partner.

Acknowledgements—This work was supported by the “Interact” French project referenced ANR-09-CORD and the “Neuro-robot” French ANR project.

REFERENCES

- [1] A. Kendon. Conducting Interaction: Patterns of behaviors in focus encounters. Cambridge University press, Cambridge, UK. 1990.
- [2] J. Nadel, I. Carchon, C. Kervella, D. Marcelli, D. Reserbat. Expectancies for social contingency in 2 months old. *Developmental science* 2 (1999) 164-173.
- [3] J. Issartel, L. Marin, M. Cadopi. Unintended Interpersonal Co-ordination: Can we march to the beat of our own drum?. *Neuroscience Letters*, 411:174-179, 2007.
- [4] Z. Nda, E. Ravasz, T. Vicsek, Y. Brechet, and A.-L. Barabasi, The sound of many hands clapping, *Nature*, 403, 850 (2000).
- [5] Varlet M, Marin L, Raffard S, Schmidt RC, Capdevielle D, et al. (2012) Impairments of Social Motor Coordination in Schizophrenia. *PLoS ONE* 7(1): e29772. doi:10.1371/journal.pone.0029772
- [6] G.J. Stephens, L.J. Silbert, U Hasson. Speaker-listener neural coupling underlies successful communication. *Proc Natl Acad Sci USA* 2010;107(32):14425:30.
- [7] G. Dumas, J. Nadel, R. Soussignan, J. Martinerie, L. Garnero. Inter-Brain Synchronization during Social Interaction. *PLoS ONE* 2010;5(8):e12166.
- [8] A. Pikovsky, M. Rosenblum and J. Kurth, Synchronization. A Universal Concept in Nonlinear Science, Cambridge (2001).
- [9] P. Andry, A. Blandchard, P. Gaussier. Using the rhythm of non-verbal human-robot interaction as a signal for learning. *IEEE Transactions on Autonomous Mental Development* 3(1): 30-42 (2011).
- [10] K. Prepin and P. Gaussier. How an agent can detect and use Synchrony parameter of its own interaction with a human. *COST 2102 Training School 2009*: 50-65.
- [11] A. Blanchard and L. Caamero. Using visual velocity detection to achieve synchronization. In Y. Demiris (Ed.), *AISB05 third International Symposium on imitation in Animals and Artifacts*. Hatfield UK. 2005.
- [12] L. Marin, J. Issartel and T. Chaminade. Interpersonal motor coordination, from humanhuman to humanrobot interactions. In: Dautenhahn, Kerstin (ed.), *Robots in the Wild: Exploring human-robot interaction in naturalistic environments: Special Issue of Interaction Studies* 10:3 (2009).
- [13] J. P Lachaux, E. Rodriguez, J. Martinerie and F.J. Varela. Measuring phase synchrony in brain signals. *Human Brain Mapping* (1999).
- [14] A. Revel, P. Andry : Emergence of Structured Interactions: From theoretical model to pragmatic robotics. *Neural Networks*, vol. 22, no. 2, pp. 116–125 - 2009.
- [15] J. Nagumo. A learning method for system identification. *IEEE Trans. Autom. Control* 12(3) (1967) 282-287 NLMS reference.
- [16] P. Rochat. Ego functions of early imitation. In A. Meltzoff & W. Prinz (Eds.), *The imitative mind* (pp. 85-97). Cambridge University Press, 2002.